# Study of metabolic interactions between plant and xylem-colonizing bacteria using constraint-based modeling

Léo Gerlin

▶ **To cite this version:**

Léo Gerlin. Study of metabolic interactions between plant and xylem-colonizing bacteria using constraint-based modeling. Life Sciences [q-bio]. Université Paul Sabatier - Toulouse III, 2021. English. ⟨NNT : 2021TOU30177⟩. ⟨tel-04839734⟩

**HAL Id: tel-04839734**

**https://hal.inrae.fr/tel-04839734v1**

Submitted on 1 Jan 2026

**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.

Distributed under a Creative Commons CC0 1.0 - Universal - International License


# THÈSE

## En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

### Délivré par l'Université Toulouse 3 - Paul Sabatier

---

**Présentée et soutenue par**

**Léo GERLIN**

Le 21 octobre 2021

## Etude des interactions métaboliques entre plante et bactéries colonisatrices du xylème en utilisant la modélisation par contraintes

---

Ecole doctorale : **SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingenieries**

Spécialité : **Interactions plantes-microorganismes**

Unité de recherche :
**LIPME - Laboratoire des Interactions Plantes-Microbes-Environnement**

Thèse dirigée par
**Stephane GENIN et Caroline BAROUKH**

Jury

**Mme Alia DELLAGI,** Rapporteure
**M. Hubert CHARLES,** Rapporteur
**M. Matthieu ARLAT,** Examinateur
**Mme Anne GOELZER,** Examinatrice
**M. Stéphane GENIN,** Directeur de thèse
**Mme Caroline BAROUKH,** Co-directrice de thèse

# Study of metabolic interactions between plant and xylem-colonizing bacteria using constraint-based modeling

Léo Gerlin

**Supervisors**: Stéphane Genin, Caroline Baroukh

**Keywords**: modeling, xylem, *Ralstonia solanacearum*, metabolism, tomato, *Xylella fastidiosa*

**Defense**: 21 October 2021

**Laboratoire des Interactions Plantes-Microbes-Environnement** (LIPME)

**Université Toulouse III – Paul Sabatier**, Graduate school SEVAB

# Remerciements

Je remercie mes rapporteur·e·s Alia Dellagi et Hubert Charles, qui ont accepté d'évaluer ce manuscrit, ainsi que les membres de mon jury Anne Goelzer et Matthieu Arlat.

Je remercie également l'Université Toulouse III – Paul Sabatier, qui a financé ces travaux de thèse.

Je tiens à adresser des remerciements très chaleureux à la directrice et au directeur de thèse que j'ai eu la chance immense d'avoir, Caroline Baroukh et Stéphane Genin. Merci Caroline de m'avoir fait confiance pour un stage d'abord, puis pour cette thèse. J'ai beaucoup aimé travailler avec toi, pour tes compétences à la fois en métabolisme et modélisation, mais aussi et surtout pour ta patience, ta confiance, ta disponibilité et ton implication. Merci aussi de m'avoir incité à parfois souffler, peu de doctorants ont la chance d'avoir ces conseils-là ! J'espère que d'autres doctorants auront aussi la chance de travailler sous ton encadrement. Merci également à Stéphane, pour ton regard lui aussi toujours très bienveillant et encourageant sur mes travaux et mes avancées, merci d'avoir toujours été disponible pour répondre à mes interrogations, toujours avec calme et ouverture d'esprit.

Deux autres personnes ont grandement participé aux travaux présentés ici, il s'agit de Ludovic Cottret et Antoine Escourrou. Merci à Ludovic Cottret, expert des réseaux métabos, du SBML, de la conversion de leurs identifiants… la liste est longue ! Merci de m'avoir toujours débloqué quand j'avais besoin d'aide, et d'avoir participé à la construction des réseaux de la tomate et de *Xylella*. Merci à Antoine Escourrou, avec qui j'ai passé de nombreuses heures à découper des plants de tomates, à attendre que la sève remonte, et la prélever quand on était chanceux ! Merci pour ton implication et tes nombreux aller-retour à MetaToul et même à Bordeaux pour ce projet. Plus largement, je remercie toutes les personnes participant au bon fonctionnement des services et plateformes communes du LIPME (informatique, laverie, TPMP, gestion…) qui ont été bien sûr essentielles à la réalisation de mes travaux. Je remercie aussi l'IRHS d'Angers avec qui j'ai collaboré sur l'étude de *Xylella fastidiosa*, et en particulier Géraldine Taghouti et Sophie Cesbron. Bravo pour votre ténacité à réaliser des Biolog sur cette bactérie si fastidieuse, qui nous a donné de si précieuses données !

Je tiens à remercier les personnes qui ont suivi mes travaux lors de mes comités de thèse, qui ont été des lieux d'échange instructifs et encourageants : Sophie Colombié, Sylvain Prigent, Zhanwu Dai, Olivier Bernard, Alice Boulanger et Laurent Noel. En particulier, merci Sylvain pour l'invitation à la rédaction d'un chapitre, et merci Sophie pour tes précieux retours sur la construction et validation du modèle VYTOP. J'espère avoir d'autres occasions d'échanger avec vous par la suite.

J'ai aussi eu la chance d'enseigner durant cette thèse, en 1ère année à l'INSA Toulouse et en master 2 Bioinformatique à l'Université Paul Sabatier. Je remercie les équipes pédagogiques de ces deux formations pour m'avoir accordé leur confiance, enseigner a été un vrai plaisir.

Je remercie toutes les personnes du labo avec qui j'ai pu échanger et partager un thé l'après-midi, un repas ou une bière le soir, et qui ont rendu plus humaines ces trois années de science.

Ensuite, je tiens à remercier mes ami·e·s, ceux avec qui j'ai râlé et stressé sur mon projet ou ceux qui m'ont vu peu pendant ces trois années mais qui ne m'en ont pas voulu ! Mentions spéciales à Anna, Laurie, Adrian et Boris.

Je remercie toute ma famille, qui a toujours été si soutenante et fière de mon parcours. Merci tout particulièrement à mes parents. Vous m'avez rendu très curieux, et vous m'avez fait aimer la lecture et l'écriture, sans que vous vous en rendiez compte cela a été trois atouts essentiels pour mener à bien une thèse.

Pour finir, les plus grands remerciements vont bien sûr à Mallaury. Cela fera bientôt dix ans que nous nous soutenons dans tous les moments de notre vie, que nous vivons ensemble les joies et les déceptions de nos parcours. Tu as été un soutien essentiel à toutes les étapes. Cela fait aussi presque dix ans que je sais que je veux faire une thèse « en biologie des systèmes » (avant même de vraiment comprendre ce que ça voulait dire…), et qu'ensemble on a franchi toutes les étapes pour finalement arriver jusqu'à la fin de celle-ci. On l'a fait !

Je dédie cette thèse à Colin, mon fils, né quelques mois seulement avant la finalisation de ce manuscrit.

*Oh dear! I'd nearly forgotten that I've got to grow up again! Let me see—how is it to be managed?*
*I suppose I ought to eat or drink something or other; but the great question is, what?*
Lewis Carroll, Alice's Adventures in Wonderland

# Abbreviations

| | |
|---|---|
| **dFBA** | dynamic Flux Balance Analysis |
| **EC** | Enzyme Classification |
| **EFM** | Elementary Flux Mode |
| **EPS** | exopolysaccharides |
| **ETI** | effector-triggered immunity |
| **FBA** | Flux Balance Analysis |
| **FBPase** | fructose-1,6-bisphosphatase |
| **FCA** | Flux Coupling Analysis |
| **FVA** | Flux Variability Analysis |
| **FW** | fresh weight |
| **GPR** | Gene-Protein-Reaction |
| **HR** | hypersensitive response |
| **Hrp** | Hypersensitive response and pathogenicity |
| **LPS** | lipopolysaccharide |
| **NGAM** | Non-Growth-Associated Maintenance |
| **PFK** | phosphofructokinase |
| **PTI** | pattern-triggered immunity |
| **QSSA** | Quasi-Steady-State Approximation |
| **RSSC** | *Ralstonia solanacearum* Species Complex |
| **SBML** | Systems Biology Markup Language |
| **T3E** | type-three effector |
| **T3SS** | type-three secretion system |
| **TAL** | transaldolase |
| **VYTOP** | Virtual Young TOmato Plant |
| ***Xcc*** | *Xanthomonas campestris* pv. *campestris* |

# Figures

# Tables

# Contents

# Chapter 1

# Bacterial plant pathogens: from crop health to pathogenicity and metabolism

## 1.1 Plant health and pathogens

According to the last FAO report (FAO et al., 2020), the number of undernourished people in the world is still increasing, and reached 687.8 million in 2019. Food security is one of the lever suggested to reduce this urgent issue. In particular, biotic agents altering plant health, divided into plant pathogens and pests, cause crop losses up to 40%, according to the FAO, so they could heavily disrupt food security.

The threat of these biotic agents is complex to prevent as it increases with the growth of human activity and economy. The birth of plant pathology is traditionally associated with the 19$^{th}$ century and the Irish Potato Famine. At this time, Europe was transitioning toward an industrial society and facing an unprecedented population growth. In Ireland, the issue of feeding the population led to favor the monoculture of a unique potato variety, but caused the dramatic outbreak of the late blight disease caused by the oomycete *Phytophtora infestans* (Schumann & D'Arcy, 2017). More broadly, human activity is responsible of biodiversity reduction favoring the evolution of pathogens, and of globalization that introduces biotic agents into new areas (Lucas, 2020). There are also recent evidences that global warming will be profitable for biotic agents spread as temperature increase deactivates several plant resistance mechanisms (Desaint et al., 2021). This persistent and concerning threat of plant pathogens and pests, and their complex relationship with human activity, incite us to look closely at plant health and plant interactions with biotic agents.

The concept of diseased plant is described in opposition to “normal” plant, meaning that a plant is diseased when it deviated from its normal (“healthy”) state, leading to the appearance of symptoms such as wilting or tissue necrosis. This deviation is complex to formalize as there is an inherent stochastic variation in a population (Agrios, 2005), and the boundary between diseased and healthy plant remains relative and context-dependent. Infectious diseases are caused by pathogens, which are mainly microorganisms: fungi, bacteria, viruses, protozoa, oomycetes, but include also nematodes and eventually parasitic higher plants (sometimes also ranged into plant pests). Classically, the “amount of an infectious disease” in a plant is described to be due to the interaction between three components: the pathogen (more or less virulent, active, abundant), the host plant (age, susceptibility) and the environment (condition favoring or not the disease), forming the disease triangle (Agrios, 2005).

The definition of a pathogen (Méthot & Alizon, 2014) is even more difficult to establish. In the 19$^{th}$ century, pathogens were entities distinct from non-pathogenic organisms, possessing specifically the ability to cause disease. Advances in genomics, post-genomics and ecology contradicted this binary view. It appeared impossible to clearly point what factors distinguished a pathogenic from a non-pathogenic organism. The concept of “virulence genes” emerged in the 20$^{th}$ century, notably for plant pathogens, but failed to be based on rationale criteria, as organisms with none of a selected set of virulence genes could cause disease, and conversely, organisms with a set of virulence genes could remain harmless in a given context. The diversity of host responses also makes impossible this binary distribution. Some hosts will remain asymptomatic even with a high load of an organism described as a pathogen, while a “harmless” organism will induce in one case an exaggerated host immune response that will be seen as a disease, or will provoke disease in association with another organism. Finally, the boundary between pathogenic and non-pathogenic organisms is challenged by the consideration of evolution in pathogenesis: a harmless organism, in particular a bacteria, can evolve during an infection or during its spread in a population, becoming responsible of a severe disease. Hence, a pathogen can be seen as an organism able to cause virulence (decrease the host fitness) upon infection (Méthot & Alizon, 2014). This categorization is not fixed but is a transitory outcome of the interaction between the genotype and phenotype of the pathogen, the host genotype and phenotype plus the environment (abiotic factors as well as other interacting organisms).

Plant pathogens provoke disease by the sequence of different events in which they will interact with the plant host or its environment, named the disease cycle. The first step is the **inoculation**: the pathogen is in contact with the plant, transmitted through wind, water or insect for example. Then, the pathogen enters into the host through natural openings or wounds: it is the **penetration** step, made possible by a host recognition and attachment. The third step is the **infection**: the pathogen invades the tissues and colonizes them. The symptoms can become visible at this step, or appear only at a later stage of plant development or at a change in environmental conditions. The last step of the disease cycle is **dissemination**: the pathogen, after growth in the plant, will disseminate in the environment to reach other hosts. Additionally, a step of **survival** can also occur in the absence of the host, for example at seasonal period of the year for which it is not actively growing. Pathogens can have the ability to survive in environments such as fruiting bodies, plant debris, insects or seeds, or to remain in the soil and water.

## 1.2 Pathogenicity determinants of bacterial plant pathogens

Bacteria are unicellular microorganisms deprived of nucleus and organelles. About 100 species of bacteria (Agrios, 2005) are known to cause diseases in plants. Bacterial plant pathogens originate from diverse branches of the bacterial kingdom. Among the most scientifically and economically important bacterial plant pathogens presented by Mansfield et al. (2012), we can find γ-Proteobacteria (*Pseudomonas syringae*, *Xylella fastidiosa,* some *Xanthomonas* spp., some *Dickeya* spp. and *Erwinia amylovora*), β-Proteobacteria (*Ralstonia solanacearum*), α-Proteobacteria (*Agrobacterium tumefactiens*). While these are only gram negative and Proteobacteria, other important plant pathogens such as *Clavibacter michiganensis* and *Streptomyces scabies* are gram-positive and belong to the Actinobacteria phylum. Bellow their apparent extreme simplicity and their size (usually less than 2 µm of diameter), bacteria have developed very complex and diverse chemical and molecular weapons presented as pathogenicity determinants, which allows them to infect and colonize environments such as plants (**Figure 1-1**).

### 1.2.1 Motility and attachment systems

To penetrate into the host, bacterial pathogens use motility and attachment systems. In several *Xanthomonas* spp., proteinaceous cell-surface components named **adhesins** are involved in bacteria attachment to leaves or seeds (Darsonval et al., 2009; Das et al., 2009; Gottig et al., 2009; Ray et al., 2002). In *Ralstonia solanacearum*, twitching and swimming motility are respectively driven by **type IV pilus** and **rotating flagella**, having different roles in plant colonization (Corral et al., 2020). The external membrane of gram-negative bacteria includes **lipopolysaccharide (LPS)**, a complex macromolecule combining a lipidic part and a polysaccharide with often species or strain specific motifs. LPS plays**G** a role in cell-cell adhesion in a biofilm, as described in *X. fastidiosa* (Rapicavoli et al., 2018). Roles of LPS in virulence were also unraveled in *Xanthomonas* spp. (reviewed in Büttner & Bonas, 2010), with notably a modulation of its structure to evade plant recognition.

### 1.2.2 Cell-wall degrading enzymes

The type-two secretion system of bacteria allows the release of extracellular proteins into the extracellular environment, including a diverse panel of **cell-wall degrading enzymes** such as lipases and polygalacturonases. The importance of this system in virulence has been described in *X. campestris* pv. *campestris* (Gough et al., 1988), *X. oryzae* pv. *oryzae* (Jha et al., 2007), *R. solanacearum* (Liu et al., 2005) and *X. fastidiosa* (Nascimento et al., 2016).

### 1.2.3 Exopolysaccharides

**Exopolysaccharides (EPS)** are sugar polymers excreted by bacteria. They are involved in biofilm formation and protection against stresses. EPS of *Xanthomonas* spp. are named xanthan gum (Chou et al., 1997), while the close polymer produced by *X. fastidiosa* is named fastidian gum (Da Silva et al., 2001). Several other plant pathogens - if not all - secrete this polymer (*Erwinia, Agrobacterium, Clavibacter, Ralstonia, Pseudomonas),* as reviewed by T. P Denny (1995). For xylem-colonizing pathogens, these polymers are also suspected to contribute to reduce the water flow, potentially leading to the observed wilting symptoms they provoke.

### 1.2.4 Type-three effectors and type-three secretion system

**Type-three effectors (T3Es)** are essential pathogenicity determinants of several bacterial plant pathogens. T3Es are proteins produced by the pathogen injected

directly inside the plant cell by the **type-three secretion system (T3SS)**, which acts as a molecular syringe. Some bacteria possess a vast repertory of effectors, reaching around 70 in *R. solanacearum* strains (Deslandes & Genin, 2014; Landry et al., 2020). Some of them are very strain or species-specific while other are shared between diverse families of pathogens. Several effectors are able to suppress plant immune responses, while others act for example on plant metabolic pathways. The functional characterization of effectors is particularly challenging as there are functional redundancies between them, giving robustness to this pathogenicity determinant. Curiously, some rare bacterial pathogens miss this pathogenicity determinant, including *X. fastidiosa*, while they are largely conserved among its phylogenic family (*Xanthomonodaceae*) (Simpson et al., 2000).

### 1.2.5 Secondary metabolites

In numerous bacteria, pathogenicity and interaction with the plant is supported by the production of **secondary metabolites**. *Pseudomonas syringae* produces diverse molecules named phytotoxins, such as coronatine, that can interfere with plant metabolic pathways and contributes importantly to virulence (Ishiga, 2017). Bacterial pathogens also secrete plant hormones such as ethylene (Valls et al., 2006), which could be used to modulate the plant physiology at the advantage of the pathogen. Bacterial activity in a population is coordinated by quorum sensing, a sophisticated inter-bacteria communication system. *R. solanacearum* produces a specific fatty acid derivate as quorum sensing inducer: 3OH-PAME (3-hydroxy palmitic acid methyl ester) or 3OH-MAME (3-hydroxy myristic acid methyl ester) depending on the strain (Kai et al., 2015). At high density of 3OH PAME/MAME, the master regulator PhcA (part of the "Phenotypic conversion" system) will induce a strong phenotypic switch in *R. solanacearum*, leading to the production of several pathogenicity determinants such EPS as well as activating some physiological and metabolic behaviors, and repressing others. This makes the bacteria able to sense the level of its population density and adapt its phenotype accordingly (Clough et al., 1997; Khokhani et al., 2017).

### 1.2.6 Tumorigenic DNA

The **tumorigenic DNA** is a rare pathogenicity determinant of plant pathogens. This bacterial DNA is transferred from *Agrobacterium tumefaciens* into the plant cell through a type-four secretion system and is then integrated into the plant DNA. The

tumorigenic DNA induces the overproduction of plant hormones such as auxin, leading to the formation of a vegetal tumor, and activates the synthesis of opines, used as exclusive carbon sources by *A. tumefaciens* (Escobar & Dandekar, 2003; Tzfira & Citovsky, 2002).

## 1.3 Plant immune response to a pathogen

In response to an infection, plants have acquired molecular mechanisms to sense and respond to the presence of a pathogen. Two main mechanisms exist: **PTI (pattern-triggered immunity)** and **ETI (effector-triggered immunity)** (M. Yuan et al., 2021). Molecular elements indicative of the presence of a pathogen such as LPS or cell wall degrading enzymes (called PAMP for pathogen-associated molecular patterns) as well molecular witness of damages (called DAMP for damage-associated molecular patterns) can be sensed by receptors at the surface of the plant cell (PRRs, pattern recognition receptors) (Thomma et al., 2011). These receptors activate PTI. PTI is described as low-level and basal immunity, protecting only the plant from poorly adapted pathogen, while adapted pathogens can use effectors to deactivate PTI. Inside the plant cell, effectors can be detected by NLR (nucleotide-binding domain leucine-rich repeat) receptors, activating ETI, which is a more amplified and intense defense line against pathogens that can restore silenced PTI. PTI and ETI are based on different upstream receptors and molecular cascades, but the two immunities will lead to similar downstream outputs such as calcium flux, reactive oxygen species (ROS) burst and phytohormone signaling (Thomma et al., 2011). Ultimately, plant immunity can lead to hypersensitive response (HR), a programmed and localized cell death at the site of the pathogen entry, blocking its spread into the plant. Hence, there is an interplay in plant – pathogen interaction between activation of immunity at the detection of the pathogen through PTI, deactivation of this mechanism by some effectors, while effectors could also activate a new immune response (through ETI). This sequence of activations/deactivations of immunity is classically described as a “zigzag model” (Thomma et al., 2011). The final outcome of this molecular battle will determine whether the plant – pathogen interaction is compatible, leading to plant disease, or incompatible, leading to plant full resistance, notably through HR (Ponzio et al., 2016).

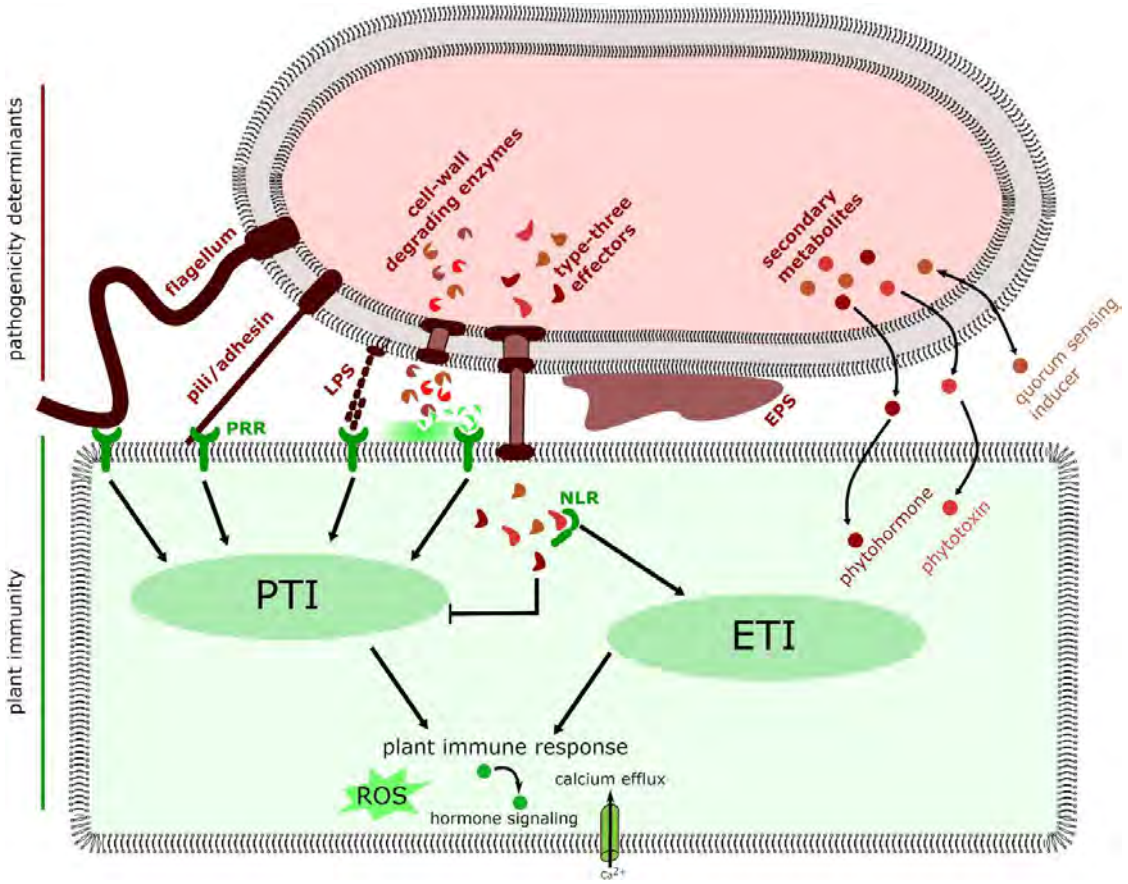



**Figure 1-1: Main pathogenicity determinants of bacterial plant pathogens, and their induction of plant immune response.**

PTI: pattern-triggered immunity, ETI: effector-triggered immunity, LPS: lipopolysaccharide, EPS: exopolysaccharide, NLR: nucleotide-binding domain leucine-rich repeat, PRR: pattern recognition receptors.

## 1.4 Metabolism in plant – bacterial pathogen interactions

### 1.4.1 Nutrition and resource allocation for virulence and immunity

Added to the vast diversity of molecular weapons described in the previous section, bacterial pathogens have the ability to proliferate very quickly compared to other living organisms, giving them important advantages in host colonization. *In vitro,* rapid proliferation is evidently due to the use of artificial media containing an abundant and optimal amount of molecules sustaining growth. To grow rapidly *in planta*, bacteria must acquire in their environment (plant tissues) sufficient amounts of building blocks (organic matter) and energy requirements. Additionally, several pathogenicity

determinants, such as EPS, also require important amount of organic matter and depend on the ability of the organism to assimilate resources in its environment.

Hence, it was demonstrated that an avirulent mutant of *R. solanacearum*, which does not synthesize virulence components, has an enhanced growth rate (198% higher) as all the resources assimilated can be used to proliferate (Peyraud et al., 2016). Adaptation of primary metabolism to the plant environment, ensuring resource assimilation, energy generation and synthesis of growth/virulence components appears to be a key element of the infection success in plant – bacterial pathogen interactions. On the plant side, the infection by a bacterium reciprocally influences its primary metabolism. The immune response is energy/matter consuming and a part of energy/matter will be deviated from their usual physiological purpose to the defense against a pathogen. Thus, there is a trade-off on the plant side to distribute the matter and energy between defense (through immunity) and growth/maintenance (Herms & Mattson, 1992; Pieterse et al., 2012; Walters & Heil, 2007). Furthermore, infection of plant tissues by pathogens also directly alters plant metabolism: the colonization of leaf surface affects photosynthesis and organic matter production, while diseased roots reduce the availability of soil nutrients, and infecting vascular tissues of the plant alters the nutrient exchanges between plant tissues. Dependence of pathogens on plant resources, and alterations of plant metabolism by pathogens lead to a paradox: bacterial pathogens need to drain plant resources to take the upper hand of the battle, but conversely, infection alters plant metabolism and will reduce the availability of nutrients for the pathogen. Thus, similarly to the thorough resources repartition between immunity and growth/maintenance that operates in the plant, the pathogen must accurately regulate a control between too drastic plant damages that will provoke too poor resource availabilities, and a too weak pathogenicity that will enable large resource availability but limit the chances to overpass plant immunity. Beyond just providing secondary metabolites, we can here highlight that metabolism, and in particular primary metabolism is tightly connected to virulence and its control on both plant and bacterium side will be determinant in the outcome of the interaction (**Figure 1-2**). This importance of the “quest for food” in host-pathogen interaction is described as “nutritional virulence” (Kwaik & Bumann, 2013) and is increasingly documented for human/animal pathogens. However, in the context of plant pathogens, it has been discussed in Berger et al. (2007) but remains only sparsely studied.

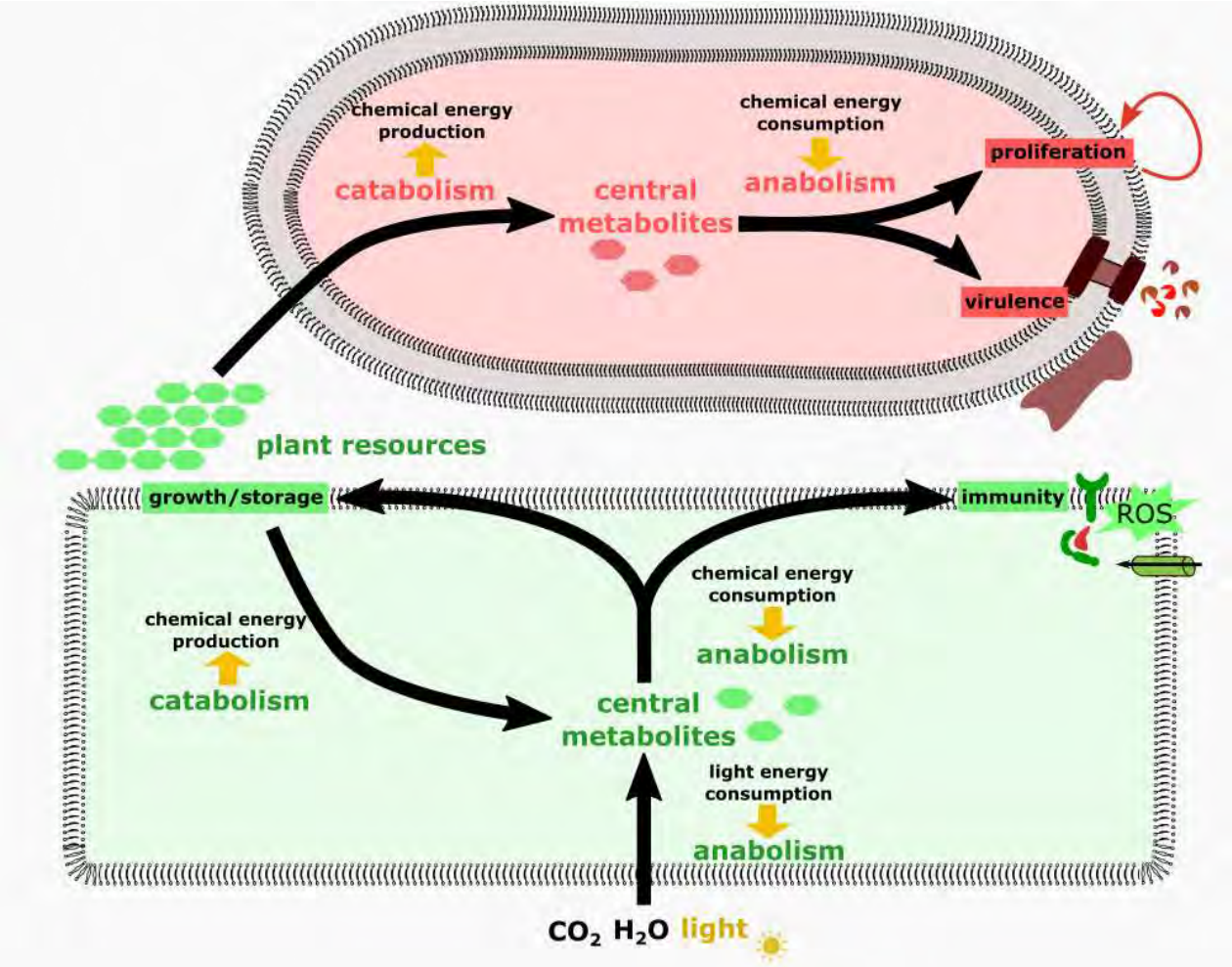



**Figure 1-2: Representation of matter/energy exchanges through metabolism in plant - bacterial pathogen interactions.**

### 1.4.2 Virulence-related elements control plant and bacterial metabolism

As plant and pathogen must regulate their primary metabolism during the interaction, we can imagine that molecular mechanisms in plant and bacteria exert a control on metabolism. Several studies showed the increasing importance of metabolic control in plant – pathogenic bacteria interaction. Plener et al. (2012) demonstrated that *R. solanacearum* switches between methionine synthase enzymes under the control of a key pathogenicity regulator or in the presence of plant cells. At a larger scale, reconstruction of virulence regulatory network of this pathogen (Peyraud et al., 2018) demonstrated that this network exerts a strong control on metabolic genes, activating functionally redundant pathways and providing robustness during plant infection. Still on *R. solanacearum*, PhcA, the master regulator of virulence activated by quorum sensing, exerts a strong control on metabolism as it represses the assimilation of 17

carbon sources (Peyraud et al., 2016). In several plant pathogens, T3Es injected in plant cells can interfere on plant metabolism. They can manipulate plant cells sugar export (reviewed in Timilsina et al., 2020 for *Xanthomonas* spp.) and plant polyamine level (Wu et al., 2019), promote GABA (gamma amino butyric acid) production (Xian et al., 2020) and either interact with metabolic enzyme such as pyruvate decarboxylase (Y. Wang et al., 2021) or be directly a metabolic enzyme (Poueymiro et al., 2014). Similarly to T3Es, in *P. syringae*, several phytotoxins secreted disrupt plant amino acid and polyamine metabolism (Ferguson & Johnston, 1980; Turner & Taha, 1984; Uchytil & Durbin, 1980). The global picture of these plant metabolic alterations and control of bacteria metabolic genes is still lacking and the conclusion of these studies remains often elusive. Indeed, it is hard to determine if plant metabolism alterations disturb immune signaling pathways or supports an increase of nutrient availability to favor the pathogen growth. Metabolism, even only primary metabolism, is composed of around 2000 reactions linked together (see Chapter 2), both in bacteria and in plants. Just analyzing the effects of one reaction is often insufficient to have an overall understanding of its impact on phenotype. This justifies the importance of using system-level approaches to understand metabolism (see Chapter 2).

## 1.5 Xylem-colonizing bacterial pathogens

In the top-ten list of the most scientifically and economically important bacterial plant pathogens (Mansfield et al., 2012), seven of them are colonize the xylem: *R. solanacearum*, *Xanthomonas oryzae*, *Xanthomonas campestris*, *Xanthomonas axonopodis*, *E. amylovora*, *X. fastidiosa*, *Pectobacterium carotovorum*. It is also the case of bacteria from the *Clavibacter*, *Curtobacterium* and *Pantoea* genera (Yadeta & Thomma, 2013). In the following section, we will describe this plant tissue and link it to the important bacterial colonization observed. Then, we will use plant xylem and xylem-colonizing bacteria as model systems for further analyses of the metabolism in plant – pathogen interactions in the other chapters of this thesis.

### 1.5.1 Xylem structure

Xylem is a part of the plant vascular system. It ensures the transport of water and minerals from the soil to the aerial parts of the plant, but also contributes to maintain the plant architecture by strengthening organs (Myburg et al., 2013). Xylem is formed by dead tubular structures allowing water and mineral flow, named tracheary elements

(Myburg et al., 2013). Pairs of small lateral openings named pits connect these dead structures. A pit membrane of primary cell wall with a certain porosity allows water and mineral flow between tracheary elements (Taiz & Zeiger, 2006). Two types of tracheary elements exist: vessels and tracheids. Vessels are wider but shorter elements, while tracheids are very long vascular structures. Vessel structures are stacked end to end, and perforation plates between stacked vessel structures transmit water flow and hence form a continuous vessel (Taiz & Zeiger, 2006) (**Figure 1-3**). In addition to dead tracheary elements, xylem includes living parenchyma cells. These cells are able to store water, minerals and carbohydrates to balance supply and demand. Also, as mechanical support and protection against pathogens and herbivores, xylem structure includes fibers (Myburg et al., 2013).

In xylem, the water and mineral transport upward is driven by i) root osmotic pressure, ii) capillary action due to the hydrophobic property of xylem surface, iii) water cohesion allowing the formation of a column of water, iv) transpiration, creating a negative water potential at the level of leaf stomata where water is evaporated (Myburg et al., 2013). Transpiration is predicted to be the major driving force of xylem flow.

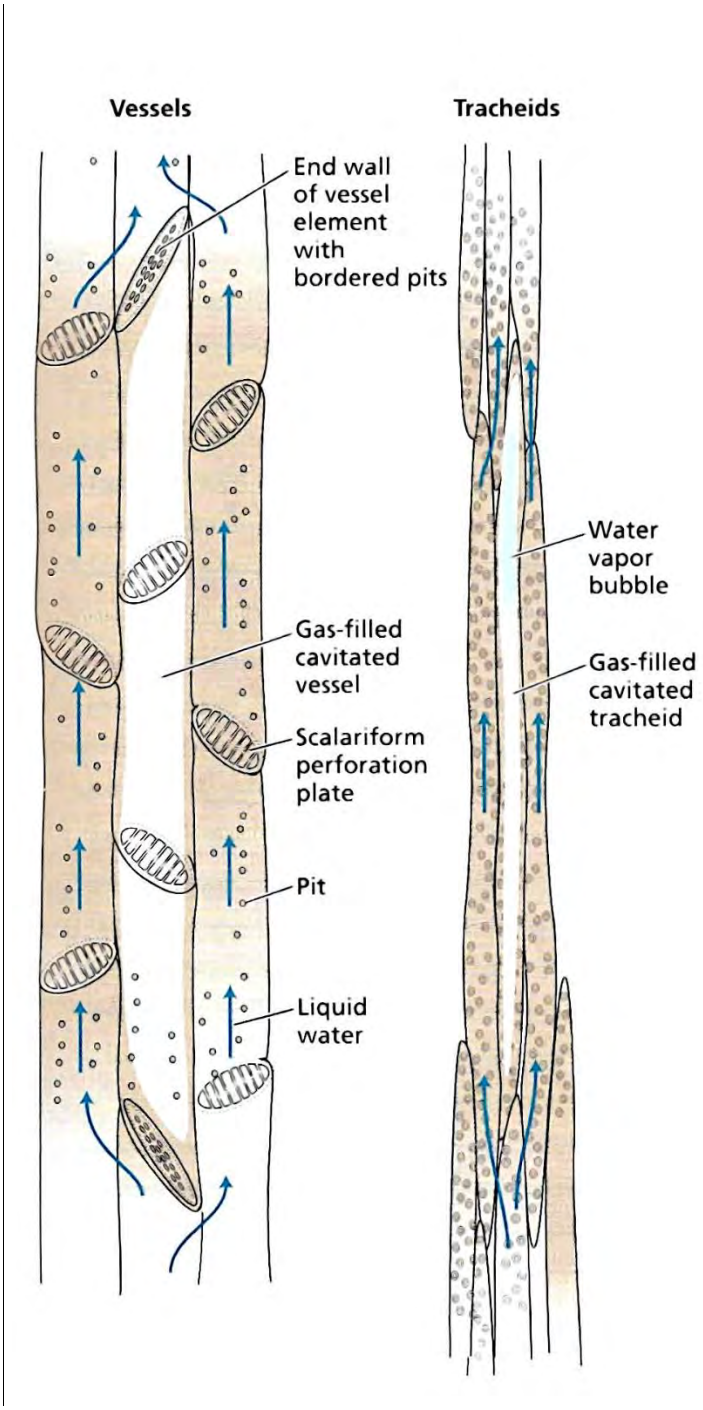



**Figure 1-3: Xylem structure and water transport.**

Interconnected tracheary elements (vessels and tracheids) allow water and nutrients transport to higher areas of the plant. Pits (present in both tracheids and vessels) and perforation plate (present in vessels only) allow water movement all along the plant. Tracheary elements can be filled with gas (embolism) and will not be able to transport water flow. The cell wall between pits prevents the spread of embolism to the whole plant. Figure taken from Taiz & Zeiger (2006).

### 1.5.2 Plant – xylem-colonizing bacteria interactions

A complete review and discussions of xylem in plant – pathogen interactions has been written by Yadeta and Thomma (2013). Briefly, xylem-colonizing bacteria can be present in soil, seeds or in insects. Some enter into the plant through wounds or natural openings, and will move into the plant to reach xylem vessels and colonize it, while other reach directly xylem vessels after transmission by xylem-sap feeding insects. Once xylem is reached, the pathogens will proliferate in vessels and spread from a vessel into another, degrading cell walls and pit membranes. The combination of plant debris, bacteria and production of biofilm will reduce water transport and induce stress and physiological alterations close to drought stress. These pathogens are thus often called “vascular wilt pathogens”.

After pathogen detection, plant can use different responses to limit the spread of the bacteria in xylem and eradicate it. Parenchyma cells release outgrowths named tyloses that block the pathogen spread, completed with gels and gums. However, this action of blocking xylem vessels could also contribute to water limitation and the induced stresses and symptoms. Plant also process to vascular coating around an infected zone, using notably callose deposition and wall swelling. As described before, HR, here named vascular HR, is also often at the origin of resistance against pathogens.

The adaptation of several bacterial pathogens to xylem could be explained by the limited presence of other microorganisms in this environment. Additionally, the low osmotic pressure compared to phloem, the vascular tissue transporting photosynthesis products, also contributes to the advantage of xylem colonization. However, xylem is described as a nutritionally poor environment. It contains abundant minerals such as nitrate, sulfate, magnesium but authors (Yadeta & Thomma, 2013) underline the lack of high amounts of carbon sources essential for bacterial growth. Yadeta and Thomma highlight the presence in low amounts of sugars, amino acids, organic acids but specify that this amount is fluctuating and is very condition dependent. They suggest that pathogens could be able to find nutritive sources by digesting cell walls or by inducing the release of carbon sources in xylem. The ability of some bacterial pathogens to modify plant metabolism through effectors previously described (see 1.4.2) could be one of the strategies. However, these ideas remain simple hypotheses, and no experimentally validated mechanism explaining *in planta* proliferation of xylem-colonizing bacteria has been proposed so far.

### 1.5.3 Examples of xylem-colonizing bacteria

In the following chapters of the thesis, I will focus on the xylem-colonizing bacteria *R. solanacearum*. *Xanthomonas campestris* pv. *campestris* (*Xcc*) and *X. fastidiosa* will be used for comparison. These three bacteria colonize the same environment but have very different lifestyles, genome size and physiological properties (such as growth rate) (**Table 1-1**).

**Table 1-1: Differences in lifestyle, genome size and growth between the three xylem-colonizing bacteria.**

Genome size are from Salanoubat et al. (2002) for *R. solanacearum*, da Silva et al. (2002) for *Xcc*, Simpson et al. (2000) for *X. fastidiosa*. Qualification of growth as "fast", "intermediate", "slow" result from important differences in the growth rates. *R. solanacearum* strain GMI1000 grows at around 0.28 $h^{-1}$ (Peyraud et al., 2016) on minimal media glutamate and repeatedly above 0.1 $h^{-1}$ on diverse carbon sources (Baroukh et al., accepted). Growth rate of *X. campestris* (Schatschneider et al., 2013) is around 0.05 $h^{-1}$ on minimal media with sugar as carbon source. *X. fastidiosa* is repeatedly qualified as fastidious since its isolation (Wells et al., 1987), and we could estimate a growth rate around 0.007 $h^{-1}$ on data from Leite et al. (2004).

| | ***R. solanacearum*** | ***Xcc*** | ***X. fastidiosa*** |
|---|---|---|---|
| **Entry into plant tissues** | Root | Leaf | Stem (direct sting through insect) |
| **Environment outside the plant** | Soil and water | Leaf surface and seed | Insect foregut |
| **Growth** | Fast | Intermediate | Slow (fastidious) |
| **Genome size (Mb)** | 5.8 | 5.1 | 2.7 |

***Ralstonia solanacearum*** (Figure 1-4, A., B., C.)

The Gram-negative β-Proteobacteria *R. solanacearum* belongs to the Burkholderiaceae family. Due to the large variability of biochemical properties and host specificities between *R. solanacearum* strains, the bacterium is described as a species complex (RSSC, *Ralstonia solancearum* Species Complex), classified by phylogenetic analyses into four phylotypes (Fegan & Prior, 2005) and more recently (Safni et al., 2014) into three species: *R. solanacearum*, *R. pseudosolanacearum* and *R. syzygii*. These classifications allow recovering roughly the geographic origin of the strains, but the host range remains usually very broad and the link with phylogenetics hard to establish (for review, see Genin, 2010; Genin & Denny, 2012). RSSC is the causal agent of bacterial wilt, a disease affecting more than 450 species of plants from

more than 50 botanical families (Hayward, 1991), including the Solanaceae (tomato, eggplant) but many others species of agronomical importance such as groundnut (Boshou et al., 1998) and ginger (Yu et al., 2007). RSSC is also responsible of plant diseases not named classically as “bacterial wilt”, such as Moko disease on banana, brown rot of potato, or Granville wilt on tobacco (Hayward, 1991).

*R. solanacearum* can survive in the soil and water, adopting a saprophytic lifestyle (using dead organic matter to grow). In contact with plant roots, it enters through wounds or natural openings (e.g. zones of root elongation) (Vasse et al., 1995). Then, the bacterium infects the intercellular spaces of root cortex, and finally reaches xylem vessels (Vasse et al., 1995). Once inside xylem vessels, *R. solanacearum* colonizes the vascular system in whole by degrading cell walls and moving upwards. The bacterium also intensely proliferates, leading to more than $10^{10}$ cells per cm of stem (Saile et al., 1997). This high cellular density, the production of EPS and potentially some plant responses (e.g. induction of tylosis and embolism) (Lowe-Power, Khokhani, et al., 2018) are suspected to be responsible of an occlusion of xylem vessels, depriving the aerial parts of water and soil nutrients and leading to wilting. After plant death, the bacteria returns to the soil and starts a new saprophytic lifestyle until it encounters a new host.

The model strain that has been the most studied by various teams worldwide is the strain GMI1000 (formally *R. pseudosolanacearum* GMI1000), belonging to the phylotype I. Because of the extensive literature in which it is named *R. solanacearum* and as a recognized reference strain of the RSSC, it will be referred by convenience hereafter as *R. solanacearum*.

***Xanthomonas campestris* pv. *campestris* (*Xcc*)** (Figure 1-4, D., E., F.)

*X. campestris* is a Gram-negative γ-Proteobacteria, belonging to the Xanthomonadaceae family. This pathogen infects majorly plants from *Brassica* species, and is divided into three pathovars: *X. campestris* pv. *campestris* (*Xcc*), causal agent of black rot of crucifers, *X. campestris* pv. *raphani*, responsible for bacterial spot on crucifers, and *X. campestris* pv. *incanae*, causing bacterial blight on ornamental *Brassica* spp. (Fargier & Manceau, 2007; Roux et al., 2015). The *Xanthomonas* genus comprises plant pathogens with differences in host range for each species, such as *X. oryzae* (Niño-Liu et al., 2006), responsible of diseases in rice, *X. albilineans*, affecting species from Poaceae family (such as sugarcane) (Birch, 2001) and *X. citri*, affecting

*Citrus* species (Brunings & Gabriel, 2003). This association between species and narrow host range allows in some cases establishing correlations between host adaptation and phylogenetics (Ryan et al., 2011), while in RSSC, this link is blurred as many strains have a broad host range.

*Xcc* has an epiphytic life, during which it survives on plant debris and leaves (Schultz, 1986). It enters into plant tissues generally through hydathodes (Hugouvieux et al., 2007), which are pores at leaf margin rejecting water transported upward by the plant xylem. The bacteria are then able to reach plant vascular system and colonize the xylem, even invading surrounding tissues of the leaf (mesophyll) (Akimoto-Tomiyama et al., 2014; Samal & Chatterjee, 2019). This bacterial propagation blackens the vascular system and rots the tissues, leading to the "black rot" symptoms. It could even contaminate seeds, contributing to a very large dissemination of the pathogen (for review, see Vicente & Holub, 2013).

***Xylella fastidiosa*** (Figure 1-4, G., H., I.)

The Gram-negative γ-Proteobacteria *X. fastidiosa* is phylogenetically close to *Xcc*, being also a member of the *Xanthomonadaceae* family. The bacterium can be found in more than 300 plant species from 63 families (Rapicavoli et al., 2018). It is described as pathogenic in more than 100 plant species, but remains harmless in the others, adopting a commensal endophyte lifestyle (Chatterjee et al., 2008; Hopkins & Purcell, 2002). *X. fastidiosa* is classified into four subspecies that correlate with symptomatic host specificities (Scally et al., 2005; Schuenzel et al., 2005). *X. fastidiosa* ssp. *fastidiosa,* is responsible of diseases in grapevine and almond, such as Pierce's disease on grapevine. *X. fastidiosa* ssp. *multiplex* colonizes almond and peach. *X. fastidiosa* ssp. *sandyi* is only found in oleander, and *X. fastidiosa* ssp. *pauca* provokes diseases in citrus, coffee and olive tree, such as the Olive Quick Decline Syndrome in Italy (Giampetruzzi et al., 2015; Marcelletti & Scortichini, 2016).

*X. fastidiosa* is limited to two environments: xylem vessels and the foregut of xylem sap-feeding insects. In insect vectors, a bacterial biofilm structure allows the adhesion to the turbulent environment of the foregut and so the subsequent transmission to plants (Killiny & Almeida, 2014). Once in a susceptible host, the plant colonization can be described by a model proposed by Chatterjee et al. (2008): *X. fastidiosa* exists in an 'exploratory' motile state, moving from one xylem vessel into another using cell wall degrading enzymes and also going upwards, and in an adhesive state (insect-acquirable

form), which involves biofilm formation, production of EPS and adhesins to favor stickiness. Generally, infection is only detected weeks after the infection, as the bacteria remain at low cell density in several parts of the plant and only partially colonize the vascular system (Newman et al., 2003). The typical symptoms of *X. fastidiosa*-related disease are marginal leaf necrosis and leaf scorching. Fruit production is also progressively affected (desiccation, hardening or size reduction) and bacteria will eventually lead to an overall plant decline (Rapicavoli et al., 2018). Interestingly, *X. fastidiosa* lacks a major pathogenicity determinant shared among most of the bacterial plant pathogens, the T3SS and associated effector substrates (see 1.2.4) (Simpson et al., 2000). The molecular bases explaining its ability to cause severe disease among a broad range of plant host without this determinant are still not fully understood.

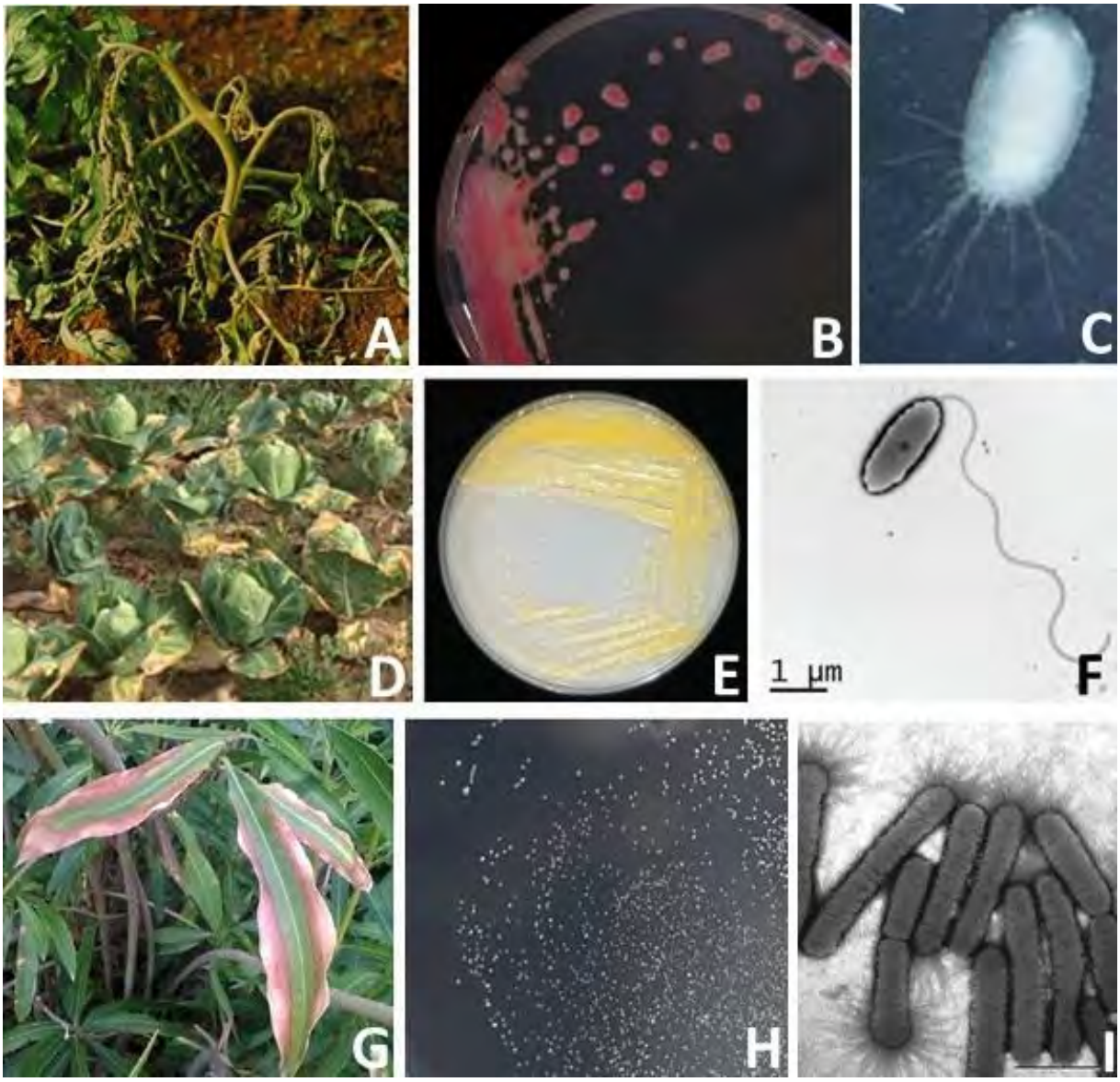



**Figure 1-4: Symptoms on plants (A, D, G), aspects on plates (B, E, H) and using microscopy (C, F, I) of respectively *R. solanacearum* (A, B, C), *X. campestris* pv. *campestris* (D, E, F), and *X. fastidiosa* (G, H, I).**

A. Wilting symptoms caused by *R. solanacearum* on tomato. B. Plating of *R. solancearum* (GMI1000) on solid medium supplemented with triphenyl – tetrazolium chloride and glucose. C. Transmission electron microscopy observation of *R. solanacearum* (GMI1000). D. Black rot symptoms caused by *Xcc* on cabbages. E. Plating of *Xcc* on Yeast Dextrose Calcium Carbonate medium. F. Electron microscopy observation of *Xcc*. G. Marginal leaf scorch symptoms caused by *Xylella fastidiosa* subsp. pauca on oleander. H. Plating of *Xylella fastidiosa* subsp. *fastidiosa* on BCYE medium. I. Transmission electron microscopy of *X. fastidiosa* (the bar represent 1 µm). Pictures from: A. Mansfield et al. (2012). B. A. Guidot, LIPME. C J. Vasse, LIPME. D., E., F. Vicente & Holub (2013). G., H. "PM 7/24 (2) *Xylella Fastidiosa*" (2016) I. Meng et al. (2005).

## 1.6 Objectives of the thesis

This state-of-start highlights that even though some molecular aspects of plant – pathogen interactions were unraveled, in both bacteria and host plants, metabolism is a non-negligible parameter of each interaction and lacks studies. Bacteria must efficiently find nutritive resource in poor and hostile environment such as xylem, and redirect matter and energy fluxes toward both proliferation and virulence. They deactivate plant physiology at their advantage but have still to maintain a basal physiology to have enough resources to pursue their colonization. On the other side, the whole plant must deal with distributing resources to immunity and still maintaining growth and maintenance process, the whole with increasing bacterial population affecting its physiology. Additionally, bacterial virulence is tightly connected with metabolism, and we presented several examples on how effectors, regulators, phytotoxins actively affected and redirected metabolism on the two sides of the interaction. Among the bacterial plant pathogens, several of them are xylem-colonizing bacteria, and use paradoxically a nutrient-poor environment to proliferate and spread inside the plant.

The general objective of this thesis, through the example of xylem-colonizing bacteria, is to unravel the nutrients used by plant pathogenic bacteria, and analyze how bacterial metabolism influences bacteria lifestyle and their interactions with the host. Our state-of-art revealed the complexity of having a global picture on metabolic interactions, as both plant and pathogen metabolism are intertwined, and gather a high number of biochemical reactions. Thus, to study metabolism in plant – pathogen interaction, I used systems biology approaches, and in particular genome-scale metabolic modeling.

Systems biology tools for plant – pathogen were already developed before the beginning of my thesis: the genome-scale metabolic model of *R. solanacearum*, one of the model organism of the laboratory, has been developed by Peyraud et al. (2016), and completed with a macromolecule module and a virulence regulatory network (Peyraud et al., 2018). The metabolic model was already calibrated *in vitro* with physiological data (Peyraud et al., 2016) for wild-type and some mutant strains. This calibration was pursued by Caroline Baroukh, who performed similar experiments on a wide diversity of carbon substrates (Baroukh et al., accepted) to better understand the trophic preferences *in planta*.

To pursue the previous systems biology tools and discoveries made in my laboratory before my arrival, my thesis focused on three main aspects, divided into four chapter (Chapter 3 to Chapter 6).

i) **How to model the environment of xylem-colonizing bacteria?** (Chapter 3)

The metabolic models already developed before the beginning of my PhD did not integrate the plant. I built a multi-organ metabolic model of tomato plant, with accurate growth and exchange fluxes. In particular, xylem fluxes has been represented in a precise manner, as it is the place of metabolic exchanges between plant and xylem-colonizing bacteria. Experiments were performed to calibrate quantitatively fluxes of matter, and the model was validated on independent data and use-cases. The model organism used in this part will be the tomato plant (*Solanum lycopersicum*), which is the natural host of the *R. solanacearum* GMI1000 strain used in the team.

ii) **How bacterial metabolism is sustained in xylem environment and affects plant physiology?** (Chapter 4 and Chapter 5)

Using the tomato plant metabolic model and the existing *R. solanacearum* metabolic model, I was able to build a plant – pathogen metabolic model. This model was used to study the plant fluxes required to sustain in planta bacterial growth, and to predict at which population density it affects plant physiology. To build a quantitative model, additional calibration was needed, including infected plant physiological measurements and xylem metabolomics. Chapter 4 presents these experimental calibrations and the new insights given on *R. solanacearum* and plant metabolic interaction. Chapter 5 will present the model developed using these calibrations.

iii) **Do xylem-colonizing bacteria have different metabolic capacities and metabolic strategies?** (Chapter 6)

After these advances on understanding the metabolic interaction between *R. solanacearum* and a tomato plant, the thesis focused on comparing *R. solanacearum* metabolic capacities with another xylem-colonizing pathogen, the bacterium *X. fastidiosa*. The genome-scale metabolic model of *X. fastidiosa* was reconstructed, and the same analysis tools was used on the two bacteria to compare their metabolic potentials. In particular, I tried to explain why these two bacteria have very different proliferation

capabilities: *R. solanacearum* proliferates spectacularly fast, while *X. fastidiosa* is called “fastidious” due to its very slow proliferation in vitro and in planta.

The Chapter 2 presents principles of genome-scale and constraint-based metabolic modeling to make the approach understandable for non-specialist readers, with an emphasis on previous existing metabolic models of plant, pathogen and plant – pathogen interactions.

# Chapter 2

# Constraint-based metabolic modeling, a mathematical framework to model metabolic responses to internal or external perturbations

## 2.1 System-based study of metabolism

Metabolism is described as the set of biochemical reactions in an organism, each converting some metabolites (named substrates) into others (named products), either through the action of a protein with catalytic activity (enzyme) or spontaneously (rarely). Metabolism also includes transport of metabolites from a cell compartment into another or between outside and inside the cell. This can be achieved passively by diffusion or through a transporter, depending of the metabolite size and chemical properties. In the bacterium *Escherichia coli*, the amount of reactions is estimated around 2,719, and the number of metabolites around 1192 (Monk et al., 2017). Yet, part of metabolism remains probably unknown or not properly characterized. In addition, more complex organisms such as plants or mammals contain also specific secondary metabolites that importantly enhance the number of biochemical reactions (e.g. hormone production pathways).

Metabolism can be represented as a network, metabolites being nodes and reactions being the links between nodes. Due to the high size of the network, metabolism as a whole remains complex to study. In addition, metabolism is not a simple linear process: numerous hub metabolites (e.g. acetyl-coA) can take part in different pathways such as lipid or vitamin biosynthesis, and other, such as water or ATP (adenosine triphosphate, which is the principal form of biochemical energy) are ubiquitous. This high connectivity complicates the study of metabolism at the system-

level. As presented in **Figure 2-1**, the visualization of the whole metabolic network of *E. coli* highlights the huge size and connectivity of its metabolism, and the afferent difficulty to apprehend it.

Bioinformatics and mathematics are possible tools to untangle this complexity. In the next parts, I will present how new metabolic networks are built for an organism of interest, and how mathematical modeling can be used to simulate the fluxes of matter at the cell level. Finally, I will present what has already been done in metabolic modeling for plant, plant pathogens and plant – pathogen interactions, and what methodological frameworks related to metabolic modeling are developed in this thesis.

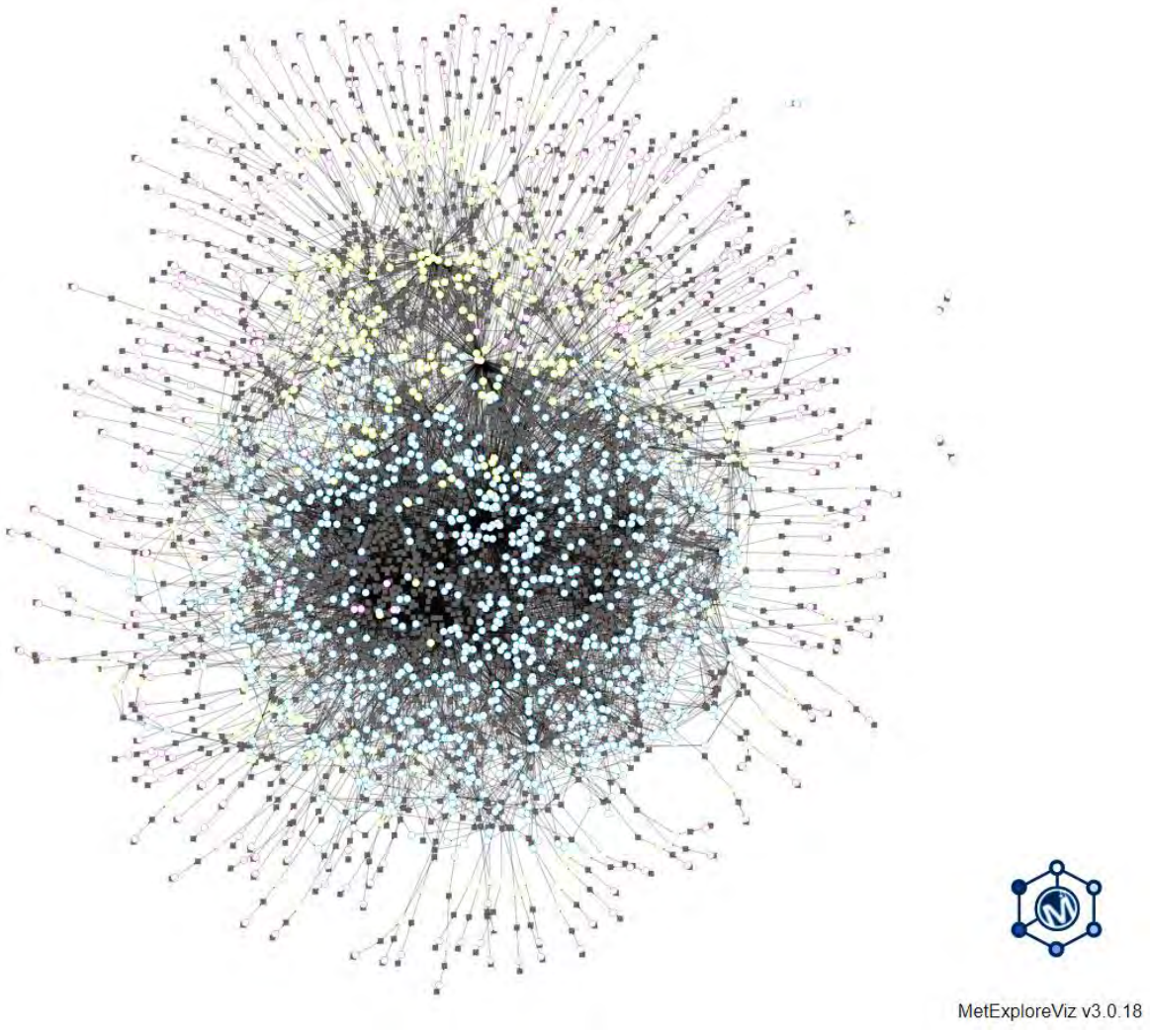



**Figure 2-1. Visualization of *Escherichia coli* metabolic network version iML1515 (Monk et al., 2017).**

Grey squares: metabolic reactions, white circles: metabolites from extracellular space (pink outline), periplasm (yellow outline), cytosol (blue outline). Grey line connects each reaction to the metabolites it produces or consumes. This figure was generated using MetExplore (Cottret et al., 2018).

## 2.2 Genome-scale metabolic network reconstruction

To build the most exhaustive metabolic network of an organism, genomics data are the usual starting point. Using a metabolic network reconstruction pipeline, it will generate automatically a genome-scale metabolic network. The resulting network is often named "draft network" and will be curated afterward to generate a final high-quality genome-scale metabolic network. A complete protocol for genome-scale metabolic network reconstruction is described by Thiele & Palsson (2010).

### 2.2.1 Automatic reconstruction of draft network

The organism of interest proteic sequences' and/or genome annotations' are used to predict the hypothetical list of metabolic reactions existing in the organism. Several tools and pipelines exist to generate this list. Many of them use orthologies with dataset containing Gene-Protein-Reaction (GPR) associations: one or several genes are associated to a protein (enzyme), and one or several proteins (enzymes) are associated to a metabolic reaction (list and stoichiometry of substrates, products, reversibility). These datasets exist online on databases such as KEGG (Kanehisa et al., 2017), TheSeed (Henry et al., 2010) or MetaCyc (Caspi et al., 2016).

Rather than using a large database, it is also possible to use already reconstructed high-quality genome-scale metabolic networks of organisms such as *E. coli*: in existing high-quality metabolic networks, GPR are also present. The advantage of using online databases is that they gather GPR on very large sets of existing organisms, providing a robust and larger list of metabolic reactions. However, these databases are not manually curated and they contain many errors and imprecisions, for example in stoichiometry or reaction direction, or incorrectly named metabolite or reaction identifiers. Using high-quality genome-scale networks have the advantage to limit the amount of errors on identifiers or stoichiometry as they were manually verified. However, by using this method we could miss some very specific metabolic pathways. This is why it is recommended to perform this method on several high-quality networks and then merge the resulting list of reactions to avoid missing some pathways. It is also possible to use different metabolic network reconstruction pipelines and then merge the results to have a final metabolic network.

These methods based on proteic sequences do not directly integrate spontaneous reactions, or reaction catalyzed by unknown or not-well annotated protein, which is notably the case for some transport systems. Some automatic pipelines add them if the associated metabolites were already recovered in the network.

### 2.2.2 Network curation and generation of high-quality network

All the pipelines described for metabolic network reconstruction could generate inaccuracies inducing major errors in further system-level study of metabolism. For example, a same metabolite could have different identifiers and some intermediate reactions required to generate a final product of metabolism could be missing. Using different databases or networks, or a mix of them, could generate conflicts: the same reaction could be written as reversible in a database and irreversible in another for example. To handle these aspects, a phase of network checking, correction and validation is required.

This phase is named “network curation” and it is recommended to do it manually. The person in charge of the curation will examine the reactions and check the consistencies of the network generated, for example pathway after pathway. Some automatic tools can be used to accelerate the process, for example scripts validating the stoichiometry of the reaction or performing gap-filling for incomplete pathways. The person will also make arbitrary decision when some reactions are uncertain. For example, if two genes are usually needed to have a metabolic reaction but only one have been recovered by orthology, they will decide, based on current literature available, if the reaction must be kept or not in the network. Also, species-specific metabolic pathways are not recovered by automatic reconstruction, and must be added manually based on the knowledge on the organism of interest. The person will also include biochemical data known for a given reaction, for example by adding the reference to a publication where the enzyme is characterized. They can give a confidence score to the each reaction: a high confidence score if it has been validated by biochemical assays, a limited confidence score if it is only based on orthology, and a minimal score if there is no proof but it is required in a pathway for modeling purposes.

To validate the functionality of pathways, and then of the whole network, it is also possible to start performing flux modeling on simple pathways, to check if it can propagate matter from the initial substrate to the final product of the pathway. Flux modeling will be used again once the curation is complete and the network qualified as

"high-quality", to perform genome-scale metabolic modeling. Flux modeling is described in 2.3.

### 2.2.3 Formats and databases of metabolic networks

Once finalized, high-quality metabolic networks can be saved in different formats such as a tabulated file with the identifiers and formula of the reactions, the associated list of genes and additional information such as comment, bibliographical references, Enzyme Classification (EC) number, confidence score. They can also be saved as SBML (Systems Biology Markup Language) files, a XML-based format developed to standardize the information generated in the diverse fields of Systems Biology. On an SBML file, metabolic network information is split into different components such as metabolites (named species), reactions, and compartments. These elements are gathered in different containers (such as <listofReactions> for the reactions), and each element can have a variety of attributes to add all the information initially present in the tabulated file (confidence score, EC number…). This format is well-suited to perform modeling or computational analyses on metabolic networks, but remains not user-friendly for biologists. It seems then necessary to provide the same network in different formats to maximize its accessibility and possibility of reuse.

The network can also be shared in databases dedicated to biological networks and models to facilitate its visibility and reuse. An important number of metabolic networks are available for download in BioModels (Glont et al., 2018) and *in silico* organisms (Feist et al., 2009). BiGG database (King et al., 2016) contains metabolic networks generated in SBML with metabolite and reaction identifiers following the BiGG nomenclature (e.g. M_atp_c for Adenosine Triphosphate in cytosol, and R_GLNS for Glutamine Synthetase reaction). It contains a user-friendly interface that allows easily searching for a specific metabolite and reaction, finding in which organism and published network it is identified, and have additional information such as its chemical formula or charge for a metabolite. It also links each metabolite and reaction to external biochemical databases for additional information, such as KEGG (Kanehisa et al., 2017), MetaNetX (Moretti et al., 2021), and RHEA (Lombardot et al., 2019). Finally, the online tool MetExplore (Cottret et al., 2018) contains also several high-quality metabolic networks. It has a user-friendly interface to search for specific information in a network and to perform online flux modeling and visualization/omics data mapping on the network of interest.

### 2.2.4 Graph representation and visualization

High-quality genome-scale metabolic networks, once reconstructed, can be modeled using a graph-based approach. A metabolic network can be represented as a weighted bipartite graph, which links metabolites (substrates) to reactions, and reactions to metabolites (products), with stoichiometric coefficients on these links. These links can be directed in one sense for irreversible reactions, or in two senses for reversible reactions (toy example in **Figure 2-2A,B**).

Graph representation is used to develop visualization methods of metabolic networks, which is challenging regarding the size and high connectivity of the network. Once a satisfying visualization is obtained, it can be efficiently used to integrate omics data related to metabolism: metabolic gene expression (that can be associated with reactions) or metabolomics (that gives a set of metabolites, and potentially their amount). This approach will not be used and developed in this thesis.

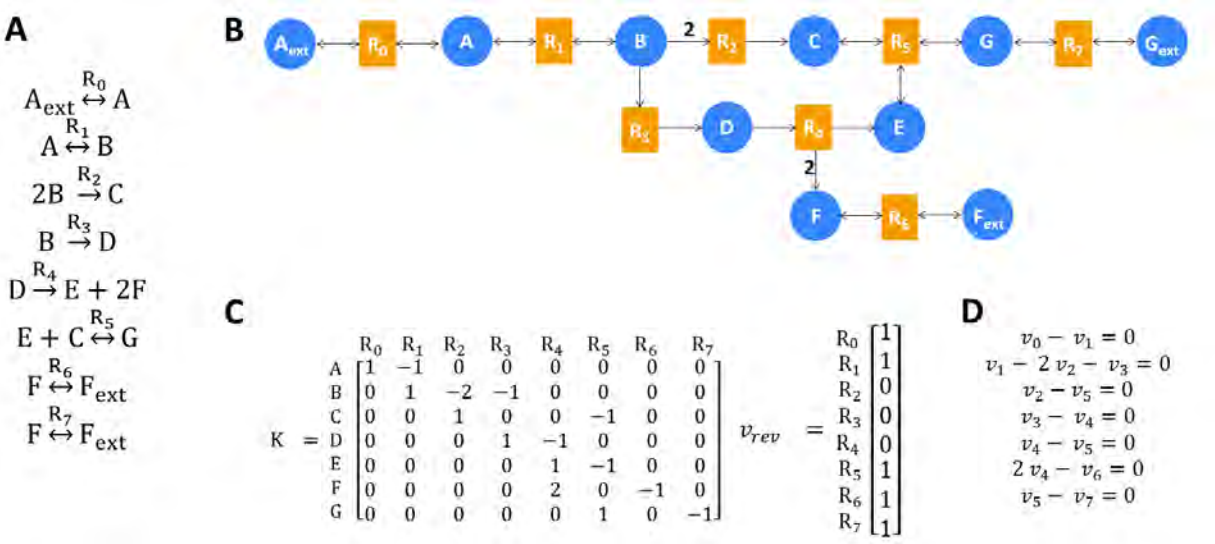



**Figure 2-2: Toy example of metabolic network and its conversion into mathematical system.**

A. List of reactions. B. Conversion of the list into a weighted mixed bipartite graph. C. Stoichiometric matrix on internal metabolites and vector of reversibility. D. System of linear equations on fluxes based on QSSA.

## 2.3 Modeling of metabolic fluxes

### 2.3.1 General formalism of biochemical kinetics

A mass balance on a metabolite (i.e. conservation of mass) inside the cell implies:

$$\frac{dm}{dt} = \sum_{p=1}^{n} k_p \cdot v_p \quad \textbf{(1)}$$

*with*
$m$: concentration of the metabolite (mmol·g dry biomass$^{-1}$)
$t$: time (h)
$n$: number of reactions in which the metabolite is involved
$k_p$: stoichiometric coefficient of the metabolite in reaction p ($1 \leq p \leq n$)
$k_p$>0 if the metabolite is a substrate, $k_p$<0 if the metabolite is a product
$v_p$: flux in mmol·h$^{-1}$·g dry biomass$^{-1}$

This local mass balance can be generalized to the whole cell metabolism, on equation (2). An example of stoichiometric matrix is given for the toy example on **Figure 2-2C**.

$$\frac{dM}{dt} = K.V \quad \textbf{(2)}$$

*with*
$M$: vector (size $n_M$) of the concentrations of the $n_M$ internal metabolites in the metabolic network (concentrations in mmol·g dry biomass$^{-1}$)
$t$: time (h)
$V$: vector (size $n_R$) of the mass fluxes of the $n_R$ reactions in the metabolic network (fluxes in mmol·h$^{-1}$·g dry biomass$^{-1}$)
$V_i \geq 0$ if reaction $i$ irreversible
$K$: stoichiometric matrix of size ($n_M$, $n_R$), $K \epsilon \mathbb{R}^{n_M \times n_R}$, extracted from the metabolic network
Value $K_{i,j}$ corresponds to the stoichiometric coefficient associated to the metabolite $i$ and the reaction $j$

To calculate the quantity of the metabolites over time, the metabolite fluxes must be known. They depend on several parameters described by the enzyme kinetics (catalytic constant $k_{cat}$, Michaelis constant $K_M$, maximum rate $V_{max}$) and their mathematical formulation can follow a Michaelis-Menten kinetics or alternative kinetics formulation for cases such as allosteric enzymes. Furthermore, these parameters are often known

in isolated enzymes, but not in an *in vivo* context. On a genome-scale metabolic network, it becomes impossible to access to each reaction kinetics parameters and formulation, because of lack of data. Hence, the system must be simplified to simulate the metabolism at a genome-scale level.

### 2.3.2 Quasi-Steady-State Approximation

To reduce the space of solution, a Quasi-Steady-State Approximation (QSSA) is formulated (Song & Ramkrishna, 2009). It consists in considering that metabolites inside a cell (elements of the vector M) are consumed at the same rate they are produced. This means that their amount is constant and implies that the time derivate is null. This leads to equation (3).

$$K.V = 0 \qquad \textbf{(3)}$$

$$V_i \geq 0 \text{ if reaction i reversible}$$

The hypothesis is named "Quasi"-Steady-State Approximation because it is only valid for metabolic fluxes inside a cell (Song & Ramkrishna, 2009). The extracellular products, including biomass, are not at steady-state, neither are the extracellular substrates. This should be kept in mind when performing dynamic modeling of metabolism, and also when we extend the approach to more complex systems such as multi-organ or ecosystem models, that cannot be approximated in a straightforward manner as a "meta-cell" at steady-state. Also, the equations 1 to 3 have been presented with units in mmol·g dry biomass$^{-1}$ and fluxes in mmol·h$^{-1}$·g dry biomass$^{-1}$, which is adequate for studying internal concentrations/fluxes. However, it becomes less adapted to study extracellular concentrations and their dynamics. This change of unit must be kept in mind when metabolic models become dynamic, are at multi-scale or are compared to experimental data.

QSSA simplifies the problem into a set of linear equation system, as presented in the toy example (**Figure 2-2D**). This reduces the space of solutions, which is now a convex polyhedral cone. However, the system is usually undeterminated: the number of reactions (corresponding to the number of variables) is superior to the number of metabolites (number of equations) and the amount of solutions remains infinite, even after defining the uptake rate of the cell substrate(s). To estimate quantitatively the metabolic fluxes, constraint-based optimization approaches are used (section 2.3.3).

Other modeling approaches, presented in 2.3.5, explore the set of solutions and their structure.

### 2.3.3 Flux Balance Analysis and constraint-based modeling: metabolic modeling through optimization

Metabolic modeling through optimization is named Flux Balance Analysis (FBA) (Orth et al., 2010) (see system (5)). It is based on the assumption that cells, by evolution, tend to optimize a specific objective. Defining an optimization problem and in addition imposing a finite uptake of carbon sources through upper bound and lower bound specifications, will allow the computation of metabolic fluxes in the network. To this end, only a linear programming solver such as Gurobi or Cplex is necessary.

On the toy example (**Figure 2-2**), maximizing the flux of reaction $R_7$ with a flux of reaction $R_0$ constrained between 0 (lower bound) and 3 (upper bound) mmol·$h^{-1}$·g dry $biomass^{-1}$ would lead to a flux of 1 mmol·$h^{-1}$·g dry $biomass^{-1}$ for $R_7$.

In the case of a bacterial growth, the objective is usually the maximization of biomass reaction flux (Feist & Palsson, 2010). Maximization of biomass reaction flux can be achieved by adding an artificial reaction to the metabolic network, representing the consumption of metabolites necessary to synthesize an average biomass composed of a proportion of proteins, carbohydrates, lipids, vitamins (among others) found in experimental data. The stoichiometric coefficients of this equation are usually defined in order to have the weight of 1 g for 1 mmol of biomass generated. This allows having the unit of $h^{-1}$ for the biomass flux, and thus easily compare it to experimental growth rate, also usually in $h^{-1}$. Using an experimental uptake rate of carbon source, the predicted biomass flux will be normally higher than the one observed experimentally, as FBA does not include the energetic cost of several processes related to growth (e.g. macromolecule synthesis) or not (e.g. maintenance of cell structure). Hydrolysis of ATP will be added to the biomass reaction to represent growth associated maintenance, and as an additional reaction for Non-Growth-Associated Maintenance (NGAM). The amount of this additional hydrolysis of ATP will be adjusted to remove the difference between experimental and predicted growth rates. The system (4) represents the usual formulation of FBA to model cell growth. FBA, also usually referred as constraint-based modeling, will allow to cartography the fluxes of matter in a whole cell, from substrate assimilation to growth and quantify them.

$$\max \text{ or } \min\ v_k$$

$$\left|\begin{array}{c} K.V = 0 \\ V_i \geq 0 \text{ if reaction i reversible} \\ l \leq V \leq u \end{array}\right. \quad (5)$$

with $l$: lower bound vector, $u$: upper bound vector, $v_k$ the reaction flux to optimize (maximize or minimize)

$$\max\ v_{Biomass}$$

$$\left|\begin{array}{c} K.V = 0 \\ V_i \geq 0 \text{ if reaction i reversible} \\ l \leq V \leq u \\ \text{with:} \\ v_{uptake\ substrate} = \text{k1} \\ v_{NGAM} = \text{k2} \end{array}\right. \quad (6)$$

with $l$: lower bound vector, $u$: upper bound vector, $v_{biomass}$ : flux of the biomass reaction including an growth-associated maintenance, $v_{uptake\ substrate}$ : flux of carbon substrate uptake, with k1 the experimentally determined value, $v_{NGAM}$ : flux of the non-growth-associated maintenance reaction, with k2 the value to have a biomass flux that fits the experimental growth rate.

### 2.3.4 Extensions of Flux Balance Analysis

Even if, by construction, the linear optimization problem (4) has a unique optimum $v_{k_{opt}}$, several sets of solutions $V_{opt}$ exists to reach this optimum. Different methods were developed to explore the alternative solutions. Flux Variability Analysis (FVA) (Mahadevan & Schilling, 2003) is the method the most used so far, and consists in solving additional optimization problems by adding the optimum $v_{k_{opt}}$ as an additional constraint and exploring the minimum/maximum flux possible for each flux of the network (see system (6)). It allows to find the maximum and minimum of each reaction flux that sustains the optimality criteria (e.g. maximal growth). This is also a way to analyze the flexibility of a metabolic network for example regarding biomass production.

$$
\begin{aligned}
&\max \text{ or } \min\ v_k \\
&\left|\begin{array}{c} K.V = 0 \\ V_i \geq 0 \text{ if reaction i reversible} \\ l \leq V \leq u \end{array}\right. \\
&\text{for i in 1 to } n_R \\
&\quad \min\ v_i \\
&\quad \max v_i \\
&\quad \text{subject to.} \\
&\quad\left|\begin{array}{c} K.V = 0 \\ V_i \geq 0 \text{ if reaction i reversible} \\ l \leq V \leq u \\ \text{with:} \\ v_k = v_{k_{opt}} \end{array}\right.
\end{aligned}
\qquad (6)
$$

with $v_{k_{opt}}$ : the solution of the first optimization problem, $n_R$: the number of reactions in the metabolic network

Classical FBA formulation allows modeling of flux at one specific time point. To perform dynamic modeling of fluxes, some frameworks were developed such as dFBA (dynamic Flux Balance Analysis). One of the dFBA methods consists in performing FBA at different time interval, using the solution after the time t as new constraint for the time t+1. This allows to follow the dynamics of nutrient availability, which is particularly adapted to better understand growth on two or more substrates (Mahadevan et al., 2002).

Extensions of FBA also allows estimating the essentiality of reactions and genes present in a metabolic network. Reaction essentiality consists in constraining successively each reaction flux to zero, and then computing the optimal objective flux. If no (or zero) solution is obtained, it means that this reaction is essential for the objective (e.g. biomass flux). This allows inventorying all essential reactions regarding a biological objective. Gene essentiality processes similarly, by listing all the metabolic genes in the network, and constraining successively to zero all the reactions where a given gene is mandatory. These two approaches will inform on the genetic and

metabolic robustness of an organism regarding an objective such as growth, and on the knockouts that could be performed to achieve a specific purpose (e.g. metabolic engineering or generation of antimicrobial compound).

Numerous other extensions of FBA exist and the field is very dynamic due to its ability to explore large metabolic networks and link them to experimental data and physiology. Among the recent and promising extensions, resource allocation models take into account the biosynthesis of enzymes and ribosomes in the computational analysis, using additional quantitative proteomic constraints (Goelzer et al., 2015; Waldherr et al., 2015).

### 2.3.5 Other metabolic modeling approaches

Other approaches can explore the cone of solutions generated without using optimization. In particular, Elementary Flux Mode Analysis consists in extracting non-decomposable steady-state pathways from the network, named Elementary Flux Modes (EFMs). It implies that the deletion of a reaction in the pathway will disable the possibility to have a steady-state flux in the pathway. EFMs are thus qualified as "minimal function building blocks" (Zanghellini et al., 2013) of the metabolic network. EFMs are interesting tools to evaluate the structure of a metabolic network, as you can access to its topology. It has been successfully applied to metabolic engineering to find the minimal deletions required to design bacterial strains overproducing a compound of interest (Trinh et al., 2011; Unrean et al., 2010). However, the main limitation of this analysis is the exponential explosion of the number of EFMs while the network size increases: EFMs generation becomes computationally challenging for genome-scale metabolic networks. Current computational developments are ongoing to overpass this limitation, but classical EFM Analysis approach must be limited to portions of a large network, to a simplified network or to a network with additional constraints (Mahout et al., 2020). Generating EFMs can also be a starting point to simplify metabolic networks, as it gives the global equation between an input and an output at a pathway scale, before using dynamic modeling based on ODEs on the simplified system. It is the case for the Lumped Hybrid Cybernetic Modeling (LHCM) approach proposed by (Song & Ramkrishna, 2010) or the Dynamic Reduction of Unbalanced Metabolism (DRUM) approach by Baroukh et al. (2013). These approaches allow analyzing the dynamics of complex metabolic processes, such as taking into account the storage/consumption of some metabolites of microalgae metabolism subjected to light.

Flux Coupling Analysis (FCA) (Burgard et al., 2004) is also a way to analyze the cone of solutions obtained from QSSA. It deciphers the coupling relation between each pair of reaction in the network. Three different coupling relationships are described: i) directional coupling (non-zero flux in reaction 1 implies non-zero flux in reaction 2 but not inversely), ii) partial coupling (non-zero flux in reaction 1 implies non-zero flux in reaction 2 and inversely but with no specific flux ratio), iii) full coupling (non-zero flux in reaction 1 implies non-zero flux in reaction 2 and inversely, with fixed flux ratio). Different computational approaches exist to perform FCA on genome-scale metabolic network, not necessarily based on optimization (David et al., 2011; Larhlimi et al., 2012). FCA can be used prior to performing gene knockout, as it links the deactivation of one reaction to whole network effects (Burgard et al., 2004), as well as to better understand gene transfer and evolutionary trajectories (Pál et al., 2005) or as a support for network curation and validation (Burgard et al., 2004).

On the following sections of the thesis, we will only focus on metabolic modeling through optimization (FBA and extensions), also usually named constraint-based modeling, due to its easy application to large metabolic networks and of its suitability to tackle the quantitative questions that emerged in Chapter 1.

## 2.4 Applications of metabolic modeling to plant pathogens and their hosts

*I contributed to a review describing metabolic modeling advances in plants and plant-interacting microorganisms and discussing on the perspectives of the field, for a chapter published in* Advances in Botanical Research *(Gerlin et al., 2020). This review is available in* ***Appendix I****. The following section presents a brief summary of the topic.*

Genome-scale metabolic modeling, and in particular FBA, has been mainly developed for metabolic engineering and remained non-used for a long time in the field of plant pathogens. However, since 2015, around one new genome-scale model is published each year for those organisms, and the field starts to benefit from these methods (**Table 2-1**). Generating and analyzing metabolic models helps understanding several aspects of the pathogen biology: the metabolic adaptation at different stages of the infectious cycle, the metabolic heterogeneity in a population or quorum sensing based regulation of metabolism. It also helps mapping omics data such as transcriptomics,

and in a more applied perspective, it could predict the effect of bactericide or nitrogen fertilizers to act directly against the pathogen spread.

**Table 2-1: Genome-scale metabolic models of plant pathogens (adapted from Gerlin et al., 2020)**

| Species | Classification | Publication | Size | Applications |
|---|---|---|---|---|
| *Pectobacterium carotovorum* | γ-proteobacteria | (C. Wang et al., 2015) | 1,113 metabolites<br>2,235 reactions<br>1,209 genes | discovery of agricultural bactericide targets |
| *Ralstonia solanacearum* | β-proteobacteria | (Peyraud et al., 2016) | 2,644 metabolites<br>2,086 reactions<br>1,476 genes | resource allocation trade-off between growth and virulence |
| *Phytophtora infestans* | eukaryota, oomycete | (Rodenburg et al., 2018) | 2,394 metabolites<br>2,685 reactions<br>1,301 genes | mapping of transcriptomics data at different life-stages |
| *Pectobacterium parmentieri* | γ-proteobacteria | (Zoledowska et al., 2019) | 2,080 metabolites<br>2,182 reactions<br>1,245 genes | metabolic behaviors at different steps |
| *Sclerotinia sclerotiorum* | eukaryota, fungus | (Peyraud et al., 2019) | 1,525 metabolites<br>1,495 reactions<br>1,039 genes | metabolic heterogeneity/division of labor |
| *Xanthomonas oryzae* | γ-proteobacteria | (Koduru et al., 2020) | 808 metabolites<br>839 reactions<br>673 genes | effects of nitrogen fertilizers, relation between quorum sensing and metabolism |
| *Xanthomonas phaseoli* | γ-proteobacteria | (D. Botero et al., 2020) | 1,527 metabolites<br>1,556 reactions<br>890 genes | analysis of transcriptomics data for wild-type and quorum sensing mutant |

As presented in chapter 1, plant and pathogen metabolism are intertwined during the infection process. Thus, plant genome-scale metabolic models are also important resources to decipher plant – pathogen interactions. Diverse plant genome-scale metabolic models are available (Table 2-2), starting in 2009 by an Arabidopsis (*Arabidopsis thaliana*) metabolic model (Poolman et al., 2009). Twelve different plant species, from various botanical families, have at least one genome-scale metabolic model. In several cases, updates of the initial one are also available. These updates focus on investigating methodological problematics such as building whole plant model (de Oliveira Dal'Molin et al., 2015; Mintz-Oron et al., 2012) or representing day/night dynamics (Shaw & Cheung, 2018) as well as biological issues: effects of environmental conditions (K. Botero et al., 2018; Pfau et al., 2018), photorespiration (De Oliveira Dal'Molin et al., 2010; Poolman et al., 2013), interpretation of omics data (De Oliveira Dal'Molin et al., 2016; Lakshmanan et al., 2015; Scheunemann et al., 2018).

**Table 2-2: Genome-scale metabolic models of plants (adapted from Gerlin et al., 2020)**

| Species | Classification | Published models |
|---|---|---|
| *Arabidopsis thaliana* | Eudicot, Rosids, Brassicaceae | 11 |
| *Oryza sativa* (rice) | Monocot, Commelinids, Poaceae | 5 |
| *Zea mays* (maize) | Monocot, Commelinids, Poaceae | 5 |
| *Sorghum bicolor* (sorghum) | Monocot, Commelinids, Poaceae | 1 |
| *Saccharum officinarum* (sugar cane) | Monocot, Commelinids, Poaceae | 1 |
| *Setaria italica* | Monocot, Commelinids, Poaceae | 1 |
| *Solanum lycopersicum* (tomato) | Eudicot, Asterids, Solanaceae | 1 |
| *Solanum tuberosum* (potato) | Eudicot, Asterids, Solanaceae | 1 |
| *Medicago truncatula* | Eudicot, Rosids, Fabaceae | 1 |
| *Mentha x piperita* (peppermint) | Eudicot, Asterids, Lamiaceae | 1 |
| *Populus trichocarpa* (poplar tree) | Eudicot, Rosids, Salicaceae | 1 |
| *Glycine max* (soy) | Eudicot, Rosids, Fabaceae | 1 |

As numerous plant and pathogen metabolic models were published (**Table 2-1, Table 2-2**), the next challenge is to combine them to build plant – pathogen metabolic models. This would help understanding trophic preferences of pathogens, the effect of one organism on the other and would also help predicting how we can limit the proliferation of a pathogen. Some previous multi-organism metabolic models were developed in the last fifteen years, for applications such as anaerobic digestion (Stolyar et al., 2007), bacterial coculture for bioprocesses (Salimi et al., 2010) and human health (Bordbar et al., 2010). Very close to the plant – pathogen interaction field, a plant – symbiont metabolic model was published in 2020 (diCenzo et al.). The model, named ViNE (Virtual Nodule Environment), gathers a previously published multi-tissue metabolic model of the plant *Medicago truncatula* (Pfau et al., 2018) and a genome-scale metabolic model of *Sinorhizobium meliloti* (diCenzo et al., 2016; diCenzo et al., 2018). Interestingly, this model integrates and analyzes several quantitative aspects of the interaction such as the different weights of plant and symbiont, the growth rates, the nitrogen fixation rates and the costs associated with nitrogen fixation and nodule formation. ViNE is an inspiring way of generating and using a plant – microorganism model, and the approach could be easily transferred to plant – pathogen models. So far, only one plant – pathogen metabolic model exists, for the *Phytophtora infestans* – tomato system (Rodenburg et al., 2019). The principal aim of this model was to integrate transcriptomics and metabolomics data generated on this system, and this approach revealed the different metabolic phases of the

interaction. It also tested several objective functions to maximize/minimize in FBA, and applied FCA to the system. This gave qualitative insights but contrary to the ViNE model, no integration of quantitative data such as organ weights or growth rates were used, neither a precise definition of the nutrient amounts available at the interaction site.

Thus, three main methodological points should be addressed in the future in metabolic models of plant – pathogen interactions:

**i) Representing the interaction quantitatively**

The absence of quantitative data complicates the use of a multi-organ metabolic model to draw conclusions on an interaction. It is hard to evaluate if metabolites assimilated by the pathogen in the model, for example, are really in enough amounts to sustain its growth. In addition, quantitative data are mandatory to estimate the effects of a pathogen on plant metabolism and physiology: not including a weight ratio would lead to a ratio 1:1 between the plant and the pathogen, and overestimate the effect of the latter on its host. This data could be rather easily acquired or found in literature, such as plant and pathogen growth rates and weight ratios, ideally at different infection steps.

**ii) Representing resource availability at the interaction site**

Basic metabolic modeling, such as on a bacterium, strongly depends on the calibration of the substrate(s) availability and uptake(s) rate. Thus, representation of the resources available at the interaction site will play an important role on the model results and its biological interpretation. Thus, it is necessary to: (a) define which resources are made available by the plant for the pathogen, (b) quantify at which amount they are present, (c) indicate if these resources could be assimilated by the pathogen, and (iv) quantify the assimilation capacity of the resources by the pathogen. These data could be difficult to find in literature, and would require some substantial and potentially complex data acquisition experiments, depending on the plant – pathogen system.

**iii) Addressing quasi-steady-state approximation**

As presented in 2.3.2, genome-scale metabolic modeling at a cell scale relies on assuming that metabolites inside the cell are at steady-state (QSSA). At a multi-organism scale, continuing to use QSSA implies that all the metabolites (in plant, pathogen and interaction site) are at steady-state: consumed immediately as they are produced. This also implies that both plant and interacting organism grow at the same

velocity. This seems inadequate for the dynamics of plant – pathogen interactions as in several cases, the pathogen's proliferation will accelerate as it deploys its virulence mechanisms against the host, and oppositely, the plant growth will decelerate. To address this, the model can be simulated at a precise time point, and we could consider that on a relatively short time interval, the system could be considered at quasi-steady-state. This instantaneous simulation would then be repeated at different time steps to access the dynamics of the interaction, by updating the biomass of each organism at each time step. This process would require the acquisition of at least growth of the two organisms throughout the experiment. Another solution is to consider that separately, the organisms are at QSSA, and allow non-QSSA behaviors for the exchanged metabolites, by implementing kinetics parameters in the model.

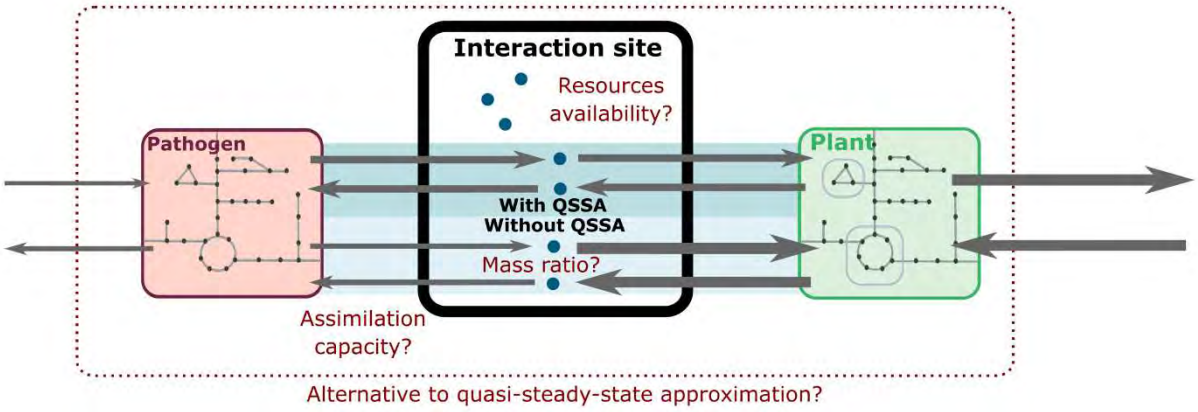



**Figure 2-3: Toward a quantitative and dynamic metabolic model of biotic interactions.**

The principal methodological and experimental issues that must be addressed are represented in red: quantification of the pathogen/plant ratio, use of a methodology that relaxes the QSSA growth hypothesis, and accurate representation of the resources available at the interaction site. Exchange fluxes are represented in two scenarios: assuming QSSA, metabolic flux from plant to interaction site and interaction site to microorganism must be equal (in mol/time unit) while without QSSA, this constraint is relaxed. Relaxing equality in exchange fluxes could be relevant in several cases, as the microorganism and the plant can have contrasted assimilation fluxes. Adapted from Gerlin et al. (2020).

## 2.5 Methodological frameworks of the thesis

Three main questions emerged in Chapter 1 from a state-of-art on plant – pathogen metabolic interaction and xylem-colonizing bacteria. These three questions constitute the main objectives of this thesis, and are addressed on the following chapters. They are: **i) How to model the environment of xylem-colonizing bacteria?** (Chapter 3), **ii) How bacterial metabolism is sustained in xylem environment and affects plant physiology?** (Chapter 4, Chapter 5) **iii) Do xylem-colonizing bacteria have different metabolic capacities and metabolic strategies?** (Chapter 6). All these questions require using metabolic modeling. I develop here the methodological frameworks related to these objectives.

### 2.5.1 Multi-organ tomato plant metabolic modeling

In **Chapter 3**, the objective was to build a multi-organ metabolic model of tomato plant during vegetative growth. Before my work, only a leaf metabolic model of tomato existed (H. Yuan et al., 2016). The first step was to adapt the model to build a metabolic model representing leaf, stem, root, exchanges between organs and with the environment. Also, the identifiers used in the metabolic model of Yuan et al. (2016) did not follow the same nomenclature than the ones for the model pathogen *Ralstonia solanacearum*. An update of the network with appropriate identifiers was thus performed. Plant metabolic modeling can be achieved using FBA only if i) we can apply QSSA at the whole plant level, meaning that during vegetative growth we can consider the whole plant as a "meta-cell" with no dynamics in exchange ii) we can find a global objective that the plant must minimize/maximize. These two points were checked using experimental data. In addition, accurate representation of the interaction site of plant – pathogen is crucial to understand the interaction, as presented in 2.4. On the current multi-organ metabolic models of plants, xylem is poorly detailed and exchanges between organ (such as leaf and root) are represented through a global "exchange pool", with no discrimination between phloem (transporting some nutrients from leaf to root) and xylem (transporting other nutrients from root to leaf). The objective was also to build a plant metabolic model that predicts xylem fluxes in response to specific constraints such as plant physiological state or plant nutrition. Finally, to have a reliable calibration of exchange fluxes, growth rates, weight ratios experiments on plant growth were also performed and integrated in the model.

### 2.5.2 Plant-pathogen metabolic modeling

In **Chapter 5**, the objective was to build a plant – pathogen metabolic model to understand bacterial metabolism in plant and its effects on plant physiology. This model is based on the plant multi-organ from Chapter 3 and on the *Ralstonia solanacearum* metabolic model (Peyraud et al., 2016). To calibrate this tomato – *Ralstonia* model, experimental data on the interaction were acquired and are presented in **Chapter 4**, to access to information such as resource availability, weight ratio between the plant and the pathogen, and assimilation capacities that are crucial for plant – pathogen metabolic modeling (as discussed in 2.4).

All these data were integrated in the metabolic model presented in **Chapter 5**. The model is thus able to simulate the impact of different pathogen/plant ratio on metabolic fluxes at different densities. However, using QSSA and considering the whole system as a "meta-cell" is not reasonable as plant growth stops while bacteria importantly proliferates, according to the results in **Chapter 4**. The model is thus limited in representing matter of fluxes at short and different time intervals, but I propose some methodological perspectives to develop dynamical modeling on the system.

### 2.5.3 *Xylella fastidiosa* genome-scale metabolic model

Chapter 6 is dedicated to build and analyze *Xylella fastidiosa* genome-scale metabolic model. The methodology of metabolic network reconstruction presented in 2.2 was applied, and this generated a high-quality model calibrated and refined with experimental data, made available in different formats (SBML, table file). Then, the model is analyzed through FBA and extensions (e.g. gene essentiality, FVA), and compared to other metabolic models such as *E. coli* and, more importantly, the plant pathogen *R. solanacearum*.

# Chapter 3
# Multi-organ metabolic model of a tomato plant and analysis of xylem fluxes and composition

## 3.1 Modeling the environment of xylem-colonizing bacteria

To understand metabolic interactions between plant and xylem-colonizing bacteria, it is essential to represent, analyze and predict in a precise manner the nutrients in the xylem available for bacterial growth. As xylem links roots and aerial parts of the plant, we must keep in mind that xylem composition is dependent on the whole plant metabolism, and dependent on environmental conditions (e.g. access to nutrients, light, watering). To predict or represent accurately xylem metabolic composition, we develop a multi-organ metabolic model of plant that integrates xylem fluxes in the model.

Several multi-organ metabolic models of plants exist for *Arabidopsis thaliana* (de Oliveira Dal'Molin et al., 2015; Mintz-Oron et al., 2012; Shaw & Cheung, 2018), barley (Grafahrend-Belau et al., 2013), soybean (Moreira et al., 2019), *Setaria viridis* (Shaw & Cheung, 2019). However, the biological questions investigated are far from nutrient availability in vascular tissues (xylem, phloem) and vascular tissues are usually rudimentary represented. In most of the models cited, xylem and phloem are not dissociated and studied, excepted in the soybean model developed by Moreira et al. (2019). In the soybean model, xylem and phloem are depicted as dissociate fluxes and there is some focus on phloem, but xylem remains poorly studied. Developing a multi-organ metabolic model with an emphasis on xylem fluxes and metabolic composition is thus a new approach in the field.

To study the composition of xylem fluxes in the context of plant – pathogen interaction, I choose as model plant the tomato plant (*Solanum lycopersicum*). Tomato plant has many advantages over other plants: i) tomato is the natural host of the bacterial model strain *Ralstonia solanacearum* GMI1000, internationally used to study bacterial wilt, ii) tomato morphology (rigid and large stem) allows easy xylem sap extraction compared to Arabidopsis, iii) my laboratory developed large expertise on cultivating, phenotyping and manipulating tomato, iv) tomato is one of the plant with the highest agronomical interest. Providing a tool to model whole tomato metabolism could thus be beneficiary for several research projects beyond my thesis.

## 3.2 Multi-organ metabolic model of tomato plant [article]

**Article**

A multi-organ metabolic model of tomato predicts plant responses to nutritional and genetic perturbations

**Journal**

*Plant Physiology* (published online 19 October 2021).

**Contributions**

- Constructing, calibrating and validating the metabolic model (with C. Baroukh and L. Cottret)
- Performing the phenotypic experiments (with A. Escourrou) and analyzing the experimental data (phenotypic experiments, metabolomics) (with A. Escourrou and C. Baroukh)
- Writing the complete draft article and designing the figures
- Editing the article (with the co-authors)

**Supplementary Materials** available in **Appendix II**.

**Summary**

The article presents VYTOP (Virtual Young TOmato Plant), a multi-organ metabolic model of tomato that integrates xylem and phloem fluxes. We detail the construction process of VYTOP as well as its use on several use cases, which include an analysis of the predictions of xylem fluxes.

The sections of Results “VYTOP, a multi-organ metabolic model”, “VYTOP calibration with physiological and metabolic data” are of great importance for the next chapters of the thesis, as they present the framework used, that will also be the starting point of the plant – pathogen metabolic model. The section “Use case 3: Infer the composition of tomato xylem sap by applying physiological constraints” discusses xylem composition, which is also important for the rest of the thesis. The other sections present applications of the multi-organ that are not directly connected to the thesis.

https://doi.org/10.1093/plphys/kiab548 PLANT PHYSIOLOGY 2021: 00: 1–15

Plant Physiology®

Research Article

# A multi-organ metabolic model of tomato predicts plant responses to nutritional and genetic perturbations

**Léo Gerlin [1], Ludovic Cottret [1], Antoine Escourrou [1], Stéphane Genin [1] and Caroline Baroukh [1,*,†]**

1 LIPME, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France

*Author for communication: caroline.baroukh@inrae.fr

†Senior author

C.B. conceived the project and designed the experiments; C.B. and L.C. supervised the construction, calibration, and validation of the model; L.G. constructed, calibrated, and validated the model; C.B. and S.G. supervised the experiments; A.E. and L.G. performed the phenotypic experiments and A.E. performed the metabolomics analysis; L.G., A.E., and C.B. analyzed the data; L.G. wrote the draft article; C.B., L.C., and S.G. supervised and completed the writing. C.B. agrees to serve as the author responsible for contact and ensures communication.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plphys/pages/general-instructions) is Caroline Baroukh (caroline.baroukh@inrae.fr).

## Abstract

Predicting and understanding plant responses to perturbations require integrating the interactions between nutritional sources, genes, cell metabolism, and physiology in the same model. This can be achieved using metabolic modeling calibrated by experimental data. In this study, we developed a multi-organ metabolic model of a tomato (*Solanum lycopersicum*) plant during vegetative growth, named Virtual Young TOmato Plant (VYTOP) that combines genome-scale metabolic models of leaf, stem and root and integrates experimental data acquired from metabolomics and high-throughput phenotyping of tomato plants. It is composed of 6,689 reactions and 6,326 metabolites. We validated VYTOP predictions on five independent use cases. The model correctly predicted that glutamine is the main organic nutrient of xylem sap. The model estimated quantitatively how stem photosynthetic contribution impacts exchanges between the different organs. The model was also able to predict how nitrogen limitation affects vegetative growth and the metabolic behavior of transgenic tomato lines with altered expression of core metabolic enzymes. The integration of different components, such as a metabolic model, physiological constraints, and experimental data, generates a powerful predictive tool to study plant behavior, which will be useful for several other applications, such as plant metabolic engineering or plant nutrition.

## Introduction

Systems biology can predict and explain how a whole organism responds to a stimulus or a perturbation and how different components of the organism are connected. In particular, one way to unravel physiological mechanisms is the use of constraint-based metabolic modeling. Constraint-based metabolic modeling relies on genome-scale metabolic networks, which gather all the identified metabolic reactions of an organism according to its genomic sequences and current metabolic knowledge summarized in databases (Gu et al., 2019). It predicts and quantifies the metabolic pathways used, thanks to mathematical constraints that encompass physiological states, such as the uptake rate of

Received September 30, 2021. Accepted October 27, 2021. Advance access publication November 19, 2021

© American Society of Plant Biologists 2021. All rights reserved. For permissions, please email: journals.permissions@oup.com

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021

nutrients present in the environment. Initially developed for unicellular organisms, constraint-based metabolic modeling has been extended to diverse plant species such as *Arabidopsis thaliana* (Mintz-Oron et al., 2012; Gomes de Oliveira Dal'Molin et al., 2015; Shaw and Cheung, 2018), *Medicago truncatula* (Pfau et al., 2018), barley (*Hordeum vulgare*; Grafahrend-Belau et al., 2013), *Steraria viridis* (Shaw and Cheung, 2019), and soybean (*Glycine max*; Moreira et al., 2019). These studies analyzed mechanisms such as the effect of the nitrogen (N) source (Arnold and Nikoloski, 2014; Arnold et al., 2015; Gomes de Oliveira Dal'Molin et al., 2015; Seaver et al., 2018), the N fixation by symbiotic bacteria (Pfau et al., 2018), the effect of diurnal cycles (Gomes de Oliveira Dal'Molin et al., 2015; Shaw and Cheung, 2018), or the genetic modifications to perform in order to enhance a product of interest such as vitamin E (Mintz-Oron et al., 2012; Saha et al., 2011). Some of these studies were supported with labeling data, such as Robaina-Estévez et al. (2017), which studied the metabolic differences between guard cells and mesophyll cells. For a full review of all models of plants developed so far, refer to Gerlin et al. (2021) and Shaw and Cheung (2020).

The representation of the multi-organ structure of plants through modeling is challenging (Clark et al., 2020), but is required to better understand plant metabolism. To model exchanges between organs, a global exchange pool is usually used (Gomes de Oliveira Dal'Molin et al., 2015; Shaw and Cheung, 2018), but the veracity of the exchanged fluxes has not been studied. Another major limitation is the lack of experimental calibrations since most of the methodologies gather heterogeneous data from literature, issued from multiple species and different growth scenarios, which could imply bias and inaccuracies. In addition, tomato (*Solanum lycopersicum*), a major agricultural crop, has not benefited from the advances in multi-organ metabolic models, as only a genome-scale metabolic model of leaf cell has been published (Yuan et al., 2016), as well as a metabolic model of the fruit (Colombié et al., 2017, 2015; Li et al., 2020).

In this study, we developed a multi-organ metabolic model of tomato during vegetative growth, in order to provide to the plant research community a comprehensive tool adapted for many biological questions such as the impact of environmental factors on the early development of the tomato plant, the exchange fluxes of matter between organs, or the prediction of the reactions to be modified to obtain a desired phenotype. This model, named Virtual Young TOmato Plant (VYTOP), includes leaf, stem, root and, contrary to existing multi-organ models, dissociated xylem and phloem compartments. It is able to simulate all the major metabolic reactions used in the plant and the sink/source relationships between organs. The model was calibrated with homogeneous data from experiments we performed, using an automated phenotyping platform, gathering both physiological data (growth and transpiration) and metabolomics (xylem sap chemistry and organ biomass composition). We chose five independent use cases to validate the usefulness of our model for a broad range of biological questions: (1) prediction of core metabolic fluxes in each tomato organ; (2) contribution of stem to photosynthesis; (3) organic xylem composition; (4) impact of N limitation on growth; and (5) predictions of metabolic changes in transgenic lines. Beyond its use to model the fluxes between tomato organs, this model could serve as a template for modeling other plants of interest as well as predicting and understanding the impact of diverse perturbations.

## Results and Discussion

### VYTOP, a multi-organ metabolic model

The multi-organ metabolic model, named VYTOP (Figure 1) was built by aggregating genome-scale metabolic models of each organ. The published tomato leaf genome-scale metabolic model iHY3410 (Yuan et al., 2016) was used as a starting point. We curated the model to integrate missing pathways such as the catabolism of different metabolites (lysine, ethanol [ETOH], isoleucine, beta-alanine, leucine, and cysteine). As the iHY3410 metabolic model identifiers (IDs) mostly followed the MetaCyc IDs nomenclature (Caspi et al., 2016), we converted them into the Biochemical Genetic and Genomic format (BiGG) (King et al., 2016) using a semi-automatic conversion framework. BiGG is widely used for genome-scale metabolic models from bacteria to microalgae and humans (Norsigian et al., 2020). The converted model was also manually curated to achieve mass balance for carbon (C), N, oxygen ($O_2$), phosphate, and sulfur for each reaction. We name this tomato metabolic model Sl2183; it includes 2,183 reactions, 2,097 metabolites and 3,433 genes. It is available in different formats: table format (Supplemental File S1), Systems Biology Markup Language (SBML) format (Supplemental File S2), in the database BioModels (Glont et al., 2018) under the ID MODEL2111120001, and in the MetExplore database to enable pathway visualization and omics mapping (Cottret et al., 2018) https://metexplore.toulouse.inrae.fr/metexplore2/?idBioSource=6353.

To transpose this genome-scale metabolic model into a multi-organ model, we considered the whole plant as three main organs (Figure 1). Each organ was modeled with a replicate of the Sl2183 genome-scale metabolic model with an organ-specific biomass equation. In addition, each organ has specific physiological roles: photosynthesis/organic matter production in leaf, partial photosynthesis in stem, and minerals/water uptake in root represented by accurate physiological constraints. The transport tissues xylem and phloem were defined as exchange compartments: xylem represents exchanges from root to leaf and phloem represents exchanges from leaf to root. Stem can exchange with both compartments in the two directions: uptake or contribution to xylem and phloem. Leaf and stem assimilate photons (i.e. light) while root assimilates water and minerals. An ATP cost was set up for transport reactions between organs and xylem/phloem to include the transfer costs. The final

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021

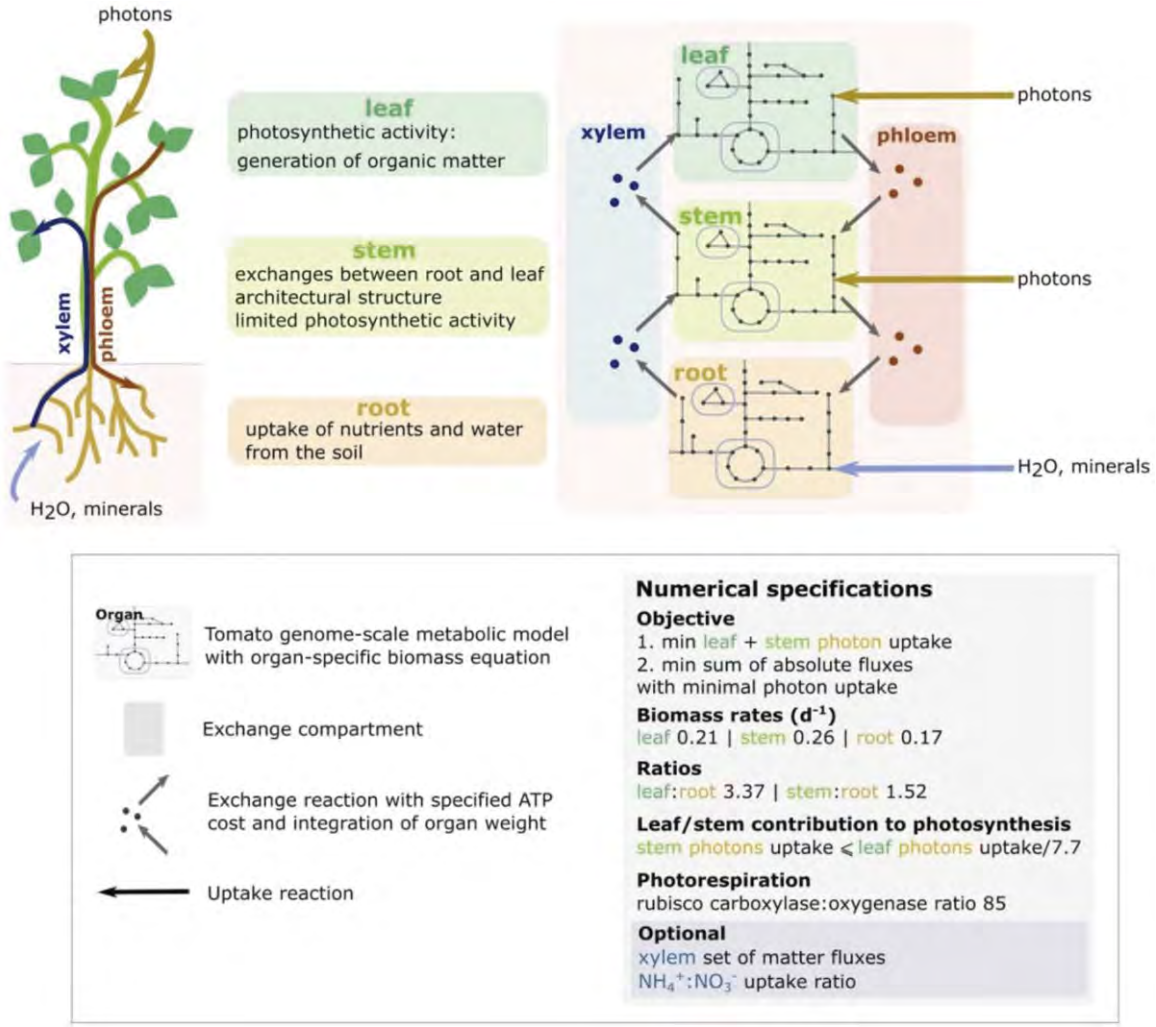



**Figure 1** Generation and calibration of VYTOP based on a tomato genome-scale metabolic model. Top left panel: role of each organ as modeled in VYTOP. Top right: Schematic view of the different organs and exchange compartments as modeled in VYTOP. Bottom: legend, numerical constraints, and objective functions used in VYTOP.

VYTOP model includes 6,689 reactions and 6,326 metabolites.

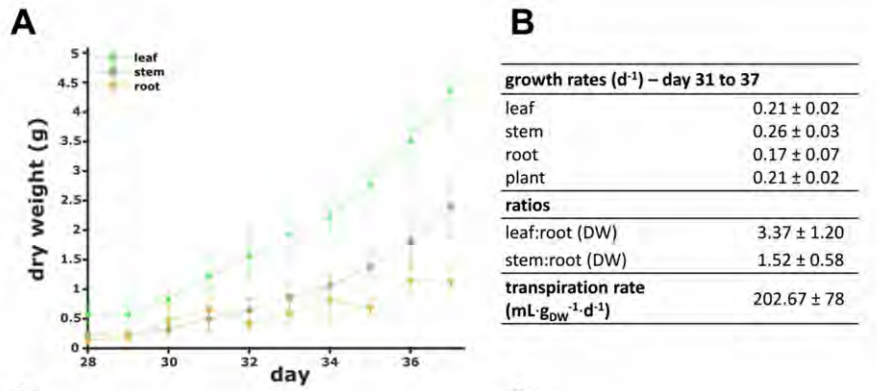
## VYTOP calibration with physiological and metabolic data

To calibrate VYTOP, we performed physiological and biochemical measurements (Figure 2) on 90 tomato plants. The growth of 90 plants was monitored on an automatic phenotyping facility, which allowed automatic monitoring of plant transpiration rates. Sampling of six plants per day was performed during 9 d, starting at 28 d after seeding. Xylem organic metabolite concentrations and organ weights were determined on sampled plants. The biomass composition of each organ was measured in another experiment performed in the same environmental conditions, with the same tomato variety and a similar plant age.

A weight dataset of sampled plants was used to measure the growth rate for each organ, the organ dry weight ratios (leaf:root and stem:root) (Figure 2, A and B), and the fresh:dry weight ratio for each organ (Supplemental Figure S1). The growth was considered as exponential between Days 31 and 37 after sowing (Supplemental Figure S1) and regressions of fresh/dry, leaf/root, stem/root weights showed high correlation coefficients (Supplemental Figure S1). Growth rates were used in VYTOP to constrain biomass fluxes; organ ratios were implemented in exchange reactions to appropriately represent the difference of weight between the three organs (Figure 1). Transpiration was constant per mass unit, and was converted into a transpiration rate (Figure 2B). The transpiration rate was further used to compute organic xylem metabolite fluxes (see "Materials and methods").

We measured organ biomass composition on fresh samples of growing tomato (Figure 2D). As expected, the leaf accumulates the highest proportion of starch and has the highest chlorophyll content. The data obtained was completed with data for lipids, minerals, (hemi)cellulose, and lignin from the literature to generate accurate organ-specific biomass equations (Supplemental File S1). For each organ, the biomass equation yields 1 g of dry weight biomass (Supplementary File S1).

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021




| C | metabolite | concentration (mM) | inferred flux ($\mu mol \cdot g_{DW}^{-1} \cdot d^{-1}$) |
|---|---|---|---|
| | alanine | 0.019 ± 0.009 | 3.94 ± 1.86 |
| | arginine | 0.093 ± 0.038 | 18.89 ± 7.63 |
| | asparagine | 0.213 ± 0.170 | 43.19 ± 34.44 |
| | aspartate | 0.005 ± 0.006 | 1.01 ± 1.28 |
| | ethanol | 0.113 ± 0.094 | 22.99 ± 19.03 |
| | fumarate | 0.001 ± 0.002 | 0.24 ± 0.36 |
| | glucose | 0.024 ± 0.045 | 4.92 ± 9.16 |
| | glutamine | 4.019 ± 1.419 | 814.59 ± 287.61 |
| | isoleucine | 0.054 ± 0.018 | 10.89 ± 3.65 |
| | leucine | 0.095 ± 0.036 | 19.24 ± 7.24 |
| | lysine | 0.177 ± 0.080 | 35.95 ± 16.14 |
| | phenylalanine | 0.027 ± 0.015 | 5.52 ± 3.03 |
| | proline | 0.099 ± 0.052 | 20.12 ± 10.60 |
| | sucrose | 0.041 ± 0.039 | 8.22 ± 7.96 |
| | threonine | 0.106 ± 0.092 | 21.45 ± 18.74 |
| | tyrosine | 0.025 ± 0.011 | 5.07 ± 2.22 |
| | valine | 0.079 ± 0.026 | 15.99 ± 5.36 |

**Figure 2** Measurement of physiological and biochemical parameters on 4-week-old tomato plants for calibration of VYTOP. A, Organ dry weight evolution between Days 31 and 37. The weight (mean ± standard deviation) of six plants was measured by day. B, Growth rate of each organ (from Days 31 to 37), dry weight ratios between organs (from Days 29 to 37), and transpiration rate (from Days 31 to 37). Values (mean ± standard deviation) are based on three replicates for growth rate, 54 plants for dry weights ratios and on 42 plants for transpiration rate. C, Average organic xylem sap composition. Concentrations (mean ± standard deviation) are based on 33 samples analyzed using $^{1}H$ NMR. d, Organ composition. Measurements are based on at least 19 samples per organ. Mean weights are plotted and bars indicate standard deviation.

In comparison with the Yuan et al. model iHY3410 (2016), our updated biomass equation for leaf had a C:N ratio closer to the one observed from experimental data (6.5) in a high $CO_2$ level and high N provided as a nitrate ($NO_3^-$):ammonium ($NH_4^+$) mixture (Royer et al., 2013): 5.48 in VYTOP versus 18.52 in iHY3410.

Quantification of organic metabolites concentration in xylem sap (Figure 2C) was performed using Nuclear Magnetic Resonance (NMR). Glutamine was the major metabolite in xylem sap (4.019 ± 1.419 mM). Fifteen other organic molecules were detected: 11 amino acids, 2 sugars, ETOH, and fumarate (FUM), whose concentrations were much lower than that of glutamine (< 0.250 mM). Organic metabolic fluxes in xylem were computed using these concentration data and the transpiration rate of plants (see "Materials and methods").

We also gathered data from literature to choose which compounds are present in the phloem compartment (Hijaz and Killiny, 2014) and to estimate photosynthetic activity in the stem. Former experimental results indicated that photosynthesis in the stem is limited due to less exposed surface to light (Hetherington et al., 1998). According to this study, the leaf:stem ratio of contribution to photosynthesis is around 7.7 (considering here the petioles as stems), and this value was added in the model as a constraint. The $NH_4^+$:$NO_3^-$ uptake ratio can be optionally constrained in the model to take into account different fertilizer conditions. Since the fertilizer used for the experiments contained both $NH_4^+$ and $NO_3^-$ and we did not measure the uptake rate of each N source, the ratio was unconstrained for our simulations except for Use Case 4.

## Model framework of VYTOP validated by experimental data

We used FBA (Orth et al., 2010), a constraint-based modeling approach to compute the metabolic fluxes in VYTOP. FBA requires a quasi-steady-state approximation (QSSA) in the whole system. This approximation is commonly

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021

accepted for bacterial growth modeling, but to extend the approach to plant vegetative growth, we checked if QSSA was still reasonable. We examined:

i. the daily weight ratios between organs and observed no substantial discrepancies between days (Supplemental Figure S1),
ii. the organic chemical composition of xylem, and also observed no important daily variation (Supplemental Figure S2),
iii. the three organ growths, which followed, as expected, an exponential growth (Supplemental Figure S1),
iv. the biomass composition of each organ, and found that it had no substantial daily variation (Supplemental Figure S3).

These experimental observations performed over a 10-d period proved that we could reasonably extend QSSA to a tomato plant at the vegetative growth stage. We did not take into account the day/night variation as we measured plant growth on the scale of a day and not of the hour, considering that our simulation results represented an overall average of the metabolic fluxes over a 24-h period (Bénard et al., 2015).

Simulation of metabolic fluxes using FBA implies defining a linear optimization problem: the formulation of linear mathematical constraints (equalities or inequalities) and a linear objective function (maximization or minimization of one or several fluxes). The choice of this objective function is still debated for plants (Sweetlove and George Ratcliffe, 2011; Collakova et al., 2012). We assumed that plant metabolism is regulated to favor the most efficient use of light uptake and C fixation to grow, which was translated into minimization of the photon uptake flux, while growth was constrained via the biomass flux value. In addition, we assumed that plant metabolism was also efficient in the use of enzymes and solved a second optimization problem with objective function minimization of the sum of all the model's fluxes (absolute value). This second simulation was also performed to avoid stoichiometric balanced cycles. The resulting fluxes of this double optimization problem were analyzed in the following sections.

## Use Case 1: VYTOP predicts metabolic fluxes in the whole plant

Figure 3 depicts the central metabolic fluxes obtained through modeling (Supplemental File S3). We used flux variability analysis (FVA) on the central metabolic fluxes to estimate how much they were constrained and if many alternative pathways existed. FVA provides for each reaction the minimum and maximum flux values that still sustain the optimal solution found in FBA (Mahadevan and Schilling, 2003). We found no variation on the fluxes obtained, except for some reactions of glycolysis (pyruvate [PYR] kinase and phosphofructokinase) that can either be performed in the chloroplast or in the cytosol (Supplemental File S3).

The analyses established that core metabolic fluxes are consistent with the current knowledge on plant physiology:

i. Metabolic activity is more intense in leaves than in stems and roots. This makes sense since leaves generate the major part of organic matter necessary for whole plant growth. In particular, there is intense photosynthetic activity and organic matter production in leaves with major fluxes in the Calvin cycle, while this pathway is not used in the root and is poorly used in the stem.
ii. The generated organic matter is primarily converted into sugars (sucrose [SUCR], glucose [GLC], and fructose), either transformed into starch or transferred to the stem and root through the phloem flux. Most of the organic matter (>50% of leaf assimilated C) is used directly by the leaf, as it is the organ with the highest weight (Figure 2).
iii. Stems need a lower SUCR uptake flux compared to roots (40% per gram of stem versus 60% per gram of root), because it has some photosynthetic activity allowing the production of part of the needed C and energy (as ATP and NADPH). Yet, the stem is still dependent on leaf C production.
iv. Glycolysis and the TCA cycle have high fluxes in the stem and root, allowing the generation of energy and biomass precursors from phloem SUCR and citrate.

Previous multi-organ metabolic models of plants were developed (Shaw and Cheung, 2020). The term multi-organ model refers here to models that represent several plant organs with different roles for each organ and with metabolite exchanges between each organ and not to tissue-specific models which simulate one type of organ at a time. For the multi-organ metabolic models developed so far, the calibration of constraints (organ ratio, growth rate, and authorized fluxes) to accurately predict the behavior of a plant organ remains rudimentary and does not rely on extensive biochemical and physiological data. A question that also arises for these models is how to represent accurately metabolism differences between each organ. To this end, several frameworks were developed and rely on different strategies. Some of these frameworks such as the model of Arabidopsis developed by Mintz-Oron et al. (2012) used -omics data (RNAseq, proteomics, and metabolomics) to take into account organ/tissue specificities. The initial methodology consisted of pruning inappropriate reactions using an expression threshold on RNAseq data to decide whether or not the reaction takes place in the organ/tissue, and therefore should give a closer biological relevance. However, it is limited by the fact that the transcription does not include subsequent regulation layers (translation, posttranslational modifications, and enzyme activation), which can be particularly important for central C metabolism (de Groot et al., 2007). Also, choosing this expression threshold can be sometimes arbitrary, particularly for nonmodel plants. Other frameworks were developed since, which minimize these drawbacks, such as GIMME-like, iMAT-like, and MBA-like

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021

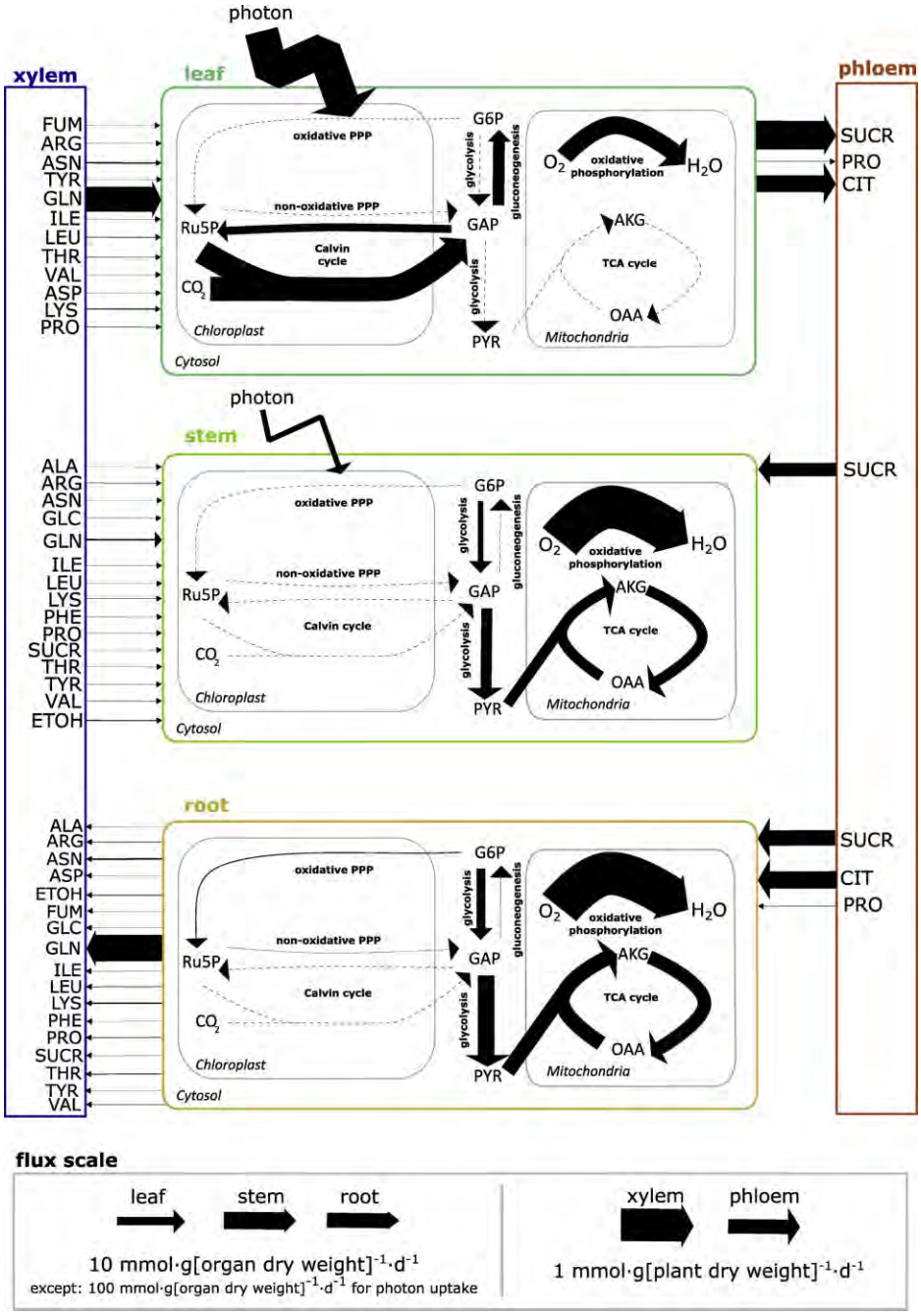



**Figure 3** Core C metabolic fluxes of a tomato plant at the vegetative growth stage predicted by VYTOP, with xylem fluxes constrained from experimental data. Each pathway flux was estimated based on a reaction flux value representative of the pathway (listed in Supplemental File S3). OAA, oxaloacetate; GAP, glyceraldehyde 3-phosphate; G6P, glucose-6-phosphate; Ru5P, ribulose-5-phosphate. Other abbreviations follow standard abbreviations for amino acids.

methods and the RegrEx method (Robaina Estévez and Nikoloski, 2014; Scheunemann et al., 2018). Even if thresholds can still be used in some of these methods, pruning is not as strict as the initial method since only a minimization of the discrepancy between the model and -omics data is performed. Finally, another multi-organ plant model relies on defined constraints with a global source–sink macroscopic model representing the allocation of C between organs (Grafahrend-Belau et al., 2013).

In our approach, we chose not to constrain each organ using -omics data, to allow all the possible chemical reactions a priori. To constrain VYTOP, we chose instead to use an important set of external physiological constraints (growth rates, ratios, well-calibrated biomass, and photosynthetic contributions) to impose organ-specific and sink/source metabolic behaviors in the different organs, which is closer to the methodology developed for barley. The C fluxes distribution in central metabolism obtained in VYTOP

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021

demonstrated that this approach, relying only on physiological constraints, allows an overall good prediction of metabolic fluxes between the different organs of a whole plant, despite the omission of -omics data. Its disadvantage is that it relies on the acquisition of physiological (and ideally metabolic) data that are not always available (whereas numerous sets of RNAseq data are available). However, it advantageously removes some of the methodological questions required to build -omics-based metabolic models (Richelle et al., 2019). In the future, it would be interesting to compare the predictions of VYTOP with models exploiting -omics data in order to reveal more precisely the advantages and drawbacks of each methodology.

## Use Case 2: VYTOP assesses the impact of physiological changes on the plant

We used VYTOP to assess the influence of the stem on plant metabolism, namely its contribution to photosynthesis (Figure 4) and the effect of the stem/whole plant ratio (Figure 5).

The contribution of the stem to photosynthesis represents the ability to perform photosynthesis comparatively to the leaf and depends on the surface of the stem (Hetherington et al., 1998). The model shows that this contribution strongly affects photon demand since the latter decreases rapidly while the stem photosynthesis capacity increases (Figure 4). In agreement with this observation, photosystem I metabolic flux strongly decreases in the leaf while it increases in the stem, as it becomes more advantageous to perform photosynthesis in the stem (also true for photosystem II, see Supplemental Figure S4). We also observed a proportional decrease of leaf export to phloem and stem uptake from phloem. However, there is no C contribution of the stem to phloem, because the range of values shown here did not allow sufficient photosynthetic activity of the stem to support both stem and root growth. Yet, a SUCR flux from stem to phloem appeared when stem contribution to photosynthesis exceeds 40% (Supplemental Text). Simulations also showed that increasing the stem weight proportion in the plant impacted progressively the photon demand (Figure 5). In this case, the plant stem plays a role of sink for C sources, which progressively becomes a burden affecting the photon demand (i.e. synthesis of C sources).

VYTOP does not integrate the architectural role of the stem in plant growth, enabling a better access to light that counterbalances its cost as sink of matter. A model integrating both plant geometry and metabolic fluxes would therefore be of great interest to study the tradeoff between the architectural role and the C sink, but remains a challenging issue.

## Use Case 3: VYTOP infers the relative composition of organic metabolites of tomato xylem sap by applying physiological constraints

To evaluate the relevance of VYTOP on exchange fluxes between organs, we simulated plant metabolic fluxes with the

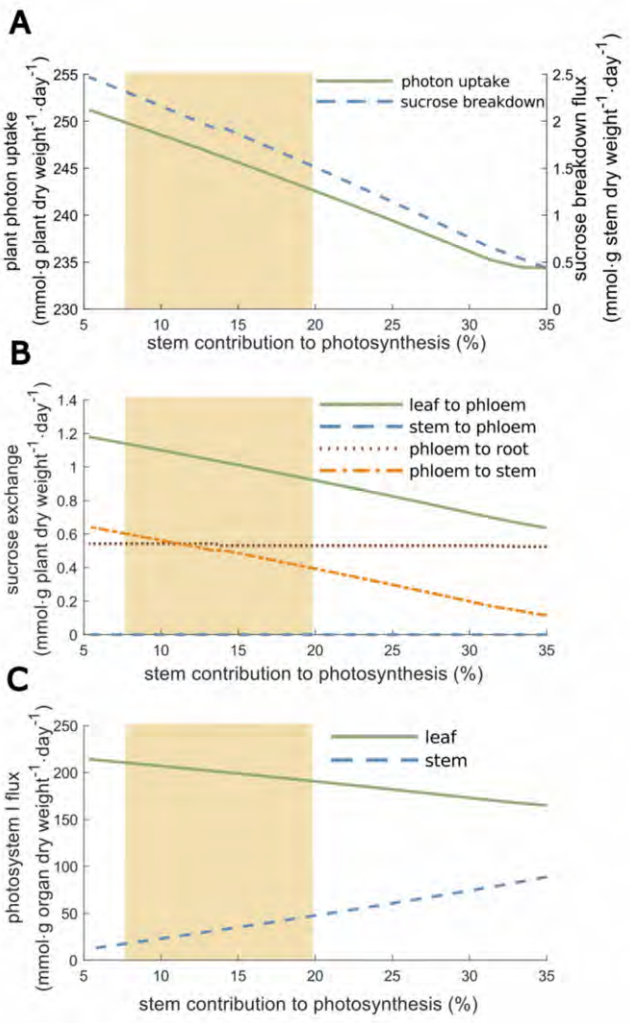



**Figure 4** Effect of different leaf:stem photosynthetic capabilities on photon uptake, SUCR exchange, and photosystem I. Results are presented in percentage of stem contribution to photosynthesis. A, Effect of stem contribution to photosynthesis on plant photon uptake and SUCR breakdown fluxes. B, Effect of stem contribution to photosynthesis on SUCR exchange reactions between organs. C, Effect of stem contribution to photosynthesis on the photosystem I flux. The yellow area represents the range of experimentally observed values: mean ± standard deviation.

experimentally measured organic xylem fluxes on the root to xylem export reactions as additional constraints in the model. We analyzed how these additional constraints impacted the global photon demand (Figure 6A). We observed a very low effect on photon demand since photon demand increased by 0.63% if a constraint is set on glutamine, the major organic molecule in concentration. Photon demand increased by 0.92% if a constraint was set for each metabolite measured experimentally. Therefore, experimental values do not disturb much the photon demand. This emphasized that VYTOP is a consistent model of matter exchange between plant organs.

In a second analysis, we compared the VYTOP composition of organic xylem fluxes to the experimental data

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021

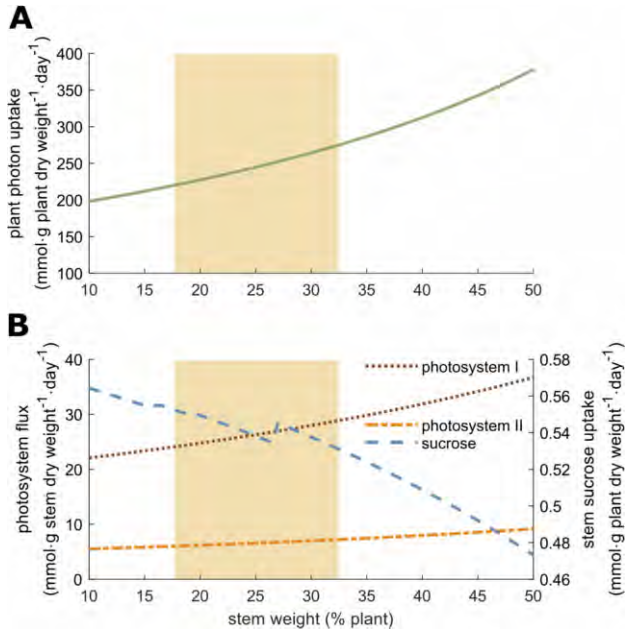



**Figure 5** Effect of stem weight on photon uptake, the stem SUCR uptake flux, and the stem photosynthesis. A, Effect of stem weight on global plant photon uptake. B, Effect of the stem weight on the stem SUCR uptake flux and the stem photosystems I and II fluxes. The yellow area represents the range of experimentally observed values: mean ± standard deviation.

(Figure 6B) with, this time, no constraints on root to xylem exports. The global features of xylem organic chemistry are well predicted by VYTOP: predominance (around 80%) of glutamine, a central metabolite found in xylem sap from diverse plant species is inferred by VYTOP. VYTOP also predicts the presence of other amino acids, in accordance with reported experimental data (Andersen and Brodbeck, 1989; Zuluaga et al., 2013; Anguita-Maeso et al., 2021). FVA results (Supplemental File S3) showed no variability for the different organic fluxes predicted, as represented in Figure 6B by the absence of error bars. FVA thus confirmed that VYTOP predicts the importance of glutamine in xylem sap, as observed by metabolomics experiments.

With only global physiological constraints, specific biomass equation, and resource use optimization (photon minimization and sum of flux minimization; Figures 1 and 2), VYTOP efficiently predicts transport of metabolites to aerial parts. This result shows that the organic composition of xylem sap is probably driven by plant physiology and resource optimization. The predominance of glutamine in both the model and experiments is remarkable. Glutamine is, with glutamate, the precursor of other amino acids and can be directly used as a C source since it is directly connected to the TCA cycle. Transporting a central metabolite, branched with both central metabolism and amino acid biosynthesis, thus appears more advantageous and robust. The predominance of glutamine over glutamate in xylem sap may be due to the acidic property of the latter, which could be deleterious at high concentrations. The presence of other amino acids is also a conserved property of xylem sap. We hypothesize that this redundancy of N sources could bring robustness to plant metabolism.

Several other organic molecules (organic acids, sugars, and ETOH) were observed experimentally but not predicted by the presented flux distribution (Figure 6B). This must be due to a certain level of exchange and porosity between compartments that exists in nature and is not taken into account in the model, such as possible exchange fluxes between xylem and phloem via stem cells. Furthermore, an ascending phloem is observed in tomato (Bonnemain, 1980), contributing to the observation of sugars in xylem sap. The presence of ETOH may reflect an $O_2$ limitation in certain tomato cells, a local phenomenon not depicted in the model. Organic molecules can also act as shuttles for some ions, such as citrate for iron (Rellán-Álvarez et al., 2010). VYTOP does not represent these complex behaviors and provides a simplified and more schematized view of matter fluxes, nevertheless consistent for around 95% of the organic xylem sap content.

## Use Case 4: VYTOP predicts the impact of nutritional limitation on the tomato growth rate

We used VYTOP to predict how a whole tomato plant responds to changes in a nutrient resource. De Groot et al. (2002) analyzed how the tomato growth rate was affected by N limitation. They monitored growth rates associated with different N supplies (Supplemental File S3), which revealed a progressive decrease of plant growth rate (Figure 7). To simulate the impact of N limitation in VYTOP, we used the N uptake flux and photon uptake computed in optimal conditions as constraints and progressively decreased the N uptake flux as in De Groot et al. (2002), upon the same photon availability. The simulation was performed with only $NO_3^-$ as the N source, as in the experimental study. We assumed as a new objective function the maximization of leaf biomass growth, with a constant root/leaf and stem/leaf ratio, obtained from our experimental results in N replete conditions. The simulation predicted a linear decrease of growth rate, which followed the same pattern as for the experiment, with very close values. For example, both VYTOP and the experimental data showed a reduction of the plant growth rate from 0.24 to around 0.13/day when the N content is reduced by 50%. Thus, VYTOP agrees with the experimentally observed impact of N limitation on tomato plant growth.

## Use Case 5: VYTOP predicts and helps to understand the physiology and metabolism of transgenic tomato lines

We assessed how efficiently VYTOP can predict the effect of a genetic perturbation on plant physiology (i.e. mutant behavior). We selected four studies presenting transgenic tomato lines. Three lines express a fragment of a TCA cycle gene in the antisense orientation, resulting in a reduced enzymatic activity of, respectively, mitochondrial citrate

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021

A

| condition | plant photon uptake mmol·g plant dry weight$^{-1}$· d$^{-1}$ | uptake increase % |
|---|---|---|
| no xylem flux constraint | 247.687 | - |
| glutamine flux constrained | 249.257 | 0.63 |
| amino acid fluxes constrained | 249.901 | 0.89 |
| organic molecule fluxes constrained | 249.977 | 0.92 |

B

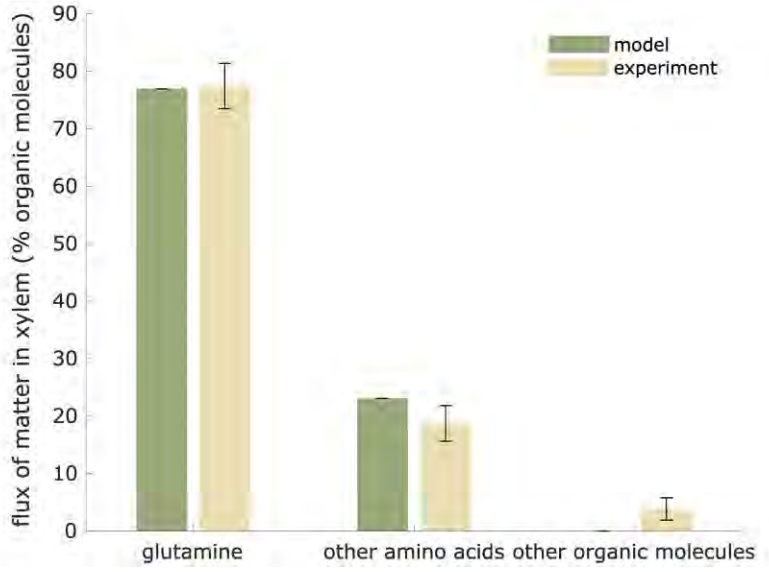



**Figure 6** Prediction of the relative composition of organic metabolites in tomato xylem fluxes by VYTOP. A, Impact of the addition of constraints on organic xylem fluxes (set from experimental data) on plant photon demand. B, Comparison of predicted fluxes by VYTOP and experimental fluxes of organic matter in xylem sap. Modeling results were obtained from a simulation without organic xylem flux constraint. Error bars indicate FVA range for the modeled distribution and mean ± standard deviation range for the experimental distribution (sample size: 33 plants).

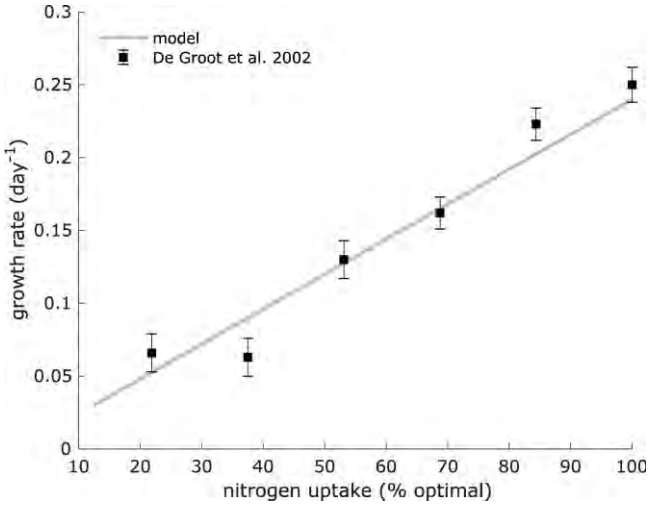



**Figure 7** Impact of N limitation on plant growth. Experimental data were taken from De Groot et al. (2002) and are represented as mean ± standard deviation (sample size: 6). Growth rates were predicted using VYTOP model with a limited N uptake flux.

synthase (mCS; Sienkiewicz-Porzucek et al., 2008), mitochondrial succinyl-CoA ligase (SCoAL; Studart-Guimarães et al., 2007) and mitochondrial alpha-ketoglutarate (AKG) dehydrogenase (AKGDH; Araújo et al., 2012). The fourth line is a SUCR phosphate synthase (SPS) gene knockout (C.M Rick TGRC https://tgrc.ucdavis.edu/Data/Acc/GenRepeater.aspx?Gene=Sus). We modeled these transgenic lines with VYTOP by forcing the matter flux to zero for the knockout line or performing simulation at 50% of the optimal flux for attenuated lines, in agreement with the activity reduction presented in the three studies. We then analyzed the impact of these changes on metabolic fluxes (Table 1).

One major experimental result on the three lines with attenuated activity of the TCA cycle is that the impact on vegetative growth and young tomato physiology appears to be weak, with slightly reduced growth and photosynthetic activity similar to the wild-type line (Studart-Guimarães et al., 2007; Sienkiewicz-Porzucek et al., 2008; Araújo et al., 2012). Similarly, using VYTOP, we found that plant growth was the same. We also analyzed photosynthesis flux predicted in the model and found only a very limited impact (< 1.5% of flux increase) compared to the wild-type model.

In the different experimental studies, authors suggested that there was a bypass to compensate the enzyme deficiencies in the transgenic plants. In line with this hypothesis, VYTOP predicted an activation of peroxisomal citrate synthase (pCS) flux for the line attenuated in mCS. The experimental study on the same line confirms this prediction since an enhanced activity of pCS is observed. Similarly, for

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021

**Table 1** Effects of transgenic tomato lines on plant physiology and metabolism and comparison between experimentally observed effects and VYTOP predictions

| Enzyme | Activity | Effect | VYTOP prediction | Consistency |
|---|---|---|---|---|
| mCS | Attenuated | Minor effect on growth | Optimal growth achieved | + |
| mCS | Attenuated | Unaltered leaf photosynthesis | $<1.5\%$ increase | + |
| mCS | Attenuated | Enhanced CSp activity | CSp activation | + |
| mCS | Attenuated | Reduction of $NO_3^-$ metabolism | No change | − |
| SCoAL | Attenuated | Minor effect on growth | Optimal growth achieved | + |
| SCoAL | Attenuated | Unaltered leaf photosynthesis | $<1.5\%$ increase | + |
| SCoAL | Attenuated | Evidences of GABA shunt use | GABA shunt activation | + |
| AKGDH | Attenuated | Unaltered leaf photosynthesis | $<1.5\%$ flux increase | + |
| AKGDH | Attenuated | Evidences of GABA shunt use | GABA shunt activation | + |
| AKGDH | Attenuated | Minor effect on growth | Optimal growth achieved | + |
| SPS | Suppressed | Lethal | Growth infeasible | + |

Experimental results for mCS, SCoAL, AKGDH, and SPS are, respectively, from Sienkiewicz-Porzucek et al. (2008), Studart-Guimarães et al. (2007), Araújo et al. (2012), and C.M Rick TGRC (https://tgrc.ucdavis.edu/Data/Acc/GenRepeater.aspx?Gene=Sus).

the lines attenuated in mSCoAL and mAKGDH activities, VYTOP predicts the activation of three reactions: gamma-amino-butyric acid (GABA)-Transaminase, Succinic Semialdehyde Dehydrogenase, and Glutamate Decarboxylase, constituting the GABA shunt. This pathway can generate succinate from AKG without the use of AKGDH and SCoAL. Interestingly, metabolomics samplings and $^{14}CO_2$ in the experimental studies also point toward the use of this pathway.

Conversely, no bypass exists after SPS gene knockout, as its silencing led to no growth in the resulting homozygous line. VYTOP also predicts this dependence as no growth can be obtained with silenced SPS flux. Thus, VYTOP seems to be an adequate tool to predict the alternative pathways that could be used to bypass a bottleneck in a transgenic line, or to predict the essentiality of some other reactions and the absence of alternative pathways.

VYTOP could thus be both a predictive tool to anticipate the effect of genetic engineering before generating a transgenic line, or a way to understand and analyze a phenotype observed in the transgenic line. However, it is worth noting that metabolism is often tightly linked with regulation, and that some effects are not predictable by VYTOP yet. As an example, the model could not predict the reduction of $NO_3^-$ metabolism observed in plants with attenuated mCS activity, and the authors suggested that it was due to post-translational regulation. The development of hybrid metabolic-regulation models is challenging, particularly because regulation is difficult to predict and reconstruct at a genome-scale level.

## Conclusion

Modeling the complex interactions between nutritional sources, genes, energetic requirements, and exchange of matter is an ongoing challenge in plant systems biology. In this study, we developed a modeling framework that integrates all these components, based on a genome-scale metabolic model and physiological characterizations. The resulting tomato model VYTOP predicts consistent flux distribution and global behavior for each organ without constraining the internal fluxes, highlighting that the methodology used is relevant to build organ-specific models. We demonstrated that VYTOP gave consistent results in five different use cases: prediction of core metabolic fluxes and exchanges, role of the stem in the plant, analysis of the organic xylem sap composition, estimation of nutrient limitation impact on growth, and impact of genetic modifications on metabolism and physiology. This tool should be useful to guide genetic/metabolic engineering of plants as well as to understand at a system level how plants respond to nutritional variations. Other modeling approaches conducted on tomato linked plant architecture and physiology to environmental variations such as temperature (Cieslak et al., 2016; Pradal et al., 2008, 2015). It will be interesting to combine these macroscopic models with VYTOP. In addition, VYTOP could be extended to other tomato plant life-cycle stages, by adding metabolic models for flower and/or fruit. Adding other metabolic models for organisms interacting with the plant, such as fungi or bacteria (Peyraud et al., 2017), is also a promising perspective to explore plant interactions with other organisms.

## Materials and methods

### Experimental procedures

#### *Plant cultures and automatic phenotyping*

Tomato seeds (*S. lycopersicum* M82) were grown in soil (SB2, Proveen, The Netherlands) supplemented with Osmocote coated fertilizer at a rate of $4\,g\,L^{-1}$. Osmocote coated fertilizer contained nitric N (5.3% w/v), ammoniacal N (6.7% w/v), 7% of phosphorus pentoxide (w/v), and 19% of potassium oxide (w/v).

Seeds were germinated in a growth chamber (26°C, 67% Relative Humidity, 12-h LED light per day). Around 100 plantlets were transplanted in individual plastic pots ($8 \times 8$ cm) 8 d after sowing. Sixteen days after, 90 young plants were chosen and repotted in 3-L pots until the end of the experiment. Foam cover discs were placed on each pot to limit physical evaporation. The plants were loaded on the Phenoserre robot facility of the Toulouse Plant Microbe Phenotyping infrastructure. Twelve-hour-light per day at 28°C and 50% humidity were programmed.

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021

All the plants were watered with 100 mL 3 times on the loading day and weighed, in order to define a well-watered target weight. The daily transpiration was determined as the weight differences between two consecutive days, at the time of watering. Transpiration rate was deduced as following: transpiration rate at Day i (mL·g[dry weight]$^{-1}$·day$^{-1}$) = transpiration at Day i (mL·day$^{-1}$)/plant dry weight at Day i (g[dry weight]). Temperature, hygrometry, and light intensity were recorded during the whole experiment.

### Collection and preparation of plant samples

Six plants were removed each day from the conveyor belt for samplings during nine consecutive days, starting 4 d after the loading on the robot. For these plants, stems were cut just above the cotyledons node, rinsed with ∼1 mL of water and the upcoming xylem sap was harvested by repeated pipetting and collection into Eppendorf tubes placed on ice. The tubes were placed at −80°C for further quantitative NMR analyses.

### Quantitative NMR analyses

The xylem saps were analyzed by 1D 1H NMR on MetaToul analytics platform (UMR5504, UMR792, CNRS, INRAE, INSA 135 Avenue de Rangueil 31077 Toulouse Cedex 04, France), using the Bruker Avance 800 MHz equipped with an ATMA 5-mm cryoprobe. Each xylem sap sample was centrifuged to remove the residues (5 min, 13520 RCF, Hettich Mikro 200 centrifuge), then placed in 3-mm NMR tubes. TSP-d4 standard (Sodium 3-(trimethylsilyl)(1-13C,2H4)propanoate) was used as a reference. The pH 6.0 phosphate buffer was used to standardize the chemical shifts among samples. Acquisition conditions were as follows: 30° pulse angle, 20.0287 ppm spectral width, 64 scans per acquisition for a total scan time of ∼8 min per sample, and zgpr30 water presaturation sequence. The samples were kept at a temperature of 280 K (6.85°C) all along the analysis. Resonances of metabolites were manually integrated and the concentrations were calculated based on the number of equivalent protons for each integrated signal and on the TSP final concentration.

Xylem sampling and NMR analyses only give the concentration of a metabolite (in mmol·mL$^{-1}$) at a given time. However, xylem sap is a continuous flow. Using the transpiration of the plant, and assuming a constant concentration of the metabolite in the xylem on a 24-h basis (validated throughout our experiment, see "Model framework of VYTOP is validated by experimental data"), the flow of metabolite per day can be estimated using the following formulae:

Metabolite flux (in mmol·g dry weight$^{-1}$·day$^{-1}$) = Metabolite concentration (in mmol·mL$^{-1}$)· transpiration rate (mL·g dry weight$^{-1}$.day$^{-1}$).

### Biochemical analyses of metabolites

The different organs of each plant were collected separately (stems, leaves, and roots) on another experiment performed on the same platform and plant variety and close plant age (4-week old). Approximately 300 mg fresh weight for each collected organ was frozen in liquid N and stored at −20°C for further biochemical analyses.

Quantifications of metabolites were performed at the High-Throughput Metabolic phenotyping (HiTMe) platform (INRAE-IBVM-71 avenue E. Bourlaux-CS 20032-33882 Villenave d'Ornon Cedex, France). The plants samples, previously frozen in liquid N, were ground to a powder using liquid N to avoid thawing. A quantity of 20 ± 10 mg of each was weighted in previously frozen Micronic tubes. Free amino acids, GLC, fructose, malate, proteins, starch, SUCR, and chlorophylls in leaves, stems, and roots were quantified as described in (Biais et al., 2014). Briefly, ethanolic extracts from every sample were obtained using three consecutive incubations of the frozen ground powder aliquots. ETOH 80% v/v with HEPES/KOH 10 mM pH 6 buffer was used for the first two incubations, and ETOH 50% v/v with HEPES/KOH 10 mM pH 6 buffer was used for the third. Supernatants were pooled and used for the quantification of chlorophylls, GLC, fructose, SUCR, malate, and free amino acids. Pellets were used for the determination of protein and starch contents. The extracts and pellets were stored at −20°C between each quantification. For each sample, chlorophylls were quantified by measuring optical densities at 645 and 665 nm on a mix of 50 µL of extract supplemented with 150 µL of analytics grade ETOH. Amino acids were quantified using the fluorescamine method. Excitation wavelength was 405 nm and emission was measured at 485 nm. The proteins were quantified using Bradford reagent. Starch was quantified in GLC equivalent after full pellet digestion in an oven at 37°C for 18 h. For the other analytes cited above, the NADH/NADPH appearance was measured, and the analytes were quantified using a 1:1 stoichiometric coefficient.

## In silico procedures

### Tomato metabolic network curation and conversion

The tomato metabolic model iHY3410 published by Yuan et al. (2016) was used as a starting point. The authors provided the tomato metabolic network in two formats: table (.xlsx) and SBML (Hucka et al., 2003). Some reaction directions were different in the SBML and table files, and some metabolites (proton and water) were missing in the table file. Thus, we performed a first curation step to merge the files and select the appropriate reaction directions.

Then, we converted the metabolic network IDs (metabolites and reactions) to BiGG nomenclature. The advantage of BiGG nomenclature is that it is usually more explicit for the analysis as the IDs are usually an abbreviation of the metabolite/reaction names, while in MetaCyc ID it is often a generic acronym (CPD for metabolites, RXN for reactions) and a number. The web tool Semi Automatic Metabolic Identifier Reconciliation (Peyraud et al., 2016) was used to find the appropriate BiGG IDs for metabolites and reactions. The correspondences between MetaCyc and BiGG IDs are provided in the final metabolic network in table format

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021

(Supplemental File S1). For some metabolites and reactions, no BiGG ID was found even after manual searches in the BiGG database, probably because they are specific to plants. New BiGG IDs following the usual nomenclature were created for 1,094 reactions and 758 metabolites (see Supplemental File S1). The SBML file (Hucka et al., 2003) was generated using Met4j (https://forgemia.inra.fr/metexplore/met4j) and ModelPolisher (Römer et al., 2016).

The resulting metabolic model was tested on autotrophic growth and heterotrophic growth (using FBA, see "Computational Simulations") on different organic substrates before generating a multi-organ metabolic model. Growth was achievable with $CO_2$ + light, starch, SUCR, GLC, fructose, glycerol, glutamate, glutamine, AKG, FUM, asparagine, alanine, and isocitrate as C sources. Other organic (glutamine, glutamate, aspartate, asparagine, alanine, proline, histidine, leucine, methionine, lysine, and cysteine) and inorganic ($NO_3^-$ and $NH_4^+$) elements can be used as N sources. No catabolic reactions were found for the assimilation of lysine, ETOH, isoleucine, beta-alanine, leucine, and cysteine as C sources whereas they were experimentally observed in the xylem in amount superior to biomass assimilation, suggesting that they could be catabolized. Thus, we searched the catabolic pathways of these amino acids in plants and incorporated them in the network, using BLAST to find the orthologous proteins in tomato. Twenty-one reactions, linked to 34 genes were added to the model iHY3410. Fifteen reactions had a Gene-Protein-Reaction (GPR) complete with an e-value superior to 5e-131 (Supplemental File S1), 6 had none. After this second curation step, we computed the mass balance (on C, N, phosphate, sulfur, and $O_2$) and manually curated the reactions with an incorrect C, N, phosphate, sulfur, and $O_2$ balance. We generated a final tomato metabolic network Sl2183. The model is available in MetExplore (Cottret et al., 2018) https://metexplore.toulouse.inrae.fr/metexplore2/?idBioSource=6353, in the database BioModels under the ID MODEL2111120001 (Glont et al., 2018) and in a github repository https://github.com/lgerlin/slyc-metabolic-model/.

### VYTOP construction

To generate VYTOP (see flow chart in Supplemental Figure S5), we built organ-specific metabolic models for leaf, stem, and root. To this end, organ growth rates and organ-specific biomass equations were determined from our experiments and implemented as constraints in the simulations. We generated an in-house script to parse the metabolic network Sl2183 (SBML file) and triplicated it to represent three organs. In the generated model, each organ is represented as a "metacompartment": metabolite and reaction IDs have a final letter (l, s and r for leaf, stem and root, respectively) to indicate their organ location (e.g. R_NAD2 reaction becomes R_NAD2_l, R_NAD2_s, R_NAD2_r and M_gln_L_c becomes M_gln_L_c_l, M_gln_L_c_s, M_gln_L_c_r).

Our in-house script also generates exchange reactions to represent transfers between organs and exchange compartments (represented with the letters xyl for xylem and phl for phloem), such as 1 M_gln_L_c_r $\rightarrow$ *n* M_gln_L_xyl and 1 M_gln_L_xyl $\rightarrow$ *m* M_gln_L_c_l. The list of metabolites authorized to be exchanged in xylem and phloem respectively is available in Supplemental File S1. *n* and *m* represent the mass ratios between the organs and the whole plant, in order to take into account the different weights of organs and have quantitative predictions at the whole plant level. For example, for transfer from the root to the xylem, 1 $\mu mol.g^{-1}$ root dry weight per day will generate $n = 1/(1 + 1.52 + 3.37) = 0.1697\ \mu mol.g^{-1}$ plant dry weight per day in xylem, and then 1 $\mu mol.g^{-1}$ plant dry weight per day in xylem will generate $m = (1 + 1.52 + 3.37)/3.37 = 1.74\ \mu mol.g^{-1}$ leaf dry weight per day. Finally, a transport cost (ATP cost) of 0.5 was added in the model: transporting k $\mu mol.g^{-1}$ organ dry weight per day would consume k·(transport cost) $\mu mol.g^{-1}$ organ dry weigh per day of ATP. Performing several simulations of Use Case 3 with different values of ATP transport cost, we found that a transport cost between 0.02 and 1.0 mol of ATP per mol of exchanged molecule, which appears reasonable, provided the most consistent predictions in agreement with experimental data, while use Case 4 is not impacted by the different sets of values tested (Supplemental Text).

Physiological constraints were added to impose different physiological roles in organs.

i. Uptake of minerals and water is not allowed in leaf and stem, while it is allowed with no limitation in roots (assuming they are not limiting factors for well-watered plants and adequate nutritional supply in the soil).
ii. Photon uptake was allowed in leaf and stem but not in root.
iii. The $NH_4^+$:$NO_3^-$ uptake ratio can be constrained. Few data are available to predict the balance between $NH_4^+$ and $NO_3^-$ in root uptake. We analyzed the impact of varying the parameter in Supplemental Text. The ratio was not constrained in our simulations, excepted use Case 4, to agree with the experimental data.
iv. Photosynthesis is limited in stems by its reduced exchange surface compared to leaves (Hetherington et al., 1998). We then introduced a ratio (leaf/stem) of contribution to photosynthesis, to limit the ability of the stem to perform photosynthesis: photosynthesis in stems is constrained to be inferior or equal to photosynthesis in leaves/surface ratio.
v. Photorespiration was modeled by imposing a ratio between rubisco carboxylase and oxygenase fluxes. The ratio was set to 85, which represents the mean ratio of enzymatic specificity for $CO_2$ versus $O_2$ for C3 plants, according to experimental data gathered by Tcherkez et al. (2006).

Finally, ATP maintenance, which represents the global cost in a cell for nonmetabolic processes (such as housekeeping functions), was kept at the value used by Yuan

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021

et al. (2016) as no organ-specific ATP maintenance values are available.

### Computational simulations

Modeling of whole tomato plant metabolism was performed using constrained-based modeling, with the FBA methodology (Orth et al., 2010). Briefly, FBA relies on (1) the QSSA in the biological system modeled, (2) the formulation of a biologically relevant objective: flux(es) to minimize/maximize, and (3) eventually additional physiological constraints. The solution of FBA will be the optimal matter fluxes (regarding the objective) in a metabolic network assuming QSSA and respecting the constraints. Use of QSSA in a whole tomato plant is justified by our experimental results (see part "Model framework of VYTOP is validated by experimental data" of the "Results and Discussion"). We decided to define the minimization of total photon uptake as our objective. The photon uptake flux obtained from this simulation was then integrated as a new constraint in the model, and FBA was run a second time with the objective of minimizing the sum of absolute fluxes. The use of these two objectives is discussed in part "Model framework of VYTOP is validated by experimental data" of "Results and Discussion".

FBA was run with or without the integration of experimentally measured organic xylem fluxes, and with varying stem proportions, surface ratios, and transport costs.

FVA (Mahadevan and Schilling, 2003) was also performed. This extension of the FBA aims at finding alternative solutions of the one generated by FBA, as its flux distribution is usually not unique. It consists of imposing the optimal objective values found in FBA (here both photon uptake and sum of absolute fluxes) as additional constraints, and determining the minimum and maximum flux that can carry each reaction in these conditions.

Simulations were performed with Python version 3.5 scripts, the open access libraries lxml, pandas and the linear programming solver CPLEX Python API, developed by IBM and free for academic institutions. Scripts are available at https://github.com/lgerlin/slyc-metabolic-model.

### Accession numbers

The tomato genome used in this article is ITAG version 4.0 and can be found on the Sol Genomic Networks website (ftp://ftp.solgenomics.net/tomato_genome/annotation/ITAG4.0_release/).

## Supplemental data

The following materials are available in the online version of this article.

**Supplemental Figure S1.** Additional physiological data.

**Supplemental Figure S2.** Additional xylem metabolomics data.

**Supplemental Figure S3.** Additional organ metabolomics data.

**Supplemental Figure S4.** Additional representation of model responses to plant physical variations.

**Supplemental Figure S5.** Flow chart of the multi-organ modeling pipeline.

**Supplemental Text.** Additional analyses (effects of ATP cost variation, Impact of physiological changes on the plant at high percentages, Impact of the uptake ratio between nitrates and ammonium, Stoichiometric Balance Cycles).

**Supplemental File S1.** Table of tomato plant genome-scale metabolic network Sl2183.

**Supplemental File S2.** SBML file of tomato plant genome-scale metabolic network Sl2183, SBML Level 3 version 1.

**Supplemental File S3.** Table with modeling results.

## Acknowledgments

We thank Toulouse Plant Microbe Phenotyping (TPMP) platform (Castanet-Tolosan, France) and its staff Nemo Peeters, Felicià Maviane Macia, and Fabrice Devoilles for their technical support in plant cultures and imaging, as well as HiTMe platform (Villenave d'Ornon, France) for biochemical analysis of organ metabolites. We acknowledge MetaToul (Metabolomics & Fluxomics Facitilies, Toulouse, France, www.metatoul.fr) platform, which is part of the MetaboHUB-ANR-11-INBS-0010 national infrastructure (www.metabohub.fr) and its staff Cécilia Berges, Edern Cahoreau, and Lindsay Peyriga for access to NMR facilities. We also thank Sophie Colombié for helpful discussions and feedbacks on the manuscript.

## Data availability

All data supporting the findings of this study are available within the paper and within its supplementary materials published online.

## Funding

L.G. and A.E. were, respectively, funded by a grant from the French Ministry of Higher Education and Research and by the French Laboratory of Excellence TULIP (ANR-10-LABX-41 and ANR-11-IDEX-0002-02). The study was funded by the French Laboratory of Excellence (LABEX) project TULIP (ANR-10-LABX-41 and ANR-11-IDEX-0002-02). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

*Conflict of interest statement:* None declared.

## References

**Anguita-Maeso M, Haro C, Montes-Borrego M, De La Fuente L, Navas-Cortés JA, Landa BB** (2021) Metabolomic, Ionomic and Microbial Characterization of Olive Xylem Sap Reveals Differences According to Plant Age and Genotype. Agronomy **11:** 1179

**Araújo WL, Tohge T, Osorio S, Lohse M, Balbo I, Krahnert I, Sienkiewicz-Porzucek A, Usadel B, Nunes-Nesi A, Fernie AR** (2012) Antisense inhibition of the 2-oxoglutarate dehydrogenase complex in tomato demonstrates its importance for plant respiration and during leaf senescence and fruit maturation. Plant Cell **24**: 2328–2351

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021

**Arnold A, Nikoloski Z** (2014) Bottom-up metabolic reconstruction of arabidopsis and its application to determining the metabolic costs of enzyme production. Plant Physiol **165**: 1380–1391

**Arnold A, Sajitz-Hermstein M, Nikoloski Z** (2015) Effects of varying nitrogen sources on amino acid synthesis costs in *Arabidopsis thaliana* under different light and carbon-source conditions. PLoS One **10**: 1–22

**Bénard C, Bernillon S, Biais B, Osorio S, Maucourt M, Ballias P, Deborde C, Colombié S, Cabasson C, Moing A, et al.** (2015) Metabolomic profiling in tomato reveals diel compositional changes in fruit affected by source-sink relationships. J Exp Bot **66**: 3391–3404

**Biais B, Bénard C, Beauvoit B, Colombié S, Prodhomme D, Ménard G, Bernillon S, Gehl B, Gautier H, Ballias P, et al.** (2014) Remarkable reproducibility of enzyme activity profiles in tomato fruits grown under contrasting environments provides a roadmap for studies of fruit metabolism. Plant Physiol **164**: 1204–21

**Bonnemain JL** (1980) Microautoradiography as a tool for the recognition of phloem transport. Bericht Deutsch Bot Gesellsch **93**: 99–107

**Caspi R, Billington R, Ferrer L, Foerster H, Fulcher CA, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Karp PD et al.** (2016) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res **44**: D471–D480

**Cieslak M, Cheddadi I, Boudon F, Baldazzi V, Génard M, Godin C, Bertin N** (2016) Integrating physiology and architecture in models of fruit expansion. Front Plant Sci **7**: 1739

**Clark TJ, Guo L, Morgan J, Schwender J** (2020) Modeling plant metabolism: from network reconstruction to mechanistic models. Ann Rev Plant Biol **71:** 303–326

**Collakova E, Yen JY, Senger RS** (2012) Are we ready for genome-scale modeling in plants? Plant Sci **191–192**: 53–70

**Colombié S, Beauvoit B, Nazaret C, Bénard C, Vercambre G, Le Gall S, Biais B, Cabasson C, Maucourt M, Gibon Y, et al.** (2017) Respiration climacteric in tomato fruits elucidated by constraint-based modelling. New Phytologist **213**: 1726–1739

**Colombié S, Nazaret C, Bénard C, Biais B, Mengin V, Solé M, Fouillen L, Dieuaide-Noubhani M, Mazat JP, Gibon, Y, et al.** (2015) Modelling central metabolic fluxes by constraint-based optimization reveals metabolic reprogramming of developing *Solanum lycopersicum* (tomato) fruit. Plant J **81**: 24–39

**Cottret L, Frainay C, Chazalviel M, Cabanettes F, Gloaguen Y, Camenen E, Merlet B, Heux S, Portais JC, Jourdan F, et al.** (2018) MetExplore: collaborative edition and exploration of metabolic networks. Nucleic Acids Res **46**: W495–W502

**De Groot CC, Marcelis LFM, Van den Boogaard R, Lambers H** (2002) Interactive effects of nitrogen and irradiance on growth and partitioning of dry mass and nitrogen in young tomato plants. Funct Plant Biol **29**: 1319–1328

**de Groot MJL, Daran-Lapujade P, van Breukelen B, Knijnenburg TA, de Hulster EAF, Reinders MJT, Pronk JT, Hec AJR, Slijper M** (2007) Quantitative proteomics and transcriptomics of anaerobic and aerobic yeast cultures reveals post-transcriptional regulation of key cellular processes. Microbiology **153**: 3864–3878

**Gerlin L, Frainay C, Jourdan F, Baroukh C, Prigent S** (2021) Plant genome-scale metabolic networks. Advances in Botanical Research, Vol. **98**. Academic Press Inc., Cambridge, MA, pp 237–270

**Glont M, Nguyen TVN, Graesslin M, Hälke R, Ali R, Schramm J, Wimalaratne SM, Kothamachu VB, Rodriguez N, Hermjakob H, et al.** (2018) BioModels: expanding horizons to include more modelling approaches and formats. Nucleic Acids Res **46**: D1248–D1253

**Gomes de Oliveira Dal'Molin C, Quek LE, Saa PA, Nielsen LK** (2015) A multi-tissue genome-scale metabolic modeling framework for the analysis of whole plant systems. Front Plant Sci **6**: 12

**Grafahrend-Belau E, Junker A, Eschenröder A, Müller J, Schreiber F, Junker BH** (2013) Multiscale metabolic modeling: dynamic flux balance analysis on a whole-plant scale. Plant Physiol **163**: 637–647

**Gu C, Kim GB, Kim WJ, Kim HU, Lee SY** (2019) Current status and applications of genome-scale metabolic models. Genome Biol **20**: 121

**Hetherington SE, Smillie RM, Davies WJ** (1998) Photosynthetic activities of vegetative and fruiting tissues of tomato. J Exp Bot **49**: 1173–1181

**Hijaz F, Killiny N** (2014) Collection and chemical composition of phloem sap from Citrus sinensis L. Osbeck (sweet orange). PLoS One **9**: 1–11

**Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Wang J, et al.** (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics **19**: 524–531

**King ZA, Lu J, Dräger A, Miller P, Federowicz S, Lerman JA, Ebrahim A, Palsson BO, Lewis NE** (2016) BiGG models: a platform for integrating, standardizing and sharing genome-scale models. Nucleic Acids Res **44**: D515–D522

**Li Y, Chen Y, Zhou L, You S, Deng H, Chen Y, Alseekh S, Yuan Y, Fu R, Zhang Y, et al.** (2020) MicroTom metabolic network: rewiring tomato metabolic regulatory network throughout the growth cycle. Mol Plant **13**: 1203–1218

**Mahadevan R, Schilling CH** (2003) The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab Eng **5**: 264–276

**Mintz-Oron S, Meir S, Malitsky S, Ruppin E, Aharoni A, Shlomi T** (2012) Reconstruction of Arabidopsis metabolic network models accounting for subcellular compartmentalization and tissue-specificity. Proc Natl Acad Sci **109**: 339–344

**Moreira TB, Shaw R, Luo X, Ganguly O, Kim HS, Coelho LGF, Cheung CYM, Williams TC R.** (2019) A genome-scale metabolic model of soybean (*Glycine max*) highlights metabolic fluxes in seedlings. Plant Physiol **180**: 1912–1929

**Norsigian CJ, Pusarla N, McConn JL, Yurkovich JT, Dräger A, Palsson BO, King Z** (2020) BiGG Models 2020: multi-strain genome-scale models and expansion across the phylogenetic tree. Nucleic Acids Res **48**: D402–D406

**Orth JJD, Thiele I, Palsson BBØ** (2010) What is flux balance analysis? Nat Biotechnol **28**: 245–248

**Peyraud R, Cottret L, Marmiesse L, Gouzy J, Genin S** (2016) A resource allocation trade-off between virulence and proliferation drives metabolic versatility in the plant pathogen Ralstonia solanacearum. PLoS Pathog **12**: e1005939

**Peyraud R, Dubiella U, Barbacci A, Genin S, Raffaele S, Roby D** (2017) Advances on plant-pathogen interactions from molecular toward systems biology perspectives. Plant J **90**: 720–737

**Pfau T, Christian N, Masakapalli SK, Sweetlove LJ, Poolman MG, Ebenhöh O** (2018) The intertwined metabolism during symbiotic nitrogen fixation elucidated by metabolic modelling. Sci Rep **8**: 12504

**Pradal C, Dufour-Kowalski S, Boudon F, Fournier C, Godin C** (2008) OpenAlea: a visual programming and component-based software platform for plant modelling. Funct Plant Biol **35**: 751

**Pradal C, Fournier C, Valduriez P, Cohen-Boulakia S** (2015) OpenAlea: scientific workflows combining data analysis and simulation. ACM International Conference Proceeding Series. Association for Computing Machinery, New York, NY, pp 1–6

**Rellán-Álvarez R, Giner-Martínez-Sierra J, Orduna J, Orera I, Rodríguez-Castrilln JÁ, García-Alonso JI, Abadía J, Álvarez-Fernández A** (2010) Identification of a tri-Iron(III), tri-citrate complex in the xylem sap of iron-deficient tomato resupplied with iron: new insights into plant iron long-distance transport. Plant Cell Physiol **51**: 91–102

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021

**Richelle A, Joshi C, Lewis NE** (2019) Assessing key decisions for transcriptomic data integration in biochemical networks. PLoS Computat Biol **15**: e1007185

**Robaina-Estévez S, Daloso DM, Zhang Y, Fernie AR, Nikoloski Z** (2017) Resolving the central metabolism of Arabidopsis guard cells. Sci Rep **7**: 1–13

**Robaina Estévez S, Nikoloski Z** (2014) Generalized framework for context-specific metabolic model extraction methods. Front Plant Sci **5**: 491

**Römer M, Eichner J, Dräger A, Wrzodek C, Wrzodek F, Zell A** (2016) ZBIT bioinformatics toolbox: a web-platform for systems biology and expression data analysis. PLoS One **11**: e0149263

**Royer M, Larbat R, Le Bot J, Adamowicz S, Robin C** (2013) Is the C:N ratio a reliable indicator of C allocation to primary and defence-related metabolisms in tomato? Phytochemistry **88**: 25–33

**Saha R, Suthers PF, Maranas CD** (2011) Zea mays irs1563: a comprehensive genome-scale metabolic reconstruction of maize metabolism. PLoS One **6**: e21784

**Scheunemann M, Brady SM, Nikoloski Z** (2018) Integration of large-scale data for extraction of integrated Arabidopsis root cell-type specific models. Sci Rep **8**: 1–15

**Seaver SMD, Lerma-Ortiz C, Conrad N, Mikaili A, Sreedasyam A, Hanson AD, Henry CS** (2018) PlantSEED enables automated annotation and reconstruction of plant primary metabolism with improved compartmentalization and comparative consistency. Plant J **95**: 1102–1113

**Shaw R, Cheung CYM** (2018) A dynamic multi-tissue flux balance model captures carbon and nitrogen metabolism and optimal resource partitioning during Arabidopsis growth. Front Plant Sci **9**: 1–15

**Shaw R, Cheung CYM** (2020) Multi - tissue to whole plant metabolic modelling. Cell Mol Life Sci **77**: 489–495

**Shaw R, Maurice Cheung CY** (2019) A mass and charge balanced metabolic model of Setaria viridis revealed mechanisms of proton balancing in C4 plants. BMC Bioinform **20**: 1–11

**Sienkiewicz-Porzucek A, Nunes-Nesi A, Sulpice R, Lisec J, Centeno DC, Carillo P, Leisse A, Urbanczyk-Wochniak E, Fernie AR** (2008) Mild reductions in mitochondrial citrate synthase activity result in a compromised nitrate assimilation and reduced leaf pigmentation but have no effect on photosynthetic performance or growth. Plant Physiol **147**: 115–127

**Studart-Guimarães C, Fait A, Nunes-Nesi A, Carrari F, Usadel B, Fernie AR** (2007) Reduced expression of succinyl-coenzyme a ligase can be compensated for by up-regulation of the γ-aminobutyrate shunt in illuminated tomato leaves. Plant Physiol **145**: 626–639

**Sweetlove LJ, George Ratcliffe R** (2011) Flux-balance modeling of plant metabolism. Front Plant Sci **2**: 1–10.

**Tcherkez GGB, Farquhar GD, Andrews TJ** (2006) Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. Proc Natl Acad Sci USA **103**: 7246–7251

**Yuan H, Cheung CYM, Poolman MG, Hilbers PAJJ, van Riel NAWW** (2016) A genome-scale metabolic network reconstruction of tomato (*Solanum lycopersicum* L.) and its application to photorespiratory metabolism. Plant J **85**: 239–304

**Zuluaga AP, Puigvert M, Valls M** (2013) Novel plant inputs influencing Ralstonia solanacearum during infection. Front Microbiol **4**: 1–7

Downloaded from https://academic.oup.com/plphys/advance-article/doi/10.1093/plphys/kiab548/6432047 by ETH Zürich user on 21 December 2021

## 3.3 Additional discussions

### 3.3.1 Modeling whole plant metabolism

Building a multi-organ metabolic model such as VYTOP raised numerous issues at each different steps during the model development and calibration. The genome-scale reconstruction of plant metabolism into an *in silico* network is already a challenge by itself. Plant metabolic networks are usually less well reconstructed than bacteria and yeasts, notably due to the difficulty to apprehend its high compartmentation and to the limited knowledge on some (sometimes-central) pathways. In addition, several plant genes possess numerous copies, due to genome duplication events, that complicate the establishment of GPR relationship. In addition, an important diversity of secondary metabolites is present, and their complexity limits the use of short and readable IDs. Facing this complexity, an alternative would be to use core metabolic networks, which contain only central pathways and reactions well-described, such as the network developed by Arnold & Nikoloski (2014) for Arabidopsis. It would have probably been sufficient for our study on xylem fluxes and our integration of simple physiological constraints. However, this approach could be limiting in further applications of VYTOP. Indeed, experimental study of plant – pathogen interaction could reveal evidences of the production of complex secondary metabolites, which could be impossible to link to metabolism in a too reduced network. Genome-scale metabolic networks can be seen as a summary of all the biochemical reactions that can occur in an organism (knowledge databases), and on this point of view, even imperfect parts of it, such as pathways with gaps or uncertainties, should be kept. An ideal perspective would be to combine carefully reconstructed and verified central reactions (using core metabolic networks), with large genome-scale networks for the pathways still poorly characterized.

Transposing a simple cell network into a complex multi-cellular and multi-tissue organism is also an ambitious step. The division on simple and physiological roles, as in VYTOP and in previous multi-organ metabolic models, of course remains a process that could reveal too naïve to investigate some biological questions. However, in this case, the modeler will focus on developing the tissue that is of interest for their study with a higher degree of precisions, such as xylem here.

Finally, a challenging question is the diurnal change in plant metabolism. Here, we limited the study to a static model representing an average of the metabolic fluxes in a

24-hour period that remains satisfying for an estimation of the global plant behavior during vegetative growth. A dynamic version of VYTOP could be developed, for example using a dFBA framework (see Chapter 2, section 2.3.3) and optimizing growth upon different time intervals. Following this approach, photon uptake would be allowed only for 12 hours, along with starch storage, while in the next 12 hours only consumption of the starch would be allowed. This requires several additional experimental data (e.g. measurement of the evolution of starch). Similarly, xylem flow is reduced by night (Windt et al., 2006) and should vary in a similar way in the dynamic model.

### 3.3.2 Xylem content under physiological and nutritional constraints

An important insight gained from VYTOP is that physiological constraints are sufficient to predict the predominance of glutamine and the presence of additional amino acids in xylem, by adding transport cost between organ that we assumed reasonable (between 0.1 and 1 mol of ATP per mol exchanged). Hence, VYTOP offers the possibility to connect global constraints such as organ behavior and external resource availability to internal fluxes of matter in xylem vessels.

Predicting the link between xylem fluxes and global constraints appears of great interest to understand the nutrition of xylem-colonizing pathogens *in planta*. Indeed, varying the ammonium:nitrate ($NH_4^+$:$NO_3^-$) uptake ratio influences greatly the presence of organic elements, and notably glutamine, in xylem. Bialczyk et al. (2004) found an important decrease of global amino acid content and glutamine when $NH_4^+$ supply is reduced and $NO_3^-$ increased (Table 3-1). For example, glutamine is reduced by 16% from full $NH_4^+$ to 4:1 ratio of $NH_4^+$:$NO_3^-$, and by 70% from full $NH_4^+$ to full $NO_3^-$. The same pattern is obtained qualitatively in VYTOP (**Figure 3-1**): glutamine and other organic elements are at very low amounts or absent when $NO_3^-$ uptake is superior to $NH_4^+$, which likely mimics a soil with high $NO3^-$ and low $NH_4^+$.

**Table 3-1: Amino acid content (µmol·L$^{-1}$) of xylem sap for tomato plants grown with different nitrogen source, from Bialczyk et al. (2004).**

| | **$NH_4^+$ only** | **$NH_4^+$:$NO_3^-$ 4:1** | **$NH_4^+$:$NO_3^-$ 1:4** | **$NO_3^-$ only** |
|---|---|---|---|---|
| Glutamine | 815 | 684 | 320 | 247 |
| Amino acids | 1304 | 1104 | 616 | 491 |

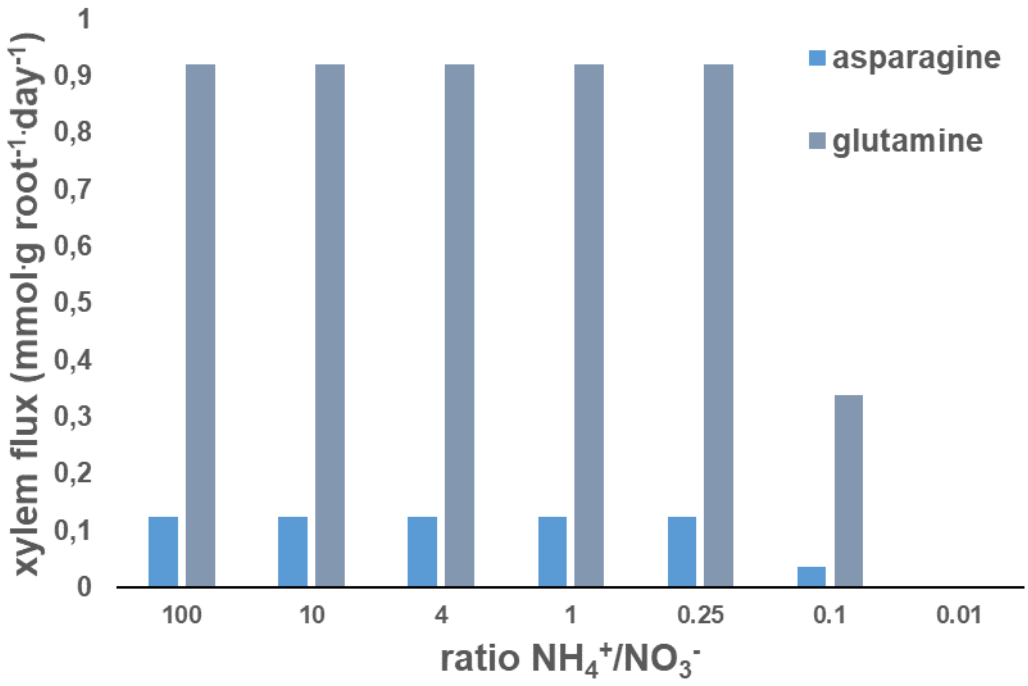



**Figure 3-1: Effect of $NH_4^+$:$NO_3^-$ uptake ratio in root on xylem fluxes.**

Uptake reactions of NH4+ and NO3- were constrained in the root part of VYTOP to be proportional at different ratio, from 100 to 0.01 ($NH_4^+$:$NO_3^-$)

Interestingly, in a study of bacterial wilt in tomato upon different nitrogen supply (Kelman, 1950), the author demonstrated that disease development by *R. solanacearum* was the slowest with a supply on $NO_3^-$ only, while the highest proportion of $NH_4^+$ lead to the most rapid disease development. Preliminary experiments performed in my team also showed that the nitrogen concentration in the nutritive solution used for tomato growth correlated negatively with the disease development (slow disease development at high nitrogen), and these aspects should be further investigated.

Nitrogen source and concentration could impact disease developments by alternative mode of actions, such as impacting the level of defenses (NO and reactive nitrogen species or defense associated enzymes) or the balance of amino acids involved in response to stresses, as reviewed by Fagard et al. (2014). However, the findings presented above suggest that plant nutrition/xylem fluxes relationship could be an important driver of xylem-colonizing pathogen nutrition *in planta*.

# Chapter 4
# Physiological and metabolic characterization of the *Ralstonia solanacearum* – tomato interaction

## 4.1 Toward a plant – pathogen metabolic model

To study metabolic interaction between a plant (tomato) and a pathogen (*R. solanacearum*), a metabolic model of *R. solanacearum* should be linked with the multi-organ tomato metabolic model VYTOP. However, VYTOP presented in Chapter 3 was calibrated using physiological data acquired on healthy tomato plants. Plant physiology is strongly disrupted by the pathogen colonization, provoking wilting symptoms. Thus, the data acquired on healthy tomato plants are not adequate to model an advanced stage of the interaction. Furthermore, some additional data are mandatory to perform quantitative flux modeling, such as pathogen/plant weight ratio, pathogen growth rate and evolution of substrate availability during plant colonization. Hence, before developing the plant – pathogen metabolic model, it was important to acquire experimental data of the interaction, necessary for the model development and calibration.

## 4.2 Physiological and metabolic characterizations of *Ralstonia solanacearum* – tomato interaction [article]

**Article**

Unravelling physiological signatures of tomato bacterial wilt and xylem metabolites exploited by *Ralstonia solanacearum*

**Journal**

*Environmental Microbiology* (published online 19 April 2021)

**Contributions**

- Performing the phenotypic experiments (with A. Escourrou) and analyzing the experimental data (phenotypic experiments, metabolomics) (with A. Escourrou and C. Baroukh)
- Writing the complete draft article and designing the figures, except: Material & Methods (written with A. Escourrou), Fig. S1 S2 S3 (F. Maviane Macia & N. Peeters) and Fig. S10 (C. Baroukh & A. Escourrou)
- Editing the article (with the co-authors)

**Supplementary Materials** available in **Appendix III**.

**Summary**

In the study, we quantified the evolution of several physiological markers (e.g. transpiration, dry weight) and found that there was a drastic change of plant physiology beyond a tipping point in the plant physiological markers. We also followed the evolution of xylem sap content during the infection. This suggested an important consumption of glutamine and asparagine during bacterial proliferation in xylem vessels and unraveled several metabolic markers of the infection. Finally, we analyzed the phenotype of a *R. solanacearum* mutant that cannot grow on xylem sugars, and found that it was only slightly delayed in proliferation and virulence.

environmental microbiology

Environmental Microbiology (2021) **00**(00), 00–00 doi:10.1111/1462-2920.15535

Special Issue Article

# Unravelling physiological signatures of tomato bacterial wilt and xylem metabolites exploited by *Ralstonia solanacearum*

**Léo Gerlin,[1] Antoine Escourrou,[1] Cédric Cassan,[2,3] Felicià Maviane Macia,[1,4] Nemo Peeters,[1,4] Stéphane Genin[1] and Caroline Baroukh[1*]**

[1]*LIPME, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France.*

[2]*UMR BFP, Université de Bordeaux, INRAE, 33882 Villenave d'Ornon, France.*

[3]*Bordeaux Metabolome, MetaboHUB, PHENOME-EMPHASIS, 33140 Villenave d'Ornon, France.*

[4]*Toulouse-Plant-Microbe-Phenotyping (TPMP), PHENOTOUL PHENOME-EMPHASIS LIPME, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France.*

**The plant pathogen *Ralstonia solanacearum* uses plant resources to intensely proliferate in xylem vessels and provoke plant wilting. We combined automatic phenotyping and tissue/xylem quantitative metabolomics of infected tomato plants to decipher the dynamics of bacterial wilt. Daily acquisition of physiological parameters such as transpiration and growth were performed. Measurements allowed us to identify a tipping point in bacterial wilt dynamics. At this tipping point, the reached bacterial density brutally disrupts plant physiology and rapidly induces its death. We compared the metabolic and physiological signatures of the infection with drought stress, and found that similar changes occur. Quantitative dynamics of xylem content enabled us to identify glutamine (and asparagine) as primary resources *R. solanacearum* consumed during its colonization phase. An abundant production of putrescine was also observed during the infection process and was strongly correlated with *in planta* bacterial growth. Dynamic profiling of xylem metabolites confirmed that glutamine is the favoured substrate of *R. solanacearum*. On the other hand, a triple mutant strain unable to metabolize glucose, sucrose and fructose appears to be only weakly reduced for *in planta* growth and pathogenicity.**

Received 9 November, 2020; revised 14 April, 2021; accepted 15 April, 2021. *For correspondence. E-mail caroline.baroukh@inrae.fr; Tel. (+33) 5 61 28 55 25.

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd.

## Introduction

The *Ralstonia solanacearum* species complex (RSSC) includes a diverse group of pathogenic strains, causative agents of bacterial wilt on many plants (Hayward, 1991; Álvarez *et al*., 2010). These strains were classified in four main evolutionary lineages also called phylotypes (Fegan and Prior, 2005) which were recently assigned to three distinct species (Safni *et al*., 2014; Prior *et al*., 2016). The bacterial wilt disease caused by RSSC strains on tomato results from a multi-stage infection process in which, after the root infection stage and colonization of the root cortex tissues, bacteria invade the xylem vessels. This allows their spread to the aerial part of the plant, concomitantly with a strong multiplication (Vasse *et al*., 1995; Caldwell *et al*., 2017; Xue *et al*., 2020). High population density in the xylem vessels, along with the production of exopolysaccharide, reduces the sap flow leading to the appearance of wilting symptoms (Hikichi *et al*., 2017). Resistance of the plant to limit bacterial colonization proceeds at different steps including root invasion (Poueymiro *et al*., 2009; Morel *et al*., 2018a), vertical movement upwards to the stem and vessel to vessel diffusion in the xylem (Caldwell *et al*., 2017; Planas-Marquès *et al*., 2020; Xue *et al*., 2020). On the other hand, *R. solanacearum* spp. relies on effectors delivered by the type 3 secretion system to circumvent plant immunity (Landry *et al*., 2020) as well as many other virulence factors which contribute to pathogenic fitness at the different infection steps (Genin and Denny, 2012; Hikichi *et al*., 2017).

Once in xylem vessels, *R. solanacearum* spp. reaches massive bacterial density up to $10^{10}$ colony-forming units per gram of fresh stem (Lowe-Power *et al*., 2018b; Perrier *et al*., 2019). Therefore, bacterial infection

success is highly dependent on the rapid assimilation of nutrients to sustain both production of biomass (proliferation) and factors subverting host immunity (virulence) (Lowe-Power *et al*., 2018b). There is evidence for intertwined regulation between virulence genes and metabolic genes (Peyraud *et al*., 2016, 2018; Khokhani *et al*., 2017), but how metabolic resources are acquired *in planta* remains poorly understood. A recent study, however, showed that sap from infected plants is enriched in several nutrients that improve *R. solanacearum* growth in sap (Lowe-Power *et al*., 2018b), notably sucrose (Hamilton *et al*., 2021).

The consumption of host resources associated with intense bacterial proliferation, as well as the effects of virulence factors, should induce strong alterations of the plant physiology. In the case of bacterial wilt, only qualitative observations of wilting are usually made, and little is known about these physiological alterations in terms of dynamic and quantitative data. The progression of the disease is generally evaluated by a disease index (disease index (DI), a visual scoring of the wilting symptoms) or by counting of the bacterial population extracted from infected plants. Although the latter method appears more sensitive for monitoring the course of the disease, it sometimes happens that the bacterial load is not correlated with disease symptom progression (Angot *et al*., 2006; Lebeau *et al*., 2011).

In this study, we aimed at quantifying the dynamics of the bacterial wilt disease on tomato plants in greenhouse conditions, using *Ralstonia pseudosolanacearum* GMI1000 (Salanoubat *et al*., 2002), a model strain representative of the RSSC. We used automated plant phenotyping and metabolomics of both xylem sap and plant organs to determine the evolution of various physiological parameters at different stages of xylem vessels colonization. This quantitative kinetic study unravelled a critical point in the *R. solanacearum* colonization process beyond which the host–pathogen interaction switches to the side of the disease, with strong physiological effects on the plant and a rapid progression of symptoms. In a 48 h window after infection, the pathogen drains the major substrates allowing for its strong proliferation. Quantitative metabolomics of infected xylem confirmed that putrescine concentration is a relevant marker of the infection process and strongly suggested that glutamine and asparagine are the major carbon sources sustaining bacterial proliferation *in planta*. On the other hand, the generation of a triple mutant strain defective for the assimilation of glucose, sucrose and fructose revealed that these sugars play little role in sustaining pathogen's growth *in planta*, as evidenced by the phenotype of this strain on tomato plants.

## Results

### *High-throughput phenotyping of tomato plants during bacterial wilt*

To evaluate the physiological and metabolic parameters associated with the infection of *R. solanacearum* strain GMI1000 of its tomato host, we monitored the progression of the bacterial wilt symptoms over 9 days in a controlled conditions greenhouse using an automated high-throughput plant phenotyping facility (http://tpmp.inrae.fr). To improve the synchronization of plant infection, the inoculation was carried out on root-damaged plants by soil drenching with a suspension of $5.10^6$ CFU per ml. Control plants (hereafter referred to as 'healthy plants') were inoculated with water only. Two independent experiments comprising, respectively, 90 and 81 plants were performed. Each experiment was analysed separately (automatic phenotyping, metabolomics) and a similar response profile was observed for both. The data presented in the following sections correspond to an experiment (3 replicates, 81 plants with 63 infected/18 healthy); the data of the second experiment are presented in Figs S4 and S6, Table S1. A larger number of infected plants was monitored given the greater variability expected in the case of the interaction with the pathogen. Automated phenotyping included weighing for the estimation of transpiration and imaging. Analysis of acquired images (Fig. S1) allowed to estimate the position (vertical) of the plant centre of gravity (Figs S2 and S3) and the chlorophyll content (Liang *et al*., 2017). Plants were also manually sampled at different days post inoculation (1, 3, 4, 5, 7, 8) to determine organ fresh weight /dry weight, as well as organ and xylem biochemical profiling and bacterial population (CFU). The appearance of bacterial wilt symptoms, from 0 (no symptoms) to 4 (complete wilting) (Arlat and Boucher, 1991) was also followed.

### R. solanacearum *alters the plant physiological balance rapidly after infection*

Physiological dynamics of the disease progression clearly revealed a tipping point in plant physiology and infection kinetic. In our experimental conditions, this point is at 4 dpi and corresponds to a colonization stage of $10^8$ bacterial colonies per gram of fresh weight (Fig. 1). Before this tipping point, plant general physiology remains similar between healthy and infected individuals: growth/transpiration rates, chlorophyll content/centre of gravity vertical position estimation (Figs 1 and 2). At 4 dpi, transpiration begins to decrease significantly for infected plants, indicating that bacterial colonization is starting to affect plant physiology in a detectable way, although symptoms are

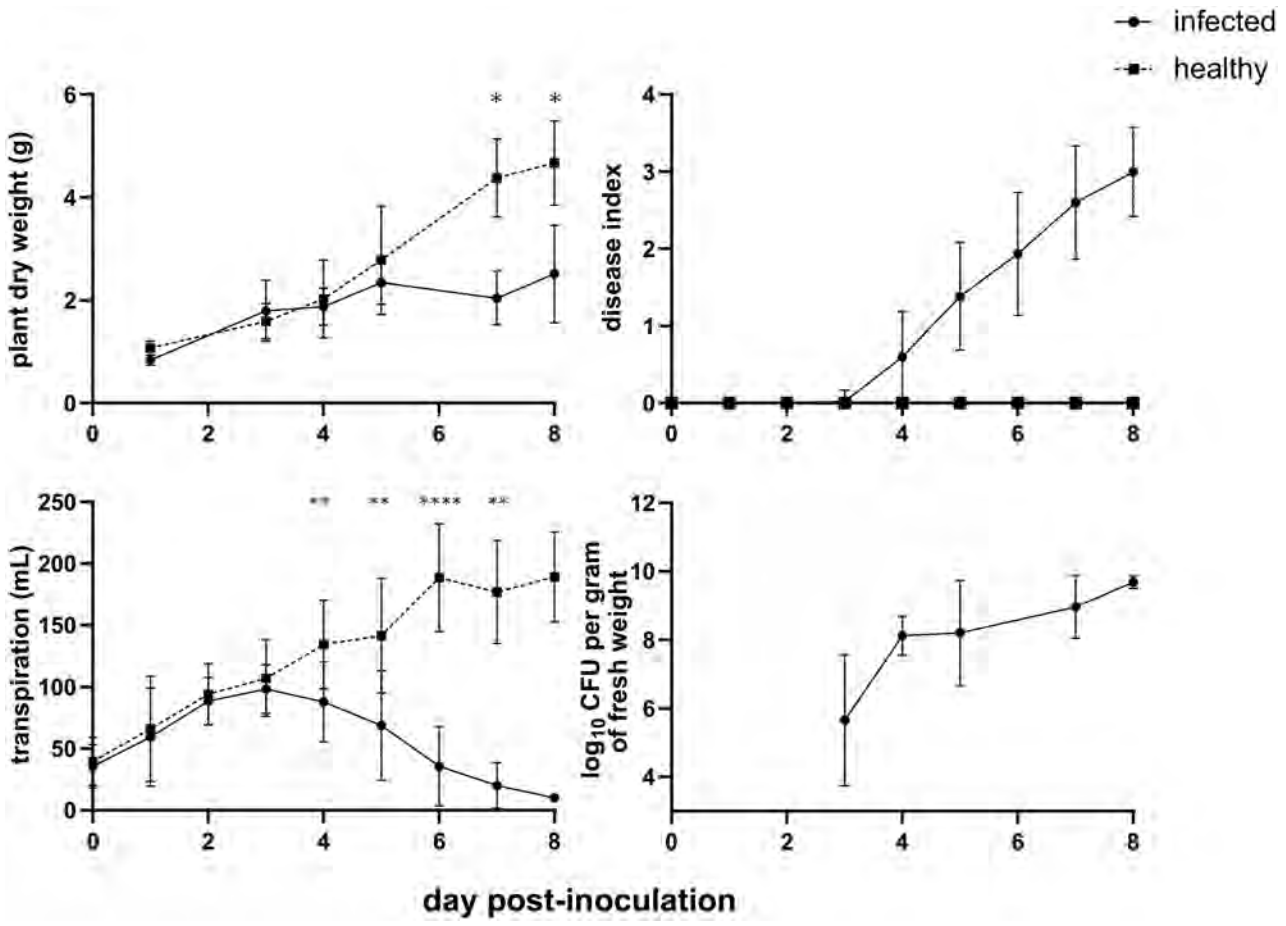



**Fig 1.** Impact of *Ralstonia solanacearum* colonization on plant physiology (dry weight and transpiration) and monitoring of disease kinetics (disease index and bacterial density). The data presented are from an experiment on 81 plants, with at least 3 plants per conditions at each sampling point. Bars indicate standard deviation. Healthy and infected plant data (transpiration and dry weight) were compared by Wilcoxon–Mann–Whitney test (*$p$ value < 0.05; **$p$ value < 0.01; ****$p$ value < 0.0001).

barely visible (50% of plants at disease index 0). Early impact on transpiration was expected as xylem occlusion limits the transmission of water to aerial parts. This very early stage of symptom development shows that before disease index score 2 the plant has already been strongly colonized by bacteria. Position of the plant centre of gravity, which is a proxy of how upwards the plant is growing, is also significantly affected (Fig. 2) at 4 dpi. Some physiology parameters such as plant growth and chlorophyll content were impacted later and correlated with disease symptom appearance (disease index of 1 and above). From 5 dpi only, plant growth began to slow down for infected plants and chlorophyll content started to decrease, further showing that a major transition occurred around 4 dpi. We will refer below to this major transition as the 'tipping point'.

After this tipping point, infected plants almost stopped growing: the growth rate was 0.044 $day^{-1}$ for infected plant versus 0.2126 $day^{-1}$ for healthy plants between 4 and 8 dpi. Transpiration also strongly decreased, from nearly 100 ml to less than 10 ml perday, while it reached around 200 ml per day for larger healthy plants. Furthermore, the fastest bacterial proliferation was observed around the tipping point (bacterial growth rate of 0.09 $h^{-1}$ between 3 and 5 dpi), before slowing down (bacterial growth rate of 0.02 $h^{-1}$ between 5 and 8 dpi). It indicated that the first colonization stage in xylem vessels is a fast proliferation of *R. solanacearum* with limited visible impacts (symptoms) but which induces a major shift in plant physiology. After this shift, bacterial proliferation slows down as the high-density population of the pathogen begins to become a burden to plant development. This is probably due to lack of nutrients because of plant stopped growth.

### *Tomato organ-specific changes induced by the bacterial wilt disease*

Fresh weight and dry weight of healthy and infected plant organs were used to determine the impact of

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

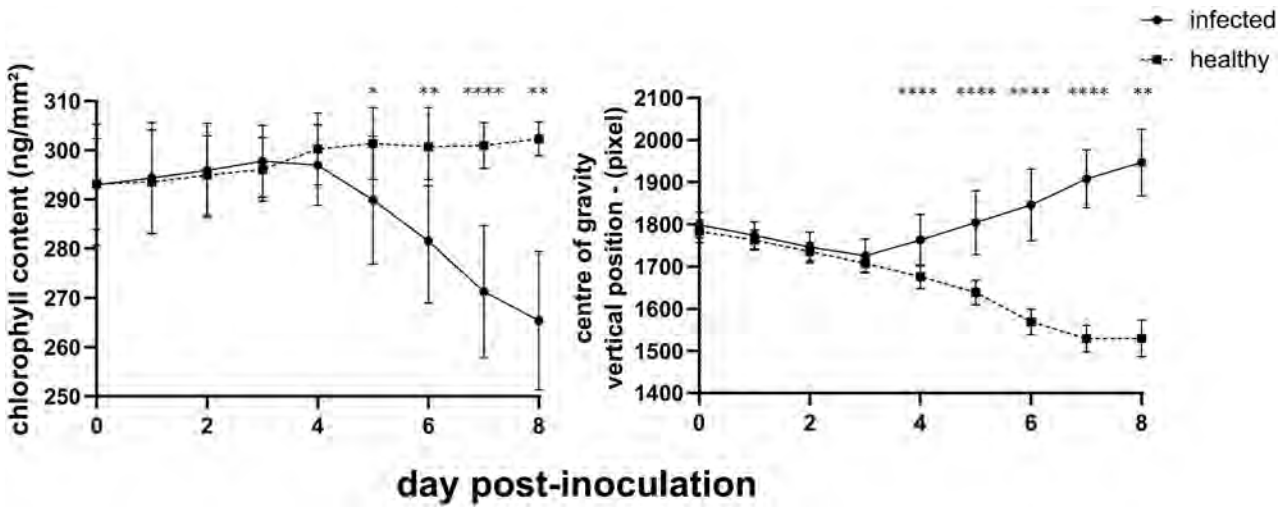



**Fig 2.** Physiological parameters determined by automatic imaging. Chlorophyll content was estimated through RGB values of plant images as described by Liang and colleagues (Liang *et al*., 2017). Centre of gravity position (vertical) was estimated through plant imaging with 0 pixel representing the top of the plant. The data presented are from an experiment on 81 plants with at least 3 plants per conditions at each sampling point. Bars indicate standard deviation. Healthy and infected plant data were compared by Wilcoxon–Mann–Whitney test (**p* value < 0.05; ***p* value < 0.01; *****p* value < 0.0001).

*R. solanacearum* proliferation on plant water status (Fig. 3A). Organ water content remains constant for healthy plants. As expected from transpiration data (Fig. 1), the percentage of water in infected plants decreased during bacterial infection for leaves and stems. Organ dehydration was observed close to the previously defined tipping point (a drop in the curve is observed at 4 dpi, although it is significant only at 5 dpi), almost simultaneously to the onset of transpiration deficiency (as observed in Fig. 1), and close to the appearance of wilting symptoms (disease index). For roots, water content was in a similar range for infected and healthy plants, confirming that major bacterial proliferation in xylem vessels does not affect root hydration (Lowe-Power *et al*., 2018b). At later infection stage (7 and 8 dpi), corresponding to a disease index score of 3 (75% of the plant wilted) with plant transpiration close to 0 ml per day, leaves and stems remained however well-hydrated, only decreasing from 87% to 80% water and 95% to 92% respectively.

Plant dehydration is suspected to affect central metabolism as high-water supply is essential for photosynthesis as well as for plant development/growth since water is necessary to constitute the turgor pressure. We thus performed quantitative metabolomics of plant tissues at 3, 5 and 7 dpi (Figs. 3B and S5), to analyse different steps of the infection process. We did not find significant differences at 3 and 5 dpi. At 7 dpi, a drastic and significant impact on leaf and stem starch was observed (Fig. 3B). In leaves, the proportion of starch was high in healthy plants (8.58% $\pm$ 6.06%) while it was depleted in infected plants (0.08% $\pm$ 0.12%). Amino acids also decreased significantly but less drastically (Fig. 3B). In accordance with the chlorophyll drop as deduced from image analysis (Fig. 2), the quantification of chlorophyll content in leaves also seems to decrease in infected tissues but not significantly (Fig. S5).

### *The bacterial wilt disease modifies metabolite availability in tomato xylem sap*

We investigated further how *R. solanacearum* is able to proliferate intensely in xylem vessels by determining the carbon sources supporting bacterial proliferation in xylem sap. We extracted xylem sap of infected and healthy plants at different dpi to perform quantitative analysis by (nuclear magnetic resonance (NMR), see Experimental Procedures). To assess which xylem metabolites are affected by infection, we performed a principal component analysis (PCA). This multivariate analysis revealed that two principal components (PCs) explained 55% of the variability (see Fig. S7), on which we focused our analysis. Each individual point plotted in Fig. 4A corresponds to a xylem sap sample (15 metabolites concentrations) transformed into the two uncorrelated PCs and associated with the corresponding bacterial population. We found that the first PC (explaining 30% of the variability) is not correlated with infection stage and reveals an important variability between plants independent of the infection. However, the second PC (explaining 25% of the variability) is strongly associated with the bacterial population level. Interestingly, uninfected xylem samples have very similar metabolic profiles as all the points are

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

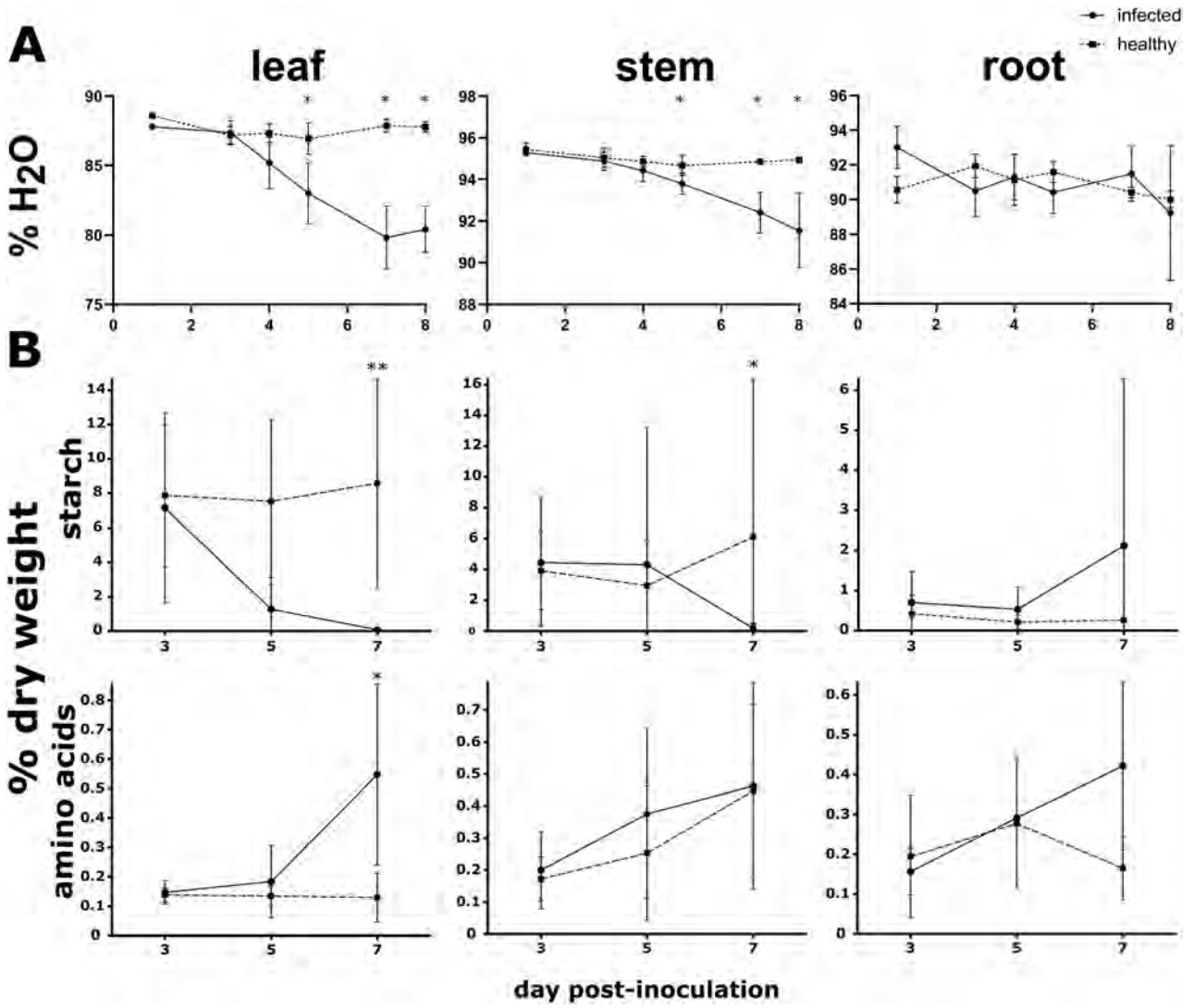



**Fig 3.** Impact of *Ralstonia solanacearum* infection on plant tissues.
A. Tissue hydration. Fresh weight and dry weight were used to determine the proportion of water in healthy and infected tissues. Bars indicate standard deviation.
B. Metabolic profile of tissues expressed in percentage of dry weight for starch and amino acids. The mean and standard deviation displayed were obtained after combining the acquired data of the two biological replicates (total of 12 samples for leaf, 10 for stem, 12 for root at each time point).
In A and B, healthy and infected plant data were compared by Wilcoxon-Mann–Whitney test (*$p$ value < 0.05; **$p$ value < 0.01).

clustered whereas infected samples display a scattered profile. This suggests that impact of bacterial infection on xylem sap composition is heterogeneous between plants and that there is therefore a certain level of variation at this stage.

We plotted the contribution of the variables (metabolite concentrations) to the two PCs in Fig. 4B. Variables differing most from zero on the horizontal axis (e.g. ethanol, lysine, phenylalanine and valine) have an important contribution to the first PC, which implies that for these compounds there is an important inter-plant variability not correlated to infection. Conversely, variables differing most from zero on the vertical axis contribute to the second PC and are therefore associated with the bacterial population level. High levels of asparagine and glutamine (bottom of the plot) are associated to healthy plants and plants with low bacterial population, while high levels of putrescine and acetate are associated with high bacterial population (see complete list in Table 1). We found that for healthy plants, concentrations are constant between 1 and 8 dpi (no statistical differences observed between sampling days, and no apparent temporal dynamics) so we determined a global mean for each plant (Tables 1 and S1). Glutamine appeared as the

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

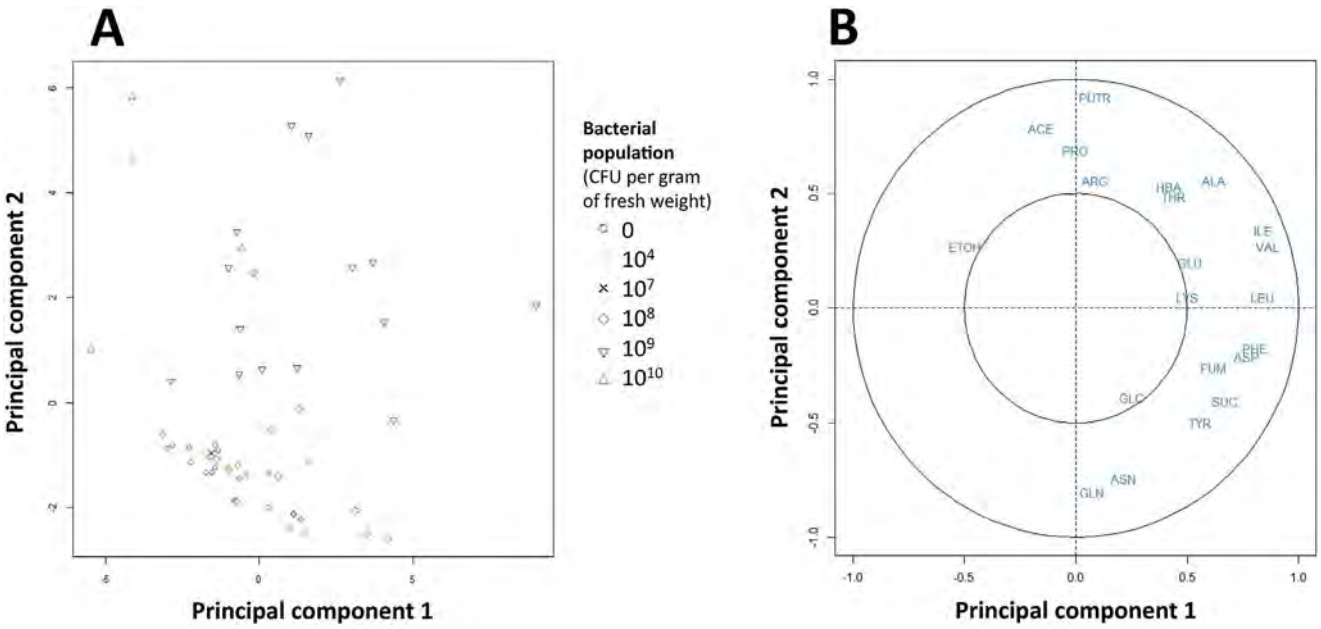



**Fig 4.** Principal component analysis on metabolite concentrations from samples at different dpi.
Representation of individual points (A) and the variables (B). Zero bacterial population corresponds to healthy plant samples. Data points at the bottom of the plot correspond to uninfected or low bacterial population samples while data points at the top correspond to high bacterial population samples. Principal component 1: 30% of variability. Principal component 2: 25% of variability. ETOH, ethanol; PUTR, putrescine; FUM, fumarate; SUC, sucrose; GLC, glucose. [Color figure can be viewed at wileyonlinelibrary.com]

**Table 1.** Xylem metabolite content and the impact of *Ralstonia solanacearum* infection.

| Metabolite | Concentration in healthy plant (mM) | Effect of the infection[a] | Growth of *R. solanacearum* as sole carbon source[b] |
|---|---|---|---|
| Glutamine | 3.29 ± 1.25 | Decrease | + |
| Asparagine | 0.178 ± 0.074 | Decrease | + |
| Tyrosine | 0.022 ± 0.010 | Putative decrease | NA |
| Glucose | 0.010 ± 0.026 | Putative decrease | + |
| Sucrose | 0.028 ± 0.018 | Putative decrease | + |
| Fumarate | 0.004 ± 0.006 | Putative decrease | + |
| Aspartate | 0.015 ± 0.014 | No effect | + |
| Phenylalanine | 0.035 ± 0.022 | No effect | − |
| Leucine | 0.088 ± 0.029 | No effect | − |
| Lysine | 0.123 ± 0.062 | No effect | − |
| Glutamic acid | 0.020 ± 0.041 | No effect | + |
| Threonine | 0.062 ± 0.019 | Putative increase | + |
| Valine | 0.083 ± 0.025 | Putative increase | + |
| Proline | 0.066 ± 0.033 | Putative increase | − |
| Alanine | 0.012 ± 0.007 | Putative increase | + |
| Isoleucine | 0.048 ± 0.015 | Putative increase | − |
| Ethanol | 0.203 ± 0.314 | Putative increase | NA |
| Arginine | 0.094 ± 0.051 | Putative increase | − |
| 3-Hydroxybutyric acid | 0.005 ± 0.040 | Putative increase | + |
| Acetate | 0.008 ± 0.033 | Increase | + |
| Putrescine | Not detected | Appearance | − |

[a]Effect of infection was deduced from the PCA presented in Fig. 4. Metabolite concentration decrease was assigned for compounds with coordinates on the second component inferior to −0.75 whereas increase in concentration was assigned for those superior to 0.75. Values ranging between −0.25 and − 0.75 correspond to compounds with putative decrease in concentration, those between 0.25 and 0.75 to putative increase. Values ranging from −0.25 and 0.25 defined compounds with no significant variation in concentration (ie not impacted by infection). Putrescine was considered as 'appearing' as it was absent in healthy plants and had a coordinate on the second component superior to 0.75 (see Table S2).
[b]Usage as carbon source (+: growth, −: no growth) was assessed from published Biolog Phenotype Microarrays data (Peyraud *et al*., 2016).

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

major component of xylem sap (75% of molar organic carbon) whereas the other compounds were found in much lower concentrations (16 times less at least).

We found that the progression of the infection was associated with glutamine and asparagine decrease and provoked the appearance of putrescine (non-detectable in healthy plants) and an increase of acetate content. We also estimated that aspartate, phenylalanine, lysine, glutamic acid and leucine were not affected by *R. solanacearum* infection as the coordinates on the second component are close to zero (inferior to 0.25). It was difficult to conclude firmly for the other variables as they were in an intermediate (and so less discriminating) area, between 0.25 and 0.75.

### *Dynamic profiling of tomato xylem fluid metabolites during* R. solanacearum *infection*

PCA showed that some tomato xylem metabolites concentration vary during infection by *R. solanacearum*. In order to get a dynamic view of these changes, we monitored metabolite concentrations over several days in healthy and infected plants. The kinetic data for the seven compounds with statistical differences for one time point or more are represented in Fig. 5 (and in Fig. S8 for alanine). We noticed that statistical differences were mostly observed at 3, 4 or 5 dpi, certainly due to a higher number of plants analysed and sometimes high variability. For additional compounds tested, no statistical difference was observed (Fig. S8).

Putrescine was detected when plant infection resulted in high bacterial proliferation (between 3 and 5 dpi, see Fig. 1) and its concentration stabilized in the last days of infection (7 and 8 dpi) when bacterial multiplication started to slow down. Putrescine therefore appears to be a sensitive and precise biomarker of the infection: its dynamics is strongly correlated with the onset and evolution of the bacterial wilt disease. Three other metabolites increase importantly and significantly in infected plants: 3-hydroxybutyric acid increases at 4 and 5 dpi, while it seems to decrease afterward, acetate increases from 4 to 7 dpi, and alanine (Fig. S8) seems to follow the same pattern than 3-hydroxybutyric acid.

Glutamine depletion is striking as the initial concentration (almost 3 mM) decreases to around 1 mM at the end of the sampling. Simultaneously, asparagine, which is observed at a smaller amount in healthy plants (around 0.15 mM), is also quickly depleted: our data indicate a drastic drop between 4 and 5 dpi. After this drop, asparagine concentration is around 0 (or at concentrations under the limit of detection). These parameters prove an intense metabolic activity over a short time window (48 h) in xylem vessels, exactly around the tipping point of the infection previously reported (4 dpi).

### *Putrescine production is a reliable marker of* R. solanacearum *infection*

The progression of bacterial wilt disease is usually monitored in two ways: either by disease index scoring or by enumerating bacteria extracted from infected tissues. We plotted the relationship between these two phenotypic markers and some of the physiological parameters we monitored during the infection assay: plant transpiration, putrescine production and glutamine concentration in xylem (Fig. 6). The calculation of the regression coefficients ($R^2$) shows that these three markers are globally well correlated with the bacterial density (in $\log_{10}$ CFU per gram of fresh weight) and the disease index, although we observe for the latter a greater dispersion of the points in the early stages of infection. As expected, bacterial population best correlated with wilting symptoms development ($R^2 = 0.81$ with disease index). Plant transpiration is an early marker of infection impact on plant physiology (Fig. 1) and accordingly a clear negative relationship was also observed with the bacterial population rate, confirming that bacterial proliferation has a strong negative impact on plant transpiration rate. It turned out that the bacterial population rate in xylem is also well correlated to the metabolic profile of infected xylem: positive correlation between putrescine concentration (in decimal logarithm) is high ($R^2 = 0.66$), indicating that appearance and accumulation of putrescine is directly associated with *R. solanacearum* infection. Negative correlation with glutamine concentration is also significant but weaker ($R^2 = 0.54$).

### *Genetic evidence that sugars are not major carbon sources supporting growth of* R. solanacearum *in xylem*

Xylem metabolomics revealed a significant decrease of glutamine over time as well as a negative correlation between bacterial density and glutamine content, whereas sugar content was found to not vary. To validate the hypothesis that *R. solanacearum* growth is mostly supported by glutamine and not sugars, we generated a triple mutant strain unable to uptake (and so metabolize) three major sugars found in plants: sucrose, glucose and fructose, glucose and sucrose being detected in our xylem content analysis. The strain GMI1000 glucose transport operon (RSp1632-RSp1635) was identified after screening a mutant library to identify genes essential for growth on glucose (Peyraud *et al*., 2018), but these genes were incorrectly annotated as a xylose transport system. Cumulative deletions in each sugar transport system were generated to create the mutant strain GRS941. We used *in vitro* growth assays on microplates to validate that strain GRS941 growth on sucrose (Fig. 7A), glucose (Fig. 7B) and fructose (Fig. S9) was severely impacted. We then performed automated

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

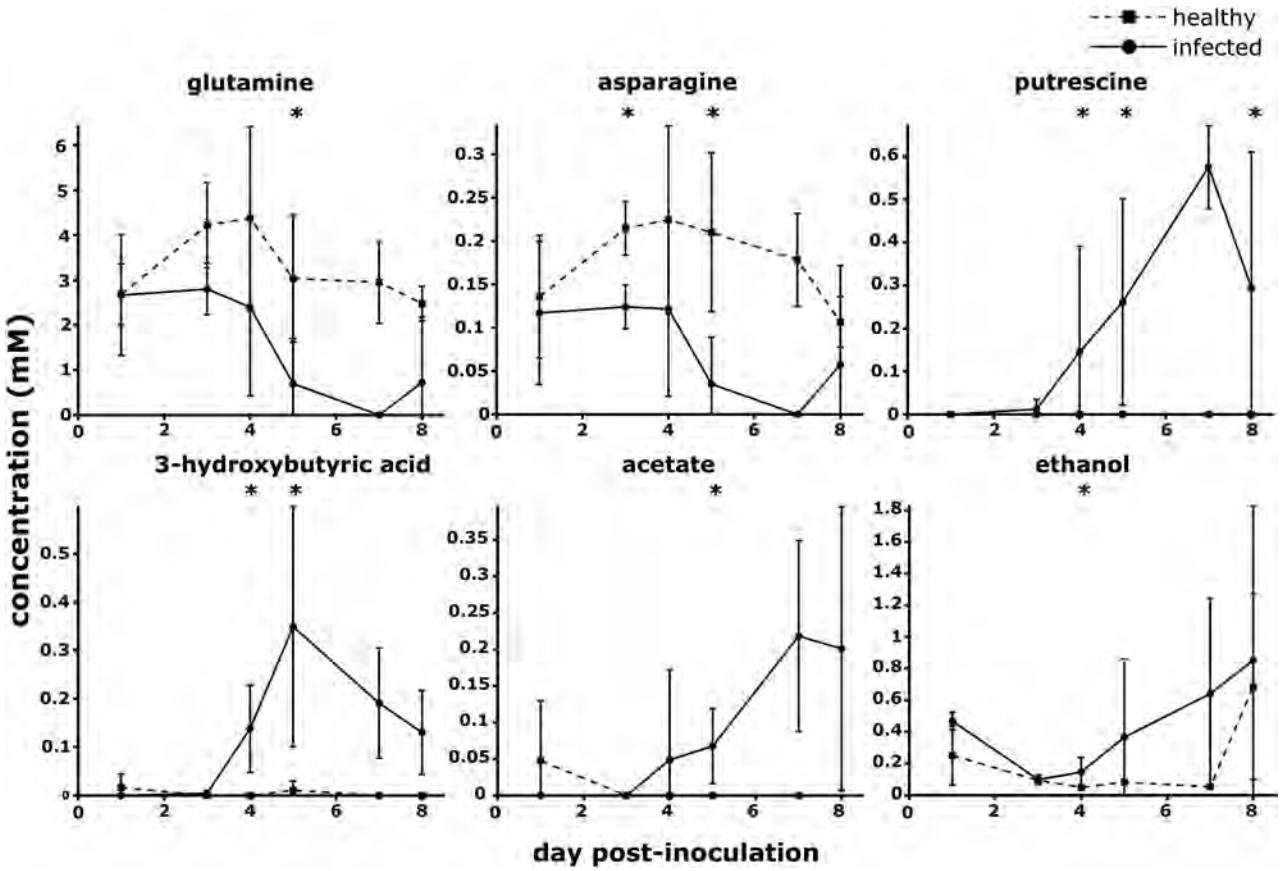



**Fig 5.** Kinetics of metabolic contents in xylem sap of infected or healthy plants based on a total of 57 xylem sap samples obtained at different dpi. Healthy and infected plant data were compared by Wilcoxon–Mann–Whitney test (*$p$ value < 0.05; **$p$ value < 0.01).

phenotyping of strain GRS941 on tomato plants under conditions similar to those used for the wild-type strain. We observed that growth *in planta* of strain GRS941 was similar to the wild-type and reached high population levels (Fig. 7C). The transpiration rate was significantly higher for plants infected by GRS941 compared with the wild-type at 6 to 8 dpi, but the difference was small with a similar overall profile and an identical tipping point at 4 dpi (Fig. 7D). Small but statistically significant variations of plant dry weight, disease index and imaging profiles were also observed (Figs 7 and S9), but the mutant strain was still able to provoke full plant wilting symptoms. We therefore concluded that glucose, fructose and sucrose are not critical carbon sources sustaining the pathogen growth during the xylem infection stage, but rather complementary nutritive sources.

## Discussion

In this study, we followed the dynamics of infection of the pathogen *Ralstonia solanacearum*at the time of xylem colonization and its impact on plant physiology. Unlike several studies that have focused on the *R. solanacearum*–plant interaction using miniaturized *in vitro* pathosystems, we have conducted this work on 4 weeks old tomato plants grown in greenhouse conditions to get closer to *in natura* interactions. In order to synchronize the infection process, we carried out infections with scarified roots, enabling easier entry of bacteria into vascular tissues. This root wounding procedure made it possible to follow the infection kinetics with relatively homogeneous values of the measured parameters (see Figs 1–3 and 5), contrary to what is observed with soil soaking inoculations where greater variability is observed (Morel *et al.*, 2018b). We used our experimental system to follow not only symptoms but also general growth and water status of the plant, as well as measure metabolites without a priori. We were then able for the first time to correlate them in a time series for a finer understanding of the overall process.

### *Physiological markers define a tipping point of bacterial wilt infection*

The use of automated phenotyping has enabled us to collect many measurement points, thus allowing for a

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

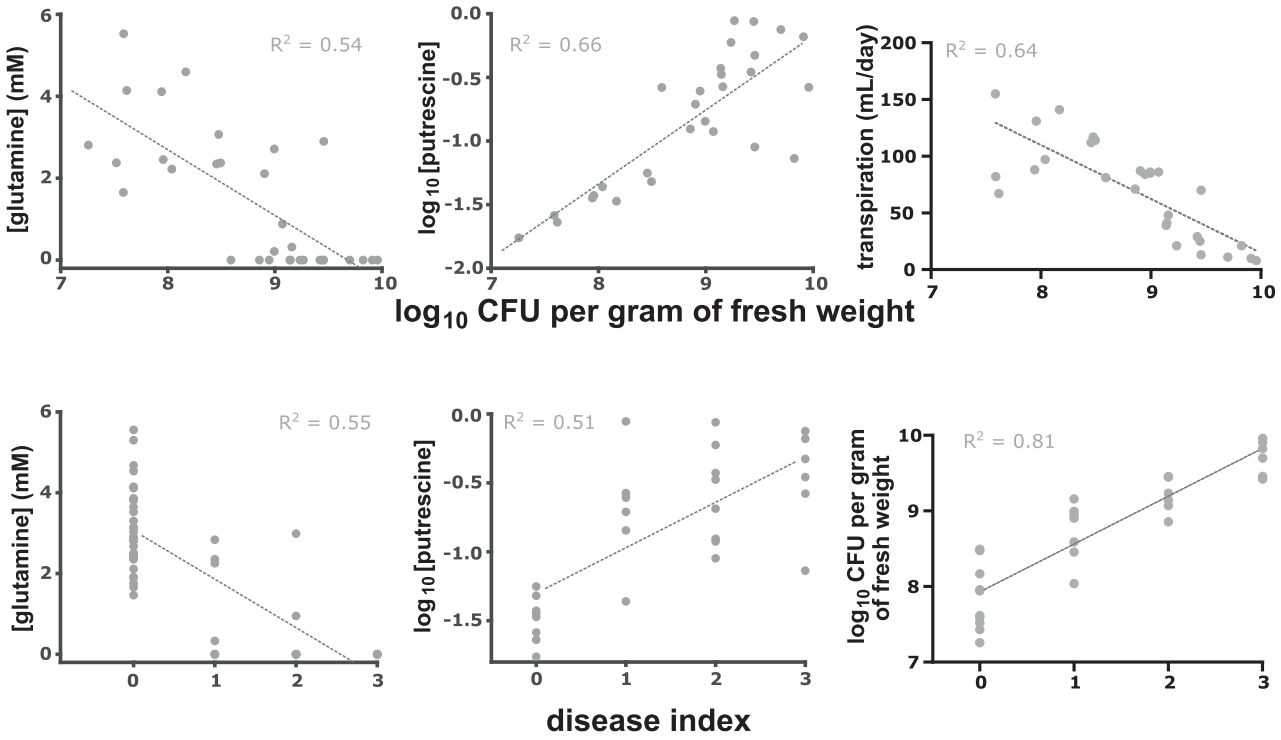



**Fig 6.** Correlations between markers of the infection. For CFU per gram of fresh weight, only data for which the counting was above $10^7$ CFU per gram of fresh weight (corresponding to the 'tipping point') were included.

sensitive estimation of major parameters linked to the infection of tomato by *R. solanacearum*. In particular, our study reveals that markers such as the plant transpiration rate in plant tissue are excellent proxies to follow the progression of bacterial infection. Moreover, the comparison of the measurements in these standardized conditions show that the quantification of the bacterial load *in planta*, despite being destructive, is much more reliable than the rating of the symptoms by disease index which is still commonly used to follow the evolution of an infection (see Fig. 6). The absence of visible symptoms is observed with plants already highly infected (up to $10^9$ CFU per gram of fresh weight) and disease index scores of 0, 1 and 2 include very diverse *in planta* bacterial loads. It is only when the disease index reaches 3 that the profiles become more homogeneous, but they correspond to an already very advanced stage of bacterial propagation and alteration of plant physiology. This emphasizes that disease index is a late marker of the infection process and some physiological markers such as the plant transpiration rate or vertical position of the centre of gravity (proxy of how upwards the plant is growing) have a much better correlation with bacterial colonization.

The fine tracking of parameters (weight, transpiration, centre of gravity, chlorophyll content) of the infection highlighted a specific step that we consider as a tipping point beyond which the disease progresses to plant death. This tipping point corresponds to a stage where macroscopic disease symptoms are barely visible (disease index = 0.5) and to the first detection of bacteria after extraction from plant tissues for counting (CFU per gram of fresh weight). This reveals that the plant can afford a load of bacteria in its vascular system before seeing its immunity overwhelmed by the pathogen, which is consistent with the observation of bacterial proliferation reaching $10^6$ CFU per gram of fresh weight in the resistant tomato plant cultivar Hawaii 7996 (Nakaho *et al*., 2004). We assume that the dynamics around a tipping point described here is a conserved trait of bacterial wilt; however, its temporal positioning could vary depending on the plant species and plant age (which could influence immune response or root infection).

The plant water loss observed during *R. solanacearum* infection (Fig. 3) is in a similar range to what is observed during drought on tomato (Sánchez-Rodríguez *et al*., 2010), around 10% of water content, showing that bacterial wilt imposes significant drought stress on the plant. Photosynthesis requires a big amount of water, so water deficit should quickly impair photosynthesis and, therefore, have an impact on the ability to store carbon such as starch during the daytime phase of the day. In

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

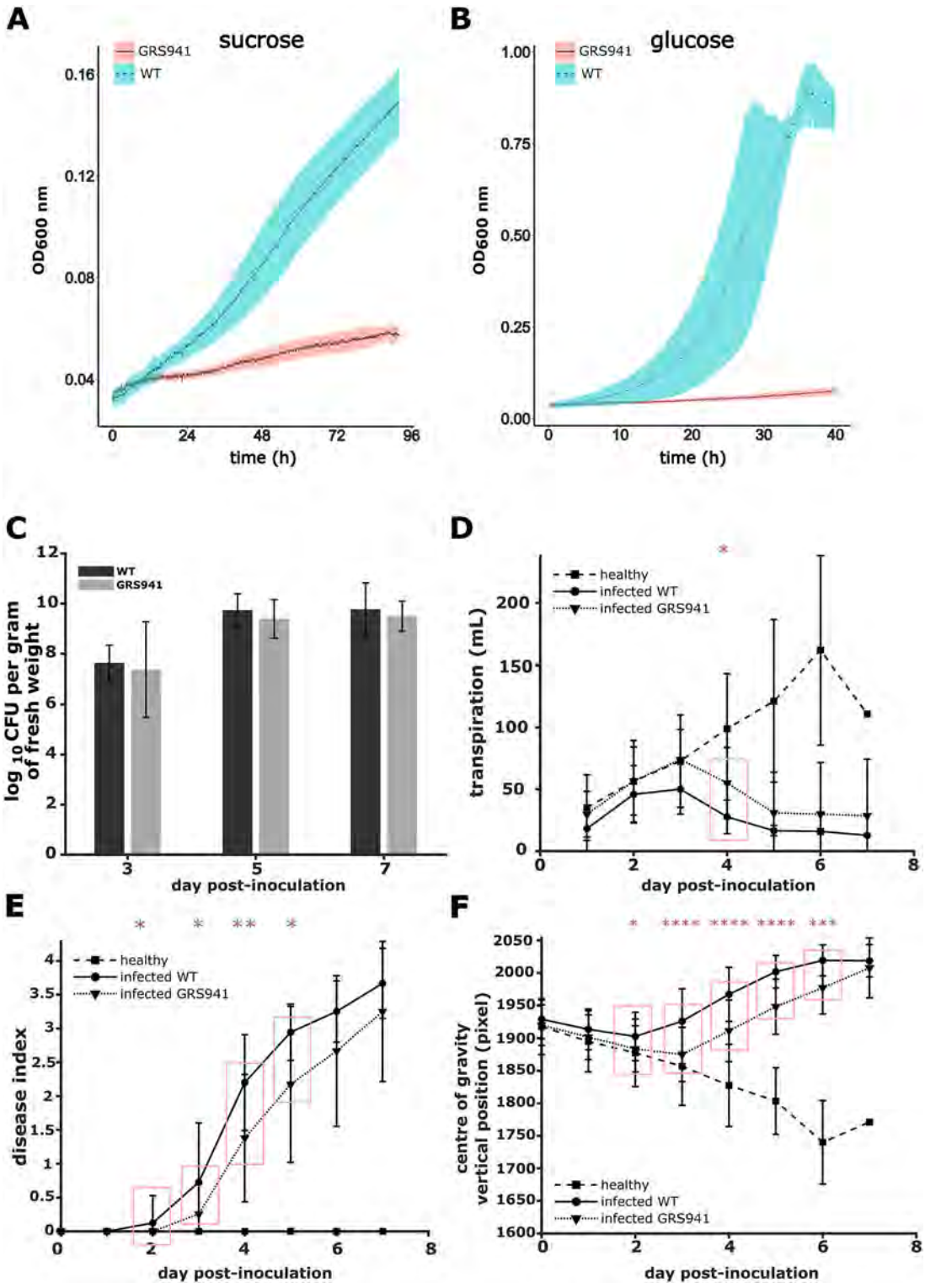



**Fig 7.** *In vitro* and *in planta* behaviour of the *Ralstonia solanacearum* mutant strain GRS941 (carrying deletion in the sucrose, glucose and fructose transporter genes), compared with the wild-type strain GMI1000.
Growth on minimal medium supplemented with sucrose (A) and glucose (B) as sole carbon source monitored in microplates with three independent biological replicates. Means and standard deviations are represented by the lines and coloured areas respectively. Bacterial density *in planta* (C) was determined on at least nine plants per condition and per day, and plant transpiration (D) on at least seven plants per condition and per day for infected WT/GRS941 (except 3 at 8 dpi) and at least three plants per day for healthy plants. Means are presented and bars indicate standard deviation. The wild-type and mutant strains were compared by Wilcoxon–Mann–Whitney test (*$p$ value < 0.05; **$p$ value < 0.01; ****$p$ value < 0.0001). [Color figure can be viewed at wileyonlinelibrary.com]

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

agreement with this, we observed in our metabolomic analysis a clear drop in starch storage in infected aerial parts at 7 dpi (Fig. 3). The estimation of chlorophyll content by imaging revealed a progressive decrease after infection, which remained limited at 7 dpi (Fig. 3). The incidence of bacterial wilt is probably less apparent on the chlorophyll content than on the accumulation of starch since chlorophyll, even though described as decreasing under stress, is subjected to a turnover slower than starch (Scialdone *et al*., 2013).

### *Quantitative insights into* R. solanacearum *preferred carbon sources in xylem*

In a second part of this work, we studied the kinetic variations of metabolites of xylem sap associated with the bacterial wilt disease. The fast proliferation of *Ralstonia solanacearum* in xylem must be sustained by nutrients that the pathogen uptakes in this environment. We therefore studied specifically the metabolites that were less abundant in the sap of infected plants compared with that of healthy plants, presumably used by bacteria for efficient growth. The PCA multivariate analysis of metabolomic data revealed that glutamine and asparagine are significantly depleted in infected plants. This is particularly striking for glutamine for which the initial concentration ($2.66 \pm 1.34$ mM at 1 dpi) decreases to lower than 1 mM at the last days of sampling. Glutamine is described as the major organic component of xylem sap in plants of different botanical families (Andersen and Brodbeck, 1989; Zuluaga *et al*., 2013; Montes Borrego *et al*., 2017). Our data are in line with these reports: glutamine in tomato xylem sap was measured $3.29 \pm 1.25$ mM for non-infected plants, which represents 75% of the molar organic carbon in this compartment. We showed that glutamine concentration was negatively correlated with bacterial growth, supporting the view that glutamine is most probably the main carbon source sustaining fast proliferation in vascular tissues. As glutamine is by far the most abundant amino acid in xylem, we hypothesize that *R. solanacearum* has adapted to preferentially metabolize this compound. The five different transport systems for this compound present in the genome of GMI1000 highlight the importance of this metabolite.

Simultaneously to glutamine, the asparagine pool (around $0.178 \pm 0.074$ mM in sap of healthy plants) is quickly depleted after 4 dpi, falling under the detection limit (between 1 and 10 μM). According to our data, sugars do not appear to play a major role as sources of carbon and energy for the bacteria in xylem vessels. Indeed, sugars are present in scarce proportions compared to amino acids such as glutamine or asparagine. Glucose can be detected but in concentration at the limit of reliable detection, and the sucrose content also appears rather low. Thus, the sugars concentration does not seem sufficient to explain the strong bacterial growth. Bacterial infection should lead to the enrichment of the sap with sugars such as sucrose and glucose, as reported by (Lowe-Power *et al*., 2018a). However, based on the kinetic and quantitative values we hypothesized that these metabolites might have a role in signalling rather than being primary carbon sources to sustain rapid growth, unless sucrose/glucose dynamics is too fast to be captured by a day by day kinetics. This was confirmed by the behaviour of a triple mutant strain that is unable to grow *in vitro* when glucose, sucrose and fructose are provided as sole carbon sources. This strain, however, is still able to colonize the plant, reaching high density in the stem, and causes physiological alteration in plants (suppression of growth and transpiration) and symptoms (disease index) delayed but similar to the wild-type strain. We thus concluded that sugars do not play a major role in the pathogen's growth at this stage.

Our metabolomics study also revealed that several potential nutrients (such as aspartate, phenylalanine and leucine) remain in xylem sap even at later infection stages in concentrations similar to that of healthy plants. This suggests that these metabolites are not consumed by *R. solanacearum*, probably due to the lack or non-expression of the appropriate transporter genes. For various other lower abundant nutrients, the high variability observed in infected plants prevents making clear conclusions. In an earlier study, Zuluaga and colleagues (2013) analysed the xylem sap content of tomato and it should be noted that some carbon sources reported in their study were not detected in ours. This mainly concerns amino acids such as gamma-aminobutyric acid, histidine and serine. The presence of fructose was also detected while our data reveal the presence of sucrose but not fructose. Higher quantities of glucose, fructose and sucrose were found in the tomato apoplast (Zuluaga *et al*., 2013). However, as the GRS491 triple mutant strain was still able to cause effective bacterial wilt, this suggests that metabolization of these sugars is also not essential for the growth of *R. solanacearum* in the apoplastic environment.

A recent study revealed glutamic acid as a major amino acid promoting expression of some *R. solanacearum* pathogenic traits (Shen *et al*., 2020). Inversely to glutamine, glutamic acid is present in trace amounts in our xylem sap analysis (Table 1) and its concentration remains constant during infection. The NMR analysis of xylem sap unambiguously discriminated between the two amino acids (Fig. S10). Therefore, we favour a role of glutamic acid in signalling to induce production of virulence factors rather than as a major carbon source metabolized in xylem vessels.

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

*Metabolic signatures of bacterial wilt*

Beside the depletion of some metabolites as mentioned above, we also observed the increase of other compounds in xylem following *R. solanacearum* infection. The emergence of such compounds could be from bacterial origin (excreted product such as putrescine; Lowe-Power *et al*., 2018a; Peyraud *et al*., 2016), or part of a plant metabolic response to infection (Zeiss *et al*., 2019). In some cases, these compounds could come simultaneously from the bacteria and from the plant. In the case of putrescine, the role of the molecule may be dual since bacterial production has been shown to contribute to the wilt symptoms (Lowe-Power *et al*., 2018a,b) and also be produced by tomato cells hijacked by a *Ralstonia solanacearum* Transcription activator-like (TAL) effector (Wu *et al*., 2019).

We found that upon infection xylem sap accumulated 3-hydroxybutyric acid, acetate, ethanol and potentially proline (discriminated by PCA, Fig. 4 although no significant difference in the kinetics on Fig. S8). Accumulation of proline and ethanol has been associated with drought stress (Sánchez-Rodríguez *et al*., 2010; Kelsey *et al*., 2014) and one can assume that the bacterial wilt disease induces a sudden and brutal water deficit for the infected plant (Zeiss *et al*., 2019). Ethanol production could also be a physiological response to xylem occlusion by *R. solanacearum*: the formation of bacterial biofilm would limit the diffusion of oxygen in the xylem vessels, thus inducing a switch to anaerobic respiration in plant cells and therefore the production of ethanol as a fermentation product. Concerning proline, its release into the xylem by plant cells may be due to the need to compensate for the depletion of glutamine/asparagine consumed by the bacteria. Acetate was also found to appear in infected xylem; similarly to ethanol, acetate could be a plant fermentation product due to a limitation of oxygen diffusion, but it could also result from an overflow of the bacterial metabolism (El-Mansi and Holms, 1989). As previously reported by Lowe-Power and colleagues (2018a), we observed an increased synthesis of 3-hydroxybutryic acid. 3-hydroxybutyric acid is a precursor of polyhydroxybutyrate, a storage polymer identified in *R. solanacearum* and phylogenetically close β-proteobacteria. Kinetics showed an increase followed by a decrease phase of 3-hydroxybutyric acid concentration (Fig. 4C), in agreement with the hypothesis that this metabolite could be excreted by *Ralstonia* and then consumed before returning to the soil (Lowe-Power *et al*., 2018a). However, it cannot be excluded that bacteria induce production of 3-hydroxybutyric acid by the plant, as the biosynthetic pathway is present in tomato plant (Yuan *et al*., 2016).

In a complementary approach, Lowe-Power and colleagues (2018a) have recently performed Gaz Chromatrography–Mass Spectrometry (GC-MS) metabolomics of tomato xylem sap infected by *R. solanacearum*. This allowed a non-targeted approach but not an absolute quantification of the metabolites as we were able to achieve it by the NMR approach. Based on the relative fold-change of metabolites in the xylem sap of healthy versus infected plants, several carbon sources such as galactose, 3-hydroxybutyrate, gluconate and glucose were identified as important carbon sources for *R. solanacearum* in the xylem (Lowe-Power *et al*., 2018a). In our study, except for 3-hydroxybutyric acid and glucose (at a concentration below 0.01 mM), the other compounds were not detected, indicating that these compounds are in concentration below 1–10 μM, which is the sensibility limit of NMR. Significant change folds can be observed more easily for low-concentration metabolites than for abundant ones. This thus suggests that the metabolites, even if they had high fold change, are probably not primary carbon sources supporting bacterial growth. Given the quantities of carbon source available in tomato xylem sap and their observed consumption by *R. solanacearum* (Fig. 5), our current model is that there is co-consumption of glutamine and asparagine to support the pathogen high proliferation in xylem. In this scenario, sugar carbon sources play a secondary role, acting as a complementary source of nutrition. Nevertheless, they may be important for the rapid expression of disease symptoms. It cannot be excluded that the importance of sugars may vary depending on the host infected. It is also possible that the perception of plant sugars, similarly to glutamic acid (Shen *et al*., 2020), may act as a signal on bacterial gene expression during infection. Comparative studies using xylems from other *R. solanacearum* host plants, as well as the use of labelled molecules, should help address these points.

## Experimental procedures

*Plant cultures and automatic phenotyping*

Two independent experiments were conducted for this study. In both cases, tomato seeds (*Solanum lycopersicum* M82) were grown in soil (SB2, Proveen, The Netherlands) supplemented with Osmocote® coated fertilizer at a rate of 4 g per L. Seed were germinated in a growth chamber (26°C, 67% HR, 12 h LED light per day). Around a hundred of plantlets were transplanted in individual plastic pots (8 × 8 cm) 8 days after sowing. Sixteen days after sowing, 90 young plants were chosen and repotted in 3 L pots until the end of the experiment. Foam cover discs were placed on each pot to limit physical evaporation.

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

On the first experiment, 90 plants were loaded on the Phenoserre robot facility of the Toulouse Plant Microbe Phenotyping infrastructure, of which 45 plants were inoculated with *R. pseudosolanacearum* GMI1000 strain (Salanoubat *et al*., 2002) and another 45 mock inoculated with sterile water (referred to as 'Healthy plants'). Twelve hours of light per day at 28°C and 50% humidity were programmed. On the second experiment, 81 plants were loaded on the Phenoserre robot, of which 63 infected and 18 healthy plants.

All the plants were watered with 100 ml three times on the loading day and weighted, in order to define a well-watered target weight. The subsequent watering was conducted automatically once a day to the target weight. Imaging was programmed four times a day (starts at 2.00 a.m., 9.00 a.m., 4.30 p.m and 7.30 p.m, for a total duration of around 1 h and a half per imaging job), the conveyor belt bringing the individual plants to a side-RGB camera. The plants were then automatically weighted after each imaging session. The daily transpiration was determined as the weight differences between two consecutive days, at the time of watering. Temperature, hygrometry and light intensity were recorded during the whole experiment.

*Plant image analysis*

Figure S1 presents the step-by-step segmentation performed using a set of python algorithms allowing for the testing of the segmentation pipeline. Every plant is imaged using an RGB camera with a blue background, under six angles (stating at 0° and then by 60° steps). Three different 'Region of Interest' (ROIs) are defined to allow easier analysis, in white containing all the plant information, with a 'safe zone' in green and a 'danger zone' in red. As the images are acquired in the exact same conditions, the ROIs are static. Masks are extracted corresponding to the bulk of the plant ('mask top'), as well as two additional masks to gather the plant information close to the pot and conveyor ('mask middle' and 'mask bottom'). All three masks are merged in order to build a coarse mask, this latter mask is then cleaned by removing additional noise to generate a clean mask and allowing for the final features extraction.

Fig. S2 displays the 'centroid y' (or Cy) morphological traits, which is the *y* coordinate of the plant area centre of gravity. The origin of the y axe is positioned at the top of the image. A larger Cy value indicates a centre of gravity closer to the pot (further from the top of the image) and is a trait that is very well associated with the bacterial wilt symptoms, as any 'lowering' of the leaves (wilting) has an impact on the Cy trait. Fig. S3 displays some Cy values of plants with different DI.

*Plant inoculation*

For both experiments, three replicates of the strain GMI1000 were grown in BG complete medium (Plener *et al*., 2010) at 28°C. Bacterial cells were washed, resuspended in sterile water and each plant was inoculated with one of the three replicates by pouring 50 ml of a suspension at $5{\cdot}10^6$ CFU per ml at the stem base 4 days after the loading on the robot. Just before inoculation, roots were scarified with a scalpel blade in order to synchronize the infection rate.

*Collection and preparation of plant samples*

For the first experiment, plants were removed each day from the conveyor belt for samplings during 9 consecutive days. On the inoculation day and at 7 and 8 dpi, three healthy and three infected plants were removed. From 2 to 6 dpi, six healthy and six infected plants were removed each day. For the second experiment, plants were removed from the conveyor belts for samplings at 1, 3, 4, 5, 7 and 8 dpi. Each sampling day, 3 healthy plants and, respectively, 3, 6, 14, 16, 14 and 10 infected plant were removed.

In both experiments, the following procedures were then applied at the same period of the day (between 10 a.m. and 2 p.m.) to avoid the effect of daily variations in xylem sap/organ chemistry. For the plant that were removed from the experimental setup, stems were cut just above the cotyledon node, rinsed with approximately 1 ml of water and the upcoming xylem sap was collected by repeated pipetting and collection in Eppendorf tubes placed on ice. For the infected plants, the tubes were centrifuged and the supernatants were transferred in new tubes while the pellets were discarded. The tubes with xylem sap were placed at −80°C for further quantitative NMR analyses.

The different organs of each plant were collected separately (stems, leaves, roots and flowers whenever they occurred) for fresh weight and dry weight measurements. Approximately, 300 mg of fresh weight for each collected organ was frozen in liquid nitrogen and stored at −20°C for further biochemical analyses.

One centimetre of stem was sampled above the cotyledons, cut up in small pieces and placed in a 2 ml centrifuge tube containing glass beads (2 mm diameter) and ground at 30 Hz with a mixer mill (MM 400, Retsch, Germany) until total grinding (1–3 min). Extracted bacteria were resuspended in 1 ml of sterile water, serial-diluted and plated on complete BG medium supplemented with triphenyl tetrazolium chloride, for counting as previously described (Guidot *et al*., 2014).

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

*Quantitative NMR analyses*

The xylem saps were analysed by 1D $^{1}$H NMR on MetaToul analytics platform (UMR5504, UMR792, CNRS, INRAE, INSA, Toulouse Cedex, France), using the Bruker Avance 800 MHz equipped with an ATMA 5 mm cryoprobe. Each xylem sap sample was centrifuged to remove the residues (5 min, 13,520 RCF, Hettich Mikro 200 centrifuge), then placed in 3 mm NMR tubes. TSP-d4 standard [sodium 3-(trimethylsilyl)(1-13C,2H4)propanoate] was used as a reference. pH 6.0 phosphate buffer was used to standardize the chemical shifts among samples. Acquisition conditions were as follows: 30° pulse angle, 20.0287 ppm spectral width, 64 scans per acquisition for a total scan time of approximately 8 min per sample, and zgpr30 water pre-saturation sequence. The samples were kept at a temperature of 280 K (6.85°C) all along the analysis. Resonances of metabolites were manually integrated and the concentrations were calculated based on the number of equivalent protons for each integrated signal and on the TSP final concentration.

*Biochemical analyses of metabolites*

Quantifications of metabolites were performed at the HitMe platform (INRAE – IBVM, Villenave d'Ornon Cedex, France). The plants samples, previously frozen in liquid nitrogen, were ground to a powder using liquid nitrogen to avoid thawing. A quantity of 20 $\pm$ 10 mg of each were weighted in previously frozen micronic tubes. Free amino acids, glucose, fructose, malate, proteins, starch, sucrose and chlorophylls in leaves, stems and roots were quantified as described in Biais *et al*. (2014). Briefly, ethanolic extracts from every samples were obtained using three consecutive incubation of the frozen ground powder aliquots. Ethanol 80% v/v with HEPES/KOH 10 mM pH 6 buffer was used for the two first incubations, and ethanol 50% v/v with HEPES/KOH 10 mM pH 6 buffer was used for the third. Supernatants were pooled and used for the quantification of chlorophylls, glucose, fructose, sucrose, malate and free amino acids. Pellets were used for the determination of protein and starch contents. The extracts and pellets were stored at −20°C between each quantification. For each sample, chlorophylls were quantified by measuring optical densities at 645 and 665 nm on a mix of 50 µl of extract supplemented with 150 µl of analytics grade ethanol. Amino acids were quantified using the fluorescamine method. Excitation wavelength was 405 nm and emission was measured at 485 nm. The proteins were quantified using Bradford reagent. Starch was quantified in glucose equivalent after full pellet digestion in an oven at 37°C for 18 h. For the other analytes cited above, the NADH/NADPH appearance was measured, and the analytes were quantified using a 1:1 stoichiometric coefficient.

*Generation and phenotyping of* R. solanacearum *triple mutant strain defective for glucose, sucrose and fructose assimilation*

Strain GRS941 was engineered by cumulating deletions in the glucose transport operon (RSp1632-RSp1635), the sucrose assimilation operon (RSp1280-RSp1286) and the fructose transport operon (RSc2861-RSc2863). Complete procedures for the triple mutant generation are available in File S1.

For *in vitro* growth assay, preculture on BG complete medium (Plener *et al*., 2010), cell centrifugation/washing and resuspension were performed. Strains were then inoculated at an OD (600 nm) of 0.01 on minimal medium (Plener *et al*., 2010) supplemented with, respectively, glucose, sucrose, fructose, glutamine as sole carbon source (at 50 mM of carbon). Growth was assessed using microplate spectrophotometer (FLUOstar Omega, BMG Labtech, Germany) during 96 h at 28°C and under shaking at 700 rpm. For *in planta* phenotyping, the same procedure than described for the GMI1000 strain was followed on two independent experiments, with, respectively, 31 and 41 tomato plants infected by the mutant strain and similar number of plants infected by the wild-type strain.

**Acknowledgements**

Léo Gerlin was funded by a PhD grant from the French Ministry of National Education and Research. Antoine Escourrou was funded by the French Laboratory of Excellence TULIP (ANR-10-LABX-41 and ANR-11-IDEX-0002-02). The study was funded by the French Laboratory of Excellence (LABEX) project TULIP (ANR-10-LABX-41 and ANR-11-IDEX-0002-02). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank Toulouse Plant Microbe Phenotyping (TPMP) platform (Castanet-Tolosan, France) and its staff Fabrice Devoilles for their technical support in plant cultures and imaging, as well as High-Throughput Metabolic Phenotyping (HiTMe) platform (Villenave d'Ornon, France) for biochemical analysis of organ metabolites. We acknowledge MetaToul (Metabolomics and Fluxomics Facitilies, Toulouse, France, www.metatoul.fr) platform, which is part of the MetaboHUB-ANR-11-INBS-0010 national infrastructure (www.metabohub.fr) and its staff Cécilia Berges, Edern Cahoreau and Lindsay Peyriga for access to NMR facilities. We also thank Patrick Barberis for his expert skills in the construction of the triple mutant strain.

**Conflict of interest**

The authors declare no conflict of interest.

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., 

## Author contributions

Conceptualization: Stéphane Genin, Caroline Baroukh. Methodology: Léo Gerlin, Antoine Escourrou, Nemo Peeters, Stéphane Genin, Caroline Baroukh. Formal analysis: Léo Gerlin, Antoine Escourrou, Felicià Maviane Macia, Caroline Baroukh. Investigation: Léo Gerlin, Antoine Escourrou, Cédric Cassan, Caroline Baroukh. Resources: Nemo Peeters. Writing – original draft: Léo Gerlin, Antoine Escourrou. Writing – review and editing: Léo Gerlin, Antoine Escourrou, Nemo Peeters, Stéphane Genin, Caroline Baroukh. Visualization: Léo Gerlin, Antoine Escourrou. Project administration: Caroline Baroukh. Funding acquisition: Caroline Baroukh.

### *Data availability statement*

All data supporting the findings of this study are available within the paper and within its supplementary materials published online.

## References


Álvarez, B., Biosca, E.G., and López, M.M. (2010) On the life of Ralstonia solanacearum, a destructive bacterial plant pathogen. *Curr Res Technol Educat Top Appl Microbiol Microbial Biotechnol* **1**: 267–279. https://pdfs.semanticscholar.org/aa85/77e213e2977a0e4eb739795e3fea51187181.pdf?_ga=2.156465595.1055873898.1504096691-1311319012.1504096691.

Andersen, P.C., and Brodbeck, B.V. (1989) Diurnal and temporal changes in the chemical profile of xylem exudate from *Vitis rotundifolia*. *Physiol Plant* **75**: 63–70. https://doi.org/10.1111/j.1399-3054.1989.tb02064.x.

Angot, A., Peeters, N., Lechner, E., Vailleau, F., Baud, C., Gentzbittel, L., *et al*. (2006) *Ralstonia solanacearum* requires F-box-like domain-containing type III effectors to promote disease on several host plants. *Proc Natl Acad Sci USA* **103**: 14620–14625. https://doi.org/10.1073/pnas.0509393103.

Arlat, M., and Boucher, C. (1991) Research notes identification of a dsp DNA region controlling aggressiveness of *Pseudomonas solanacearum*. *Mol Plant Microbe Interact* **4**: 211. https://doi.org/10.1094/mpmi-4-211.

Biais, B., Bénard, C., Beauvoit, B., Colombié, S., Prodhomme, D., Ménard, G., *et al*. (2014) Remarkable reproducibility of enzyme activity profiles in tomato fruits grown under contrasting environments provides a roadmap for studies of fruitmetabolism. *Plant Physiol* **164**: 1204–1221. https://doi.org/10.1104/pp.113.231241.

Caldwell, D., Kim, B.-S., and Iyer-Pascuzzi, A.S. (2017) *Ralstonia solanacearum* differentially colonizes roots of resistant and susceptible tomato plants. *Phytopathology* **107**: 528–536. https://doi.org/10.1094/PHYTO-09-16-0353-R.

El-Mansi, E.M.T., and Holms, W.H. (1989) Control of carbon flux to acetate excretion during growth of *Escherichia coli* in batch and continuous cultures. *J Gen Microbiol* **135**: 2875–2883. https://doi.org/10.1099/00221287-135-11-2875.

Fegan, M., and Prior, P. (2005) How complex is the "Ralstonia solanacearum species complex"? In *Bacterial Wilt Disease and the Ralstonia Solanacearum Species Complex*, July, pp. 449–461. St. Paul, MN: APS Press.

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

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. https://doi.org/10.1093/molbev/msu229.

Hamilton, C.D., Steidl, O.R., MacIntyre, A.M., Hendrich, C. G., and Allen, C. (2021) Ralstonia solanacearum depends on catabolism of myo-inositol, sucrose, and trehalose for virulence in an infection stage-dependent manner. *Mol Plant Microbe Interact*: MPMI-10-20-0298-R. https://apsjournals.apsnet.org/doi/10.1094/MPMI-10-20-0298-R.

Hayward, A.C. (1991) Biology and epidemiology of bacterial wilt caused by *Pseudomonas solanacearum*. *Annu Rev Phytopathol* **29**: 65–87. https://doi.org/10.1146/annurev.py.29.090191.000433.

Hikichi, Y., Mori, Y., Ishikawa, S., Hayashi, K., Ohnishi, K., Kiba, A., and Kai, K. (2017) Regulation involved in colonization of intercellular spaces of host plants in *Ralstonia solanacearum*. *Front Plant Sci* **8**: 967. https://doi.org/10.3389/fpls.2017.00967.

Kelsey, R.G., Gallego, D., Sánchez-García, F.J., and Pajares, J.A. (2014) Ethanol accumulation during severe drought may signal tree vulnerability to detection and attack by bark beetles. *Can J For Res* **44**: 554–561. https://doi.org/10.1139/cjfr-2013-0428.

Khokhani, D., Lowe-Power, T.M., Tran, T.M., and Allen, C. (2017) A single regulator mediates strategic switching between attachment/spread and growth/virulence in the plant pathogen *Ralstonia solanacearum*. *MBio* **8**: e00895–e00817. https://doi.org/10.1128/mBio.00895-17.

Landry, D., González-Fuente, M., Deslandes, L., and Peeters, N. (2020) The large, diverse, and robust arsenal of Ralstonia solanacearum type III effectors and their *in planta* functions. *Mol Plant Pathol* **21**: 1377–1388. https://doi.org/10.1111/mpp.12977.

Lebeau, A., Daunay, M.C., Frary, A., Palloix, A., Wang, J.F., Dintinger, J., *et al*. (2011) Bacterial wilt resistance in tomato, pepper, and eggplant: genetic resources respond to diverse strains in the *Ralstonia solanacearum* species complex. *Phytopathology* **101**: 154–165. https://doi.org/10.1094/PHYTO-02-10-0048.

Liang, Y., Urano, D., Liao, K.L., Hedrick, T.L., Gao, Y., and Jones, A.M. (2017) A nondestructive method to estimate the chlorophyll content of *Arabidopsis* seedlings. *Plant Methods* **13**: 26. https://doi.org/10.1186/s13007-017-0174-6.

Lowe-Power, T.M., Hendrich, C.G., von Roepenack-Lahaye, E., Li, B., Wu, D., Mitra, R., *et al*. (2018a) Metabolomics of tomato xylem sap during bacterial wilt reveals *Ralstonia solanacearum* produces abundant putrescine, a metabolite that accelerates wilt disease.


© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

*Environ Microbiol* **20**: 1330–1349. https://doi.org/10.1111/1462-2920.14020.

Lowe-Power, T.M., Khokhani, D., and Allen, C. (2018b) How *Ralstonia solanacearum* exploits and thrives in the flowing plant xylem environment. *Trends Microbiol* **26**: 929–942. https://doi.org/10.1016/j.tim.2018.06.002.

Montes Borrego, M., Jiménez-Díaz R.M., Trapero Casas J. L., Navas Cortés J.A.; Haro C., Rivas J.C., de la Fuente, L., and Landa, B.B. (2017). Metabolomic characterization of xylem sap of different olive cultivars growing in Spain. In *European Conference on Xylella*, November.

Morel, A., Guinard, J., Lonjon, F., Sujeeun, L., Barberis, P., Genin, S., *et al*. (2018a) The eggplant AG91-25 recognizes the type III-secreted effector RipAX2 to trigger resistance to bacterial wilt (*Ralstonia solanacearum* species complex). *Mol Plant Pathol* **19**: 2459–2472. https://doi.org/10.1111/mpp.12724.

Morel, A., Peeters, N., Vailleau, F., Barberis, P., Jiang, G., Berthomé, R., and Guidot, A. (2018b) Plant pathogenicity phenotyping of Ralstonia solanacearum strains. In *Methods in Molecular Biology*, New York, NY: Humana Press Inc, Vol. **1734**, pp. 223–239. https://doi.org/10.1007/978-1-4939-7604-1_18.

Nakaho, K., Inoue, H., Takayama, T., and Miyagawa, H. (2004) Distribution and multiplication of *Ralstonia solanacearum* in tomato plants with resistance derived from different origins. *J General Plant Pathol* **70**: 115–119. https://doi.org/10.1007/s10327-003-0097-0.

Perrier, A., Barlet, X., Rengel, D., Prior, P., Poussier, S., Genin, S., and Guidot, A. (2019) Spontaneous mutations in a regulatory gene induce phenotypic heterogeneity and adaptation of *Ralstonia solanacearum* to changing environments. *Environ Microbiol* **21**: 3140–3152. https://doi.org/10.1111/1462-2920.14717.

Peyraud, R., Cottret, L., Marmiesse, L., and Genin, S. (2018) Control of primary metabolism by a virulence regulatory network promotes robustness in a plant pathogen. *Nat Commun* **9**: 418. https://doi.org/10.1038/s41467-017-02660-4.

Peyraud, R., Cottret, L., Marmiesse, L., Gouzy, J., and Genin, S. (2016) A resource allocation trade-off between virulence and proliferation drives metabolic versatility in the plant pathogen *Ralstonia solanacearum*. *PLoS Pathog* **12**: e1005939. https://doi.org/10.1371/journal.ppat.1005939.

Planas-Marquès, M., Kressin, J.P., Kashyap, A., Panthee, D. R., Louws, F.J., Coll, N.S., and Valls, M. (2020) Four bottlenecks restrict colonization and invasion by the pathogen *Ralstonia solanacearum* in resistant tomato. *J Exp Bot* **71**: 2157–2171. https://doi.org/10.1093/jxb/erz562.

Plener, L., Manfredi, P., Valls, M., and Genin, S. (2010) PrhG, a transcriptional regulator responding to growth conditions, is involved in the control of the type III secretion system regulon in *Ralstonia solanacearum*. *J Bacteriol* **192**: 1011–1019. https://doi.org/10.1128/JB.01189-09.

Poueymiro, M., Cunnac, S., Barberis, P., Deslandes, L., Peeters, N., Cazale-Noel, A.C., *et al*. (2009) Two type III secretion system effectors from *Ralstonia solanacearum* GMI1000 determine host-range specificity on tobacco. *Mol Plant Microbe Interact* **22**: 538–550. https://doi.org/10.1094/MPMI-22-5-0538.

Prior, P., Ailloud, F., Dalsing, B.L., Remenant, B., Sanchez, B., and Allen, C. (2016) Genomic and proteomic evidence supporting the division of the plant pathogen *Ralstonia solanacearum* into three species. *BMC Genomics* **17**: 1–11. https://doi.org/10.1186/s12864-016-2413-z.

Safni, I., Cleenwerck, I., De Vos, P., Fegan, M., Sly, L., and Kappler, U. (2014) Polyphasic taxonomic revision of the *Ralstonia solanacearum* species complex: proposal to emend the descriptions of *Ralstonia solanacearum* and *Ralstonia syzygii* and reclassify current *R. syzygii* strains as *Ralstonia syzygii* subsp. *syzygii* subsp. nov., R. S. *Int J Syst Evol Microbiol* **64**: 3087–3103. https://doi.org/10.1099/ijs.0.066712-0.

Salanoubat, M., Genin, S., Artiguenave, F., Gouzy, J., Mangenot, S., Arlat, M., *et al*. (2002) Genome sequence of the plant pathogen *Ralstonia solanacearum*. *Nature* **415**: 497–502. https://doi.org/10.1038/415497a.

Sánchez-Rodríguez, E., Rubio-Wilhelmi, M., Cervilla, L.M., Blasco, B., Rios, J.J., Rosales, M.A., *et al*. (2010) Genotypic differences in some physiological parameters symptomatic for oxidative stress under moderate drought in tomato plants. *Plant Sci* **178**: 30–40. https://doi.org/10.1016/j.plantsci.2009.10.001.

Scialdone, A., Mugford, S.T., Feike, D., Skeffngton, A., Borrill, P., Graf, A., *et al*. (2013) *Arabidopsis* plants perform arithmetic division to prevent starvation at night. *Elife* **2**: e00669. https://doi.org/10.7554/eLife.00669.

Shen, F., Yin, W., Song, S., Zhang, Z., Ye, P., Zhang, Y., *et al*. (2020) *Ralstonia solanacearum* promotes pathogenicity by utilizing L-glutamic acid from host plants. *Mol Plant Pathol* **21**: 1099–1110. https://doi.org/10.1111/mpp.12963.

Vasse, J., Frey, P., and Trigalet, A. (1995) Microscopic studies of intercellular infection and protoxylem invasion of tomato roots by Pseudomonas solanacearum. *Mol Plant Microbe Interact* **8**: 241–251. https://doi.org/10.1094/MPMI-8-0241.

Wu, D., von Roepenack-Lahaye, E., Buntru, M., de Lange, O., Schandry, N., Pérez-Quintero, A.L., *et al*. (2019) A plant pathogen type III effector protein subverts translational regulation to boost host polyamine levels. *Cell Host Microbe* **26**: 638–649.e5. https://doi.org/10.1016/j.chom.2019.09.014.

Xue, H., Lozano-Durán, R., and Macho, A.P. (2020) Insights into the root invasion by the plant pathogenic bacterium *Ralstonia solanacearum*. *Plan Theory* **9**: 516. https://doi.org/10.3390/plants9040516.

Yuan, H., Cheung, C.Y.M., Poolman, M.G., Hilbers, P.A.J., and van Riel, N.A.W. (2016) A genome-scale metabolic network reconstruction of tomato (*Solanum lycopersicum* L.) and its application to photorespiratory metabolism. *Plant J Cell Mol Biol* **85**: 289–304 https://doi.org/10.1111/tpj.13075.

Zeiss, D.R., Mhlongo, M.I., Tugizimana, F., Steenkamp, P. A., and Dubery, I.A. (2019) Metabolomic profiling of the host response of tomato (*Solanum lycopersicum*) following infection by *Ralstonia solanacearum*. *Int J Mol Sci* **20**: 3945 https://doi.org/10.3390/IJMS20163945.

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

Zuluaga, A.P., Puigvert, M., and Valls, M. (2013) Novel plant inputs influencing Ralstonia solanacearum during infection. *Front Microbiol* **4**: 349. https://doi.org/10.3389/fmicb.2013.00349.

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**File. S1.** Generation of *R. solanacearum* triple mutant strain defective for glucose, sucrose and fructose assimilation.

**File. S2.** Script for the segmentation pipeline.

**Fig. S1.** Step-by-step segmentation pipeline.

**Fig. S2.** The Centroid-y trait.

**Fig. S3.** Examples of Cy values on healthy and diseased plants.

**Fig. S4.** Biological replicate of the physiological measurements.

**Fig. S5.** Impact of *R. solanacearum* infection on plant tissues metabolic content.

**Fig. S6.** PCA on the biological replicate data: experimental points (A) and variables (B).

**Fig. S7.** Explained variance of the principal components.

**Fig. S8.** Concentrations profiles for metabolites with non-statistically significant infection effect.

**Fig. S9.** *In vitro* and *in planta* phenotyping of *R. solanacearum* triple mutant strain defective for glucose, sucrose and fructose assimilation.

**Fig. S10.** NMR spectrum example and spectrum differences between glutamine and glutamic acid (or glutamate).

**Table. S1.** Biological replicate of the composition of non-infected xylem.

**Table. S2.** Loading weights on the second component of the PCA.

© 2021 Society for Applied Microbiology and John Wiley & Sons Ltd., *Environmental Microbiology*

## 4.3 Kinetics of xylem sap metabolites: the challenges of *in planta* metabolomics

### 4.3.1 Measuring the bacterial exponential growth at the right timing

The dynamics of *R. solanacearum* in xylem vessels are technically challenging to estimate. During the first few days, no difference was observed between healthy and infected sap, but once *R. solanacearum* reaches xylem vessels, in a very short time window, the plant will be so overwhelmed by bacterial virulence and proliferation that almost no xylem sap could be extracted (as bacteria decrease water flow by clogging the vascular system).

Hence, in the article, an important variability is observed in the last day of xylem sampling (Figure 5 of the article, 8 dpi), due to the facts that: i) some of the remaining plants have reached a very late stage of the disease, with almost no transpiration and consequently no possibilities to extract xylem sap and analyze it ii) the very few plants for which xylem sap extraction is possible are potentially those with delayed colonization (due to the stochasticity of the early invasion step) and thus, less clear depletion/increase of nutrient is observable.

To overpass the limitation of kinetics plots, linking bacterial density (instead of time) with the metabolomics data can be an alternative, as done in Figure 4 (on the whole set of data, using Principal Component Analysis) and Figure 6 (using two-dimension plots for the metabolites of interest). However, this analysis also has its limitations, as counting bacteria extracted from plant tissues also has biases (the entire population cannot be extracted in absolute terms). It is only by combining both of the analyses that we could decipher more clearly the metabolites that appeared to be important in the interaction.

### 4.3.2 Xylem sap sensitivity to environment, nutrition and cultivar

As mentioned in Chapter 3, organic composition of xylem sap is very dependent on nutritional or environmental conditions such as the choice of nitrogen nutrition. The xylem concentrations in the two experiments (Table 1 and Table S1 of the article) are rather close, but my team performed similar xylem extraction and analysis following the same protocol on a different tomato cultivar (cv. Marmande versus cv. M82 in the

article) without a fertilizer and in classical growth chamber. This results in differences in concentrations (4 times less glutamine in their experiment) and the detection of some metabolites only in one of the two conditions (e.g. absence of fumarate in their experiment). The relationships between environmental/nutritional conditions or cultivars and xylem sap should be further investigated as it could be a key parameter for the success of *R. solanacearum* proliferation *in planta*. However, as mentioned in the article, it is worth noting that some traits seems to be conserved even between plants of different botanical families, such as the predominance of glutamine or the presence of other amino acids at more diluted concentrations.

## 4.4 Metabolism and physiology of *Ralstonia solanacearum* in plants

### 4.4.1 Bacterial density *in planta*

Using our counting of bacteria in stem, we deduced the bacterial load (density) *in planta* using tomato plant density (estimated by weighing and volume measurement) and xylem/plant volume ratio (estimated using microscope images of xylem vessels in another project of my team, data not published) (

**Table** 4-1). We can then link bacteria enumerated *in planta* to bacterial density (cells·mL [xylem sap]$^{-1}$):

$$\text{Bacterial density (cells} \cdot \text{mL [xylem sap]}^{-1}) = \text{ cells} \cdot \text{mL [plant]}^{-1} \cdot \frac{\text{vol. plant}}{\text{vol. xyl}}$$

$$\text{Bacterial density (cells} \cdot \text{mL [xylem sap]}^{-1}) = \text{ cells} \cdot \text{mL [plant]}^{-1} \cdot \frac{100}{P_{vol}}$$

$$\text{Bacterial density (cells} \cdot \text{mL [xylem sap]}^{-1}) = \text{ cells} \cdot \text{g [fresh weight]}^{-1} \cdot d_{plant} \cdot \frac{100}{P_{vol}}$$

$$\text{Bacterial density (cells} \cdot \text{mL [xylem sap]}^{-1}) = \text{ cells} \cdot \text{g [fresh weight]}^{-1} \cdot 75.26$$

**Table 4-1: Parameters used to determine bacterial density *in planta***

| Parameter | Value |
|---|---|
| $P_{vol}$: Xylem/plant ratio (% volume) | 1.063 |
| $d_{plant}$: Tomato plant density at 4 dpi (g[fresh weight]·mL$^{-1}$) | 0.8 |

With this approach, I estimated that the bacterial density at the tipping point (physiological switch to the disease, at 4 dpi) is around $1.8 \cdot 10^{10}$ cells·mL [xylem sap]$^{-1}$. The estimation of the bacterial density in the xylem vessels at the tipping point reveals that the plant can afford a large load of bacteria in its vascular system before seeing its immunity overwhelmed by the pathogen. This threshold of $1.8 \cdot 10^{10}$ cells·mL [xylem sap]$^{-1}$ (in our conditions) is already beyond (or 360 times higher than) the cell density-sensing signal (quorum sensing) threshold activating the production of EPS and other virulence factors ($5.10^7$ cell·mL$^{-1}$ as defined by (Clough et al., 1997)). This indicates that a fully virulent pathogen population can still be controlled by the plant immune system before the tipping point is reached.

In addition, using this approach, we can better conceptualize the amount of bacteria found *in planta*: the concentration at the tipping point corresponds theoretically to an OD of 17.7 (using cells·mL$^{-1}$/OD correlation from Peyraud et al. (2016)) and the maximal amount of bacteria obtained ($9.10^9$ cell per g of fresh weight) would correspond to an OD of 673. Classical *in vitro* growth usually leads to around OD of 1 or 2 (depending on carbon substrate concentration) and the cell density activating the quorum sensing is at less than 0.1 OD. This illustrates that *in vitro* cultures, even at high density, are always very far to the population density observed *in planta*.

### 4.4.2 Putrescine in plant – *Ralstonia* interaction

The best metabolic marker of *R. solanacearum* proliferation is clearly putrescine, as previously identified in previous *in planta* and *in vitro* metabolomics essays (Lowe-Power, Hendrich, et al., 2018; Peyraud et al., 2016). We estimated a yield of putrescine *in planta* of 5.9 µmol putrescine·g [bacterial dry weight]$^{-1}$ using the formula of bacterial density previously described and dry weight/cell density correlations from Peyraud et al. (2016). *In vitro*, the yield is 78 µmol putrescine·g [bacterial dry weight]$^{-1}$ according to data by Peyraud et al. (2016). Interestingly, the yield *in vitro* is thirteen times superior to our yield *in planta*. In addition to possible imprecisions in the measurement *in planta,* three explanations could be drawn: i) plant partially consumes bacterial putrescine ii) putrescine level is diluted by a high xylem flow iii) bacterial physiology is very different *in planta* as a result of differences in oxygenation and population density for example, which leading to an excretion flux very far from the one observed *in vitro*.

This comparatively low putrescine yield *in planta* must be taken into account for discussing the hypothesis of putrescine production by plant, suggested by Wu et al. (2019). Even though it could still be possible that the plant produces a small amount of putrescine, quantitatively, bacterial production is enough to explain the amount detected *in planta* as the metabolic yield is under the metabolic yield *in vitro*.

Metabolism of putrescine, and other polyamines, is very complex to untangle in plant – pathogen interactions, as polyamines are classical products of bacterial metabolism and are essential for eukaryotic organisms. They are thus are produced by the plant, in particular in response to a stress. In addition, studies revealed that some pathogens could manipulate plant polyamine biosynthesis. I contributed to a review summarizing the known roles of polyamines in plant – pathogen interaction, available in **Appendix IV**.

### 4.4.3 *In planta* proliferation of R. solanacearum

Contrary to the growth of *R. solanacearum* in shaken-flask (batch) culture where the bacterium have access to a static source of nutrients, bacteria stuck to xylem vessels are subjected to a continuous flow of xylem sap. Hence, the xylem environment could be conceptualized as a chemostat bioreactor (continuous culture), where the microorganism is subjected to a constant input/output flow, allowing it to access to a continuous supply of nutrients. Thus, it will stay at elevated growth rate without being limited by a lack of nutrient. Alternatively, as bacteria are stuck to xylem vessels, and use EPS to block water flow to aerial parts, we could rather see xylem vessel as a fed-batch bioreactor, where there is an input flow of nutrient, but no output flow diluting the bacterial density. This system can reach high cell density compared to static (batch) culture, and helps understanding the population density observed *in planta*. In particular, using the transpiration rate presented in the article, it has been possible to determine the nutrient flow in xylem (study by Baroukh et al., accepted). This revealed that the flow should be rich-enough (in particular with the amount of glutamine quantified) to sustain the bacterial proliferation observed *in planta*, and refutes that xylem is nutrient-poor for bacterial proliferation.

The appearance of acetate in our results, if produced by *R. solanacearum*, could also be an indicator of high growth rate *in planta*. Above a growth rate threshold, bacteria usually reach a maximal rate of oxygen consumption and are not able to fully catabolize substrates using only oxidative phosphorylation, leading to a limitation in substrate

consumption and to the excretion of by-products such as acetate. My team performed several experiments on shaken-flask on xylem substrates (e.g. glucose, glutamine), but no acetate production is observed in most of them, suggesting that the nutrient provided are still under the maximal oxidative capabilities of *R. solanacearum*. Interestingly, the only experiment for which acetate was detected is a culture on a "xylem-mimicking medium" where carbon sources are a mix of the metabolites found in xylem sap, at concentrations proportional to the amount in xylem sap (unpublished, described in Chapter 5, 5.2.2). This seems to indicate that the diversity of carbon sources present in xylem, as well as their availability through fast flow, allow *R. solanacearum* reaching *in planta* its maximal capabilities in oxidative metabolism and substrate assimilation.

### 4.4.4 Are depleted metabolites indicators of trophic preferences?

In our study, the clear decrease of glutamine and asparagine (reaching the absence of detection) proves they contribute the most to *in planta* proliferation. This corroborates with the fact that glutamine concentration is essential for *in planta* fast growth and high density of the bacteria. Interestingly, among the 42 carbon sources tested in an *in vitro* study of *R. solanacearum* trophic preferences (Baroukh et al., accepted), glutamine and asparagine are among the four carbon sources sustaining the best growth (growth rate and biomass/substrate yield) along with glucose and aspartate.

However, considering xylem as a continuous flow, some metabolite depletions could have been missed as they are incessantly provided. In the last result section of the article, we use a mutant strain unable to catabolize glucose/sucrose/fructose and we showed that xylem sugars are not essential for full proliferation and virulence, but the symptoms and physiological markers of the infection are slightly delayed for the mutant strain. This seems to indicate that sugars are catabolized by *R. solanacearum in planta*, but their very low concentration makes them: i) accessory and complementary carbon sources, ii) difficult to quantify during the infection. In line with this hypothesis, no catabolic repression was reported in *R. solanacearum* from *in vitro* growth assays on combinations of carbon sources (Baroukh et al., accepted).

In conclusion, *R. solanacearum* appears to have evolved to optimize its growth in xylem vessels: it is both efficiently growing on the major source glutamine and able to assimilate most of the other metabolites present in xylem sap. By this excellent

metabolic adaptation, the pathogen is able to provoke an extremely rapid disruption of its host physiology, as described in the article. This fast growth and aggressive behavior of *R. solanacearum*, often emphasized when compared to other bacterial plant pathogens, could be strongly connected to its large and efficient metabolic capabilities.

# Chapter 5

# Analysis of host resource exploitation and limiting factors for proliferation *in planta* using a plant-pathogen model

## 5.1 Building a plant – pathogen metabolic model

In Chapter 3, I developed the model of a whole tomato plant during vegetative growth with an emphasis on xylem exchanges. In chapter 4, I characterized physiological and metabolic changes during an interaction between the tomato plant and the pathogen *R. solanacearum*. The combination of both of these results makes possible the construction of a plant-pathogen metabolic model presented in this chapter. The results remain preliminary and will be more deeply analyzed in the future.

In Chapter 2, section 2.4, I described three methodological points (summary below) that should be addressed before building a plant – pathogen metabolic model. Two of them were considered in the model presented in this chapter.

**i) Representing the interaction quantitatively**

This point was addressed by the acquisition of plant and pathogen growth rates, transpiration rate (as it is related to fluxes of matter in xylem) and a range of plant/pathogen realistic weight ratios (all presented in Chapter 3 and 4).

**ii) Representing resource availability at the interaction site**

This point was addressed by our quantification of xylem metabolite content (Chapter 3 and 4), and by a complementary study performed during my thesis consisting in *R. solanacearum* growth assay on a xylem-mimicking media. This study allowed determining assimilation and growth capacities of *R. solanacearum* in presence of the combination of carbon sources present in xylem.

I did not have time to address the third point during my thesis:

**iii) Addressing quasi-steady-state approximation**

This will limit the analysis to instantaneous simulations at different states of the infection process, considering that the system is at steady-state during short time intervals.

## 5.2 Methods

### 5.2.1 Generation of the plant – pathogen model

The cell tomato metabolic model constructed in Chapter 3 was used (SBML file). Reactions for manganese, cobalt, molybdenum, zinc assimilation, which were not represented in the initial metabolic model, were added as these metabolites are assumed to be essential for *R. solanacearum* growth. A multi-organ tomato model was generated using the same procedure than in Chapter 3. Briefly, using an in-house script, the model is replicated into three, representing respectively leaf, stem and root with adequate biomass reaction for each organ. Exchange reactions taking into account weight ratio between the different organs allow the exchange of metabolites through xylem and phloem. Xylem represents exchanges between root and leaves, phloem between leaves and roots. Stem can exchange in both directions with xylem and phloem (**Figure 5-1**). A transport cost is implemented for all of these exchange reactions. Different physiological constraints (uptakes of minerals by roots only, no photosynthesis for root, limited photosynthesis for stem) are integrated to model the different roles of each organ.

The script adds *R. solanacearum* metabolic reactions taken from Peyraud et al. (2016) to the multi-organ model, generating a four-compartment model (**Figure 5-1**): leaf, stem, root, pathogen. The boundary metabolites were removed from the metabolic model, and external metabolites of the pathogen model are considered as xylem metabolites. Sink reactions allowing the exit of some metabolites produced by *R. solanacearum* and excreted in the xylem are added. Ultimately, plant reactions for assimilation/degradation of all the metabolites potentially excreted by *R. solanacearum* (446) should be manually added, but this work was not done yet, and our simulations revealed so far that very few of sink reactions were used, and at very low fluxes (see the sink reactions activated in **Appendix V**). The weight ratio between *R. solanacearum* and plant is implemented through the exchange reactions between

xylem to the bacterial periplasm, and can be modified to model different bacterial densities *in planta*.

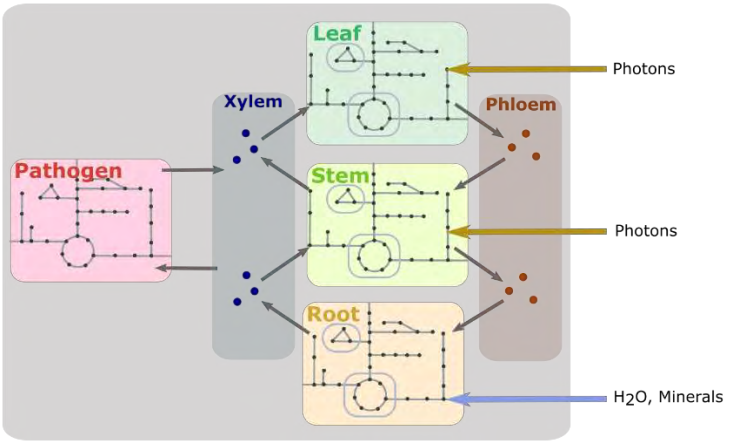



**Figure 5-1: Schematic representation of the plant - pathogen metabolic model**

Pathogen, leaf, stem and root are each modeled by a complete cell metabolic model with different constraints. Xylem and phloem are exchange compartments allowing the transport of metabolites between the cell metabolic models. The input of the system are light (photon), water and minerals that will sustain pathogen and plant growth.

### 5.2.2 Calibration of *R. solanacearum* metabolic model for growth *in planta*

To determine growth, excretion and assimilation capacities of *R. solanacearum* in the presence of xylem carbon sources, a shake-flask growth experiment was performed in a minimal medium supplemented with a mixture of organic substrates whose relative concentrations mimics a tomato xylem. Growth was monitored using OD measurements at 600 nm, and substrate assimilation and product excretions were quantified using NMR (Nuclear Magnetic Resonance) analysis. As already mentioned in Chapter 4, no catabolic repression was observed. The carbon sources were all assimilated at different rates. To determine uptake, excretion and growth rates among this complex combination of carbon sources, parameters of a macroscopic model that fits the experimental data of growth/assimilation/excretion were estimated.

I participated to the shake-flask growth experiment but not to the data analysis or the macroscopic modeling, achieved by C. Baroukh and further used in a study based on macroscopic modeling (article in preparation). The complete experimental protocol and the parameters obtained are described in **Appendix VI**.

Briefly, glutamine and glucose assimilation follows Monod-type kinetics, while other substrate assimilation is proportional to the assimilation of glutamine. Hence, growth

depends on maximal capabilities of growth on glucose and glutamine, and two maximal growth rates, one for glutamine and one for glucose, are implemented. Excretion of putrescine depends on both assimilation of glutamine and glucose, and associated excretion coefficients were determined.

Finally, the uptake/growth/excretion rates can be determined using:

- $k_{Gln}$, $k_{Glc}$: coefficient of glutamine and glucose assimilation (mmol·g [dry weight]$^{-1}$)
- $k_{Other\ substrate}$: coefficient of each other substrate assimilation, proportional to glutamine assimilation (mmol substrate·mmol Gln$^{-1}$)
- $K_{s,Gln}$, $K_{s,Glc}$: half-saturation constant of glutamine and glucose assimilation to follow Monod-type kinetics (mM)
- Concentration of glutamine and glucose (mM)
- $\mu_{max,Gln}$, $\mu_{max,Glc}$: growth rates on glutamine and glucose when other xylem substrates are present ($h^{-1}$)
- $k_{putr,Gln}$, $k_{putr,Glc}$: coefficients of putrescine excretion, respectively due to growth on glutamine and glucose (mmol·g [dry weight]$^{-1}$)

In the plant – pathogen metabolic model, uptake rates were implemented as upper bounds of the transport reaction from xylem to the periplasm of the bacterium. Fumarate and alanine uptake rates were unknown as they were not included in the shake-flask growth experiment. The uptake rates of the two were set to zero. They are among the carbon sources with the lowest concentration, so their contribution to the growth probably not impacts importantly the results.

Putrescine flux was constrained using a proportionality rule between glutamine/glucose uptake flux and putrescine excretion flux. The ATP maintenance reaction flux of the pathogen (NGAM, see Chapter section 2.3.3) was set to reach the experimentally obtained maximal growth rate. To represent correctly the differences between glutamine and glucose-dependent maximal growth rates, an additional maintenance cost (ATP consumption) was implemented on glucose uptake equation. Constraints are explicated in **Appendix V**.

### 5.2.3 Modeling of plant – pathogen interaction

Modeling of the interaction was performed using four successive FBAs (see Chapter 2, section 2.3.3) with different objective functions.

### i) Minimization of photon uptake

This first FBA is performed with leaf, stem and root growth rates constrained to experimental values, as in the simulation on the plant alone (Chapter 3). No growth of the pathogen is imposed and the pathogen ATP maintenance reaction (NGAM) is not constrained. Minimizing photon uptake allows acquiring the plant physiological capacities (photosynthesis and nitrogen uptake) in the absence of a pathogen.

### ii) Maximization of the pathogen biomass

The photon and nitrogen uptake rates determined in the previous FBA (representing plant physiological capacities) are implemented as additional constraints, to avoid that the plant uptakes more nitrogen or performs more photosynthesis than in a healthy plant. ATP maintenance reaction of the pathogen is now constrained, but the leaf and stem growth rates are relaxed, to model the fact that *R. solanacearum* resource acquisition and proliferation will hinder the possibility of the plant to grow optimally. Root growth rate remains constrained as we found in the experiments (Chapter 4, article) that root growth was poorly impacted by *R. solanacearum*. The objective function is now the maximization of the pathogen biomass reaction (i.e. maximal proliferation). This gives as output the maximal growth of *R. solanacearum* constrained by the plant physiological capacities.

### iii) Maximization of the growth of the aerial parts

The pathogen growth rate is now implemented as an additional constraint, and leaf/stem biomass reactions are maximized. To this end, maximization of leaf biomass reaction flux is set as objective, and a constant ratio between leaf and stem biomass is imposed. This allows determining the plant growth still achievable in the presence of *R. solanacearum*, as the pathogen deprives the plant from a part of its nutrients.

### iv) Minimization of the sum of absolute fluxes

To avoid the presence of Stoichiometric Balanced Cycles, a minimization of all the fluxes is performed as in the model of the plant alone (Chapter 3, article). This is computed under all the previous constraints, and also the leaf/stem growth rate obtained on the previous FBA.

Simulations were performed with Python 3.5 scripts, the open access libraries lxml, pandas and the linear programming solver CPLEX Python API, developed by IBM and free for academic institutions.

### 5.2.4 Xylem fluxes of matter and transpiration reduction

For all simulations, the xylem fluxes of matter obtained for healthy plants are integrated as upper bound in the plant-pathogen model, which allows the plant to transport less metabolites than what is observed in healthy plants, but avoids the plant to transport more.

Measurements of transpiration during plant – *R. solanacearum* interaction revealed an important decrease of transpiration when the pathogen is present at more than $10^7$ cells per g of stem fresh weight (FW) (see Chapter 4, Figures 1 and 6 of the article). I estimated by linear regression a transpiration decrease coefficient, dependent of the bacterial density (in decimal logarithm):

*Transpiration reduction coefficient = -0.2746* [log 10 cells per g of stem FW] + 2.8764*

As xylem fluxes of matter depend on the transpiration rate, the upper bound of xylem fluxes was constrained by this transpiration reduction coefficient if the bacterial density has reached more than $10^7$ cells per g of stem FW.

### 5.2.5 Modeling a *R. solanacearum* strain unable to assimilate a substrate

The phenotype of a *R. solanacearum* strain unable to assimilate one or several substrates is modeled by following the sequence of the four FBAs presented in 5.2.3, but with the upper bound of the associated transport reaction set to zero (e.g. simulating the inability to assimilate glutamine by setting the flux of glutamine transport from xylem to periplasm to zero).

### 5.2.6 Prediction of maximal colonization capacity through xylem fluxes

I also analyzed at which maximal density *R. solanacearum* can theoretically grow with xylem fluxes as source of nutrients using an alternative and simplified simulation. I removed constraints related to plant physiology: plant ATP maintenance reaction flux, organ growth rates, and limitation in photon and nitrogen uptake. The only remaining plant-related constraint were the upper bound of xylem fluxes of matter, with or without the transpiration reduction coefficient. Bacterial growth is still constrained (uptake rates, excretion rates, ATP maintenance) as presented in 5.2.2.

The simulation consisted in: i) Maximization of the pathogen biomass, ii) Minimization of the sum of absolute fluxes.

## 5.3 Results and Discussion

### 5.3.1 Bacterial density required to disrupt the plant physiology through the exploitation of nutrient resources

Modeling of the interaction by solving the multiple FBAs described in 5.2.3 allowed us to define three distinct phases of xylem colonization depending on the bacterial density (**Figure 5-2**). At each phase, different metabolic and physiological behaviors were observed.

**Phase 1: Commensal behavior**

Under $10^6$ cells per g of stem FW, *R. solanacearum* has a "commensal" behavior. At low cell density, the pathogen is able to grow optimally using plant resources, since xylem fluxes are much larger than *R. solanacearum* nutritive requirements. Also, since the cell density is really low, the assimilation of nutrients by *R. solanacearum* is negligible compared to main plant fluxes, and growth of aerial parts remains optimal and not impacted by the loss of nutrients to the pathogen. During this phase, most of the xylem carbon sources are simultaneously assimilated.

**Phase 2: Parasitic behavior**

Between $10^6$ and $10^8$ cells per g of stem FW, *R. solanacearum* has a "parasitic" behavior for the plant: its proliferation is detrimental for the plant growth. At around $10^6$ cells per g of stem FW, the pathogen density is too high to sustain both pathogen and plant growth with the resources available to the plant. Thus, the aerial parts growth rates start decreasing. However, the fluxes of matter in xylem remain under the upper limit set, suggesting that the concentration of organic carbon in xylem is not a limiting factor for *R. solanacearum* growth. At around $10^7$ cells per g of stem FW, a shift occurs: *R. solanacearum* has such a density that xylem metabolite uptake deprives totally the aerial part of the plant and thus forbids their growth. At densities above $10^7$ cells per g of stem FW, the resources start becoming also limiting for the pathogen growth, and its growth rate decreases, and reaches less than 0.05 $h^{-1}$ at $10^8$ cells per g of stem FW. In accordance with growth rate, the uptake rates of xylem metabolites per unit of pathogen biomass start decreasing.

**Phase 3: Growth arrest**

Above $10^8$ cells per g of stem FW, the pathogen reaches a stationary phase: its growth rate, already very low, will gradually reach zero as pathogen density increases and the

plant resources can no longer support proliferation. This behavior is similar to the stationary phase reached in a liquid culture with substrates exhausted. It could be considered that, due to the advanced disease state usually found at these pathogen densities, the bacteria will be able to find nutrients from the progressive necrosis of stem, sustaining an additional growth not considered in this model.

Interestingly, this model reveals that the competition for resources occurs on a relatively short bacterial density interval, between $10^6$ and $10^7$ cells per g of stem FW. In Chapter 4, tracking of physiological and metabolic parameters revealed that there was a tipping point in the disease progression when the bacterial density switches from $10^6$ to $10^8$ cells per g of stem FW, provoking an arrest of plant growth. Thus, the model is consistent with physiological experiments without any additional specifications (unless transpiration decrease, that will be discussed in the next section) to fit the data. This indicates that the calibration performed (e.g. organ ratios, growth rates, bacterial uptake rates, limitation of photon and nitrogen) is sufficient to predict quantitatively, with a metabolic model, the phases of the infection observed experimentally. Because the plant has finite capacities for photosynthesis and soil nutrient uptake, it also suggests that the hijacking of resources by the bacteria plays an important role to explain the plant growth stop and the rapid decline of plant physiology observed experimentally. Indeed, no additional constraint to represent a growth decrease induced by molecular determinants of pathogenicity (see Chapter 1, section 1.2) was necessary to fit the experimental data. Pathogenicity determinants are essential to penetrate into the plant, and to pass through the different tissues to reach xylem vessels and to potentially deactivate or attenuate immune response. However, rapid stop of plant growth, and by extension of plant "healthy" physiological state, can be explained, thanks to the model, only by the bacterial density observed *in planta* a few days after the infections. As mentioned in Chapter 4 (article Discussion), tomato plants that are tolerant to *R. solanacearum* (i.e. no wilting symptom but bacterial colonization in vascular tissues) afford up to $10^6$ cells per g of stem FW (Nakaho et al., 2004), consistent with the prediction of the model. This illustrates also that *R. solanacearum* can have, experimentally, a behavior described as commensal. It confirms that competition for resources triggered by a highest density will take an important part of the observed symptoms.

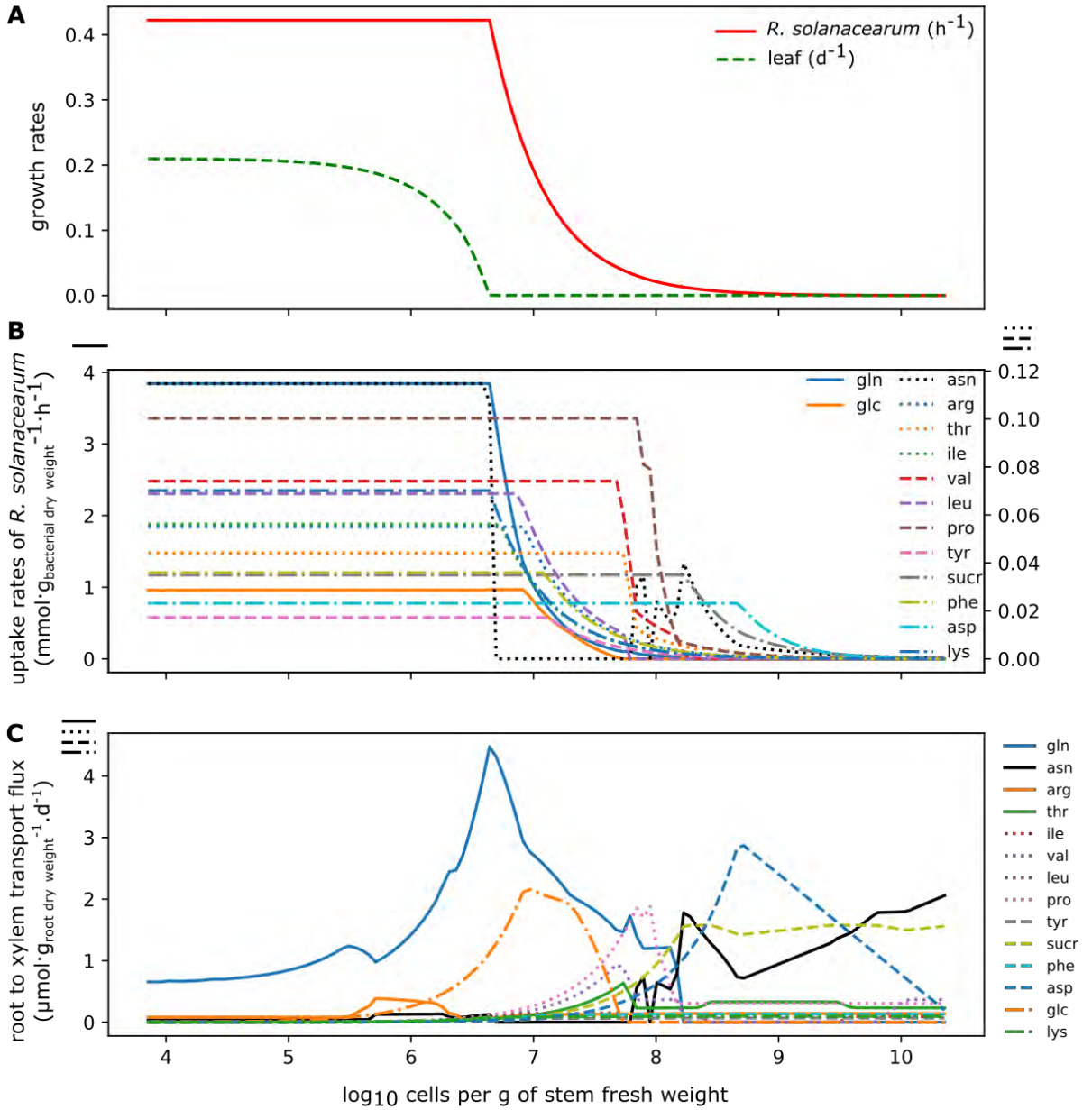



**Figure 5-2: Growth, uptake and xylem fluxes in the plant - pathogen metabolic model.**

A. Progression of leaf and *R. solanacearum* growth rates at different bacterial densities. These growth rates correspond to the biomass reaction flux in the metabolic model, after the four FBA described in Methods section. Leaf and stem growth rates are proportional in the model, so the same curve pattern would be obtained.

B. Uptake of xylem metabolites by *R. solanacearum* at different bacterial densities. The uptake rates correspond to the flux in the reaction transporting the metabolite from xylem to *R. solanacearum* periplasm, after the four FBA described in the Methods section. On this graph, full lines curves correspond to the left ordinate scale, while other lines correspond to the right ordinate scale.

C. Transport of metabolites from root to xylem vessels at different bacterial densities. This transport corresponds to the flux of metabolite from root cytoplasm to xylem, after the four FBA described in the Methods section.

### 5.3.2 The decrease in transpiration at high bacterial density is not a limiting factor in the interaction

I performed a similar simulation than in the previous section, but this time without considering the limitation of transpiration above $10^7$ cells per g of stem FW (see 5.2.4). The aim was to find if the transpiration constraint contributes to decrease of the pathogen growth rate above $10^7$ cells per g stem FW. Leaf and pathogen growth rates are similar with or without this constraint (**Figure 5-3**). This indicates that the decrease of transpiration occurs at densities too high to impact growth in our model. As it is not affecting growth, we suspect that transpiration decrease is not a cause of the plant physiological disruption that occurs earlier due to resource hijacking, but rather a consequence of it and of the xylem obstruction by *R. solanacearum*. In addition, this comparison reveals that the decrease of transpiration is not a limiting factor in the interaction: available resources (nitrogen, photon) become too low to sustain plant and pathogen growth before the transpiration is impacted.

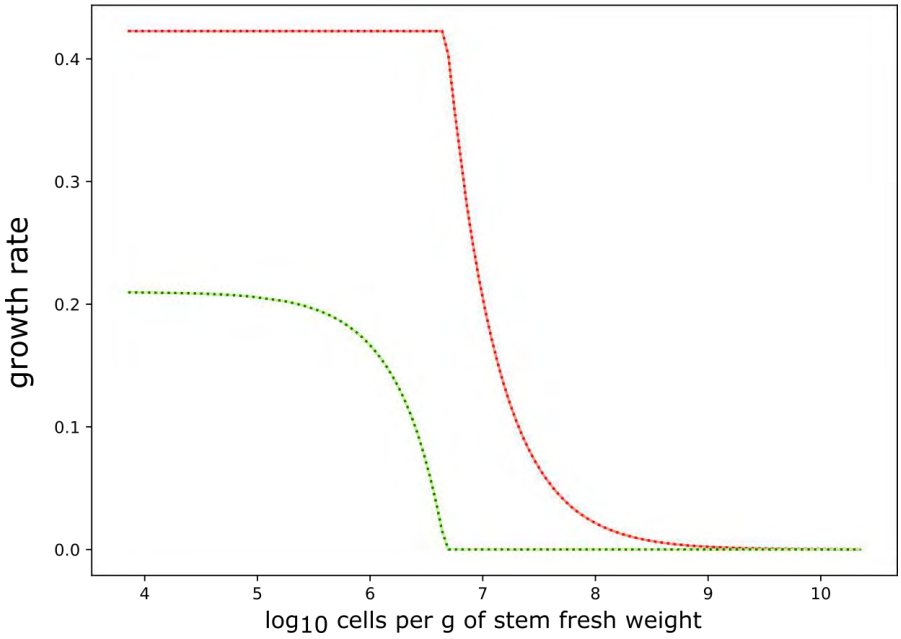



**Figure 5-3: Analysis of the impact of transpiration limitation coefficient on growth rates.**

Full lines: transpiration taken into account through reduction coefficient. Dotted lines: transpiration limitation through not taken into account. Red: *R. solanacearum* growth rate ($h^{-1}$) from biomass reaction flux. Green: leaf growth rate ($d^{-1}$) from biomass reaction flux.

### 5.3.3 Effects of the inability to assimilate sugars or glutamine

I performed the same four successive FBAs required to model plant – pathogen interaction but imposed the inactivation of xylem sugar or glutamine transport systems in the pathogen (**Figure 5-4**). This allows simulating the behavior of a mutant with deleted transporter gene(s), such as the one phenotyped in Chapter 4.

As glucose and glutamine are major contributors of growth in the first phase of the infection, when as *R. solanacearum* is at low density (under $10^6$ cells per g of stem FW), growth rates are impacted importantly on this phase. *R. solanacearum* grows at 80% of its maximal capabilities if it is not able to assimilate glucose/sucrose, versus 30% if it is not able to assimilate glutamine. These values must be taken with a grain of salt as the model is calibrated for a growth on the combination of substrates: growth on glutamine (glucose) alone, if not limiting, is around 70% (50%) of xylem-mimicking medium growth.

At the intermediate phase where resource become more limited for the pathogen, growth rates of wild-type strain or strain with assimilation inability are getting closer and will converge. It happens at lower densities for the strain prohibited of sugars uptake: at $10^7$ cells per g of stem FW, the strain is at 96% of the wild-type growth rate at the same density, while for the strain with glutamine uptake prohibited it is at 65% of the wild-type growth rate. The progressive decrease of the plant growth rate looked almost similar when simulating infection by the wild-type strain or a mutant unable to assimilate sugars. However, there is a significant difference in the density required to disrupt plant growth when comparing the wild-type and the mutant unable to uptake glutamine (**Figure 5-4**).

Overall, these results show that disabling both uptake systems (sugar/glutamine) will delay the disease progress, as a slowly proliferating *R. solanacearum* will require more days to reach the density that impacts plant physiology (around $10^7$ cells per g of stem FW). However, our simulations suggest that the metabolization of sugars is less necessary as: i) bacterial growth rate is impacted only partially (standing above 50% of the wild-type rate), ii) the growth of the wild-type and mutant strain deficient for sugar assimilation at $10^7$ cells per g of FW are similar, so sugar assimilation probably plays a significant role only at the early stage of the plant-pathogen interaction, iii) the growth rate decrease of the leaf almost overlaps with the growth rate decrease obtained with wild-type pathogen. This is in line with the mutant strain phenotyped in Chapter 4:

deletion of fructose/sucrose/glucose transport genes led to a slight delay in the disease symptoms and physiological markers, but full plant wilting was still observed at the end.

At the contrary, the inability to assimilate glutamine has a greater effect on growth, and the pathogen will remain for a longer period at low densities. Also, the density required to affect plant physiology is higher, so cumulating these constraints could give enough time for the plant to limit its proliferation to a density that is not provoking disease. No mutant strain with deleted genes of glutamine transporters was generated so far, but the predicted differences between the phenotypes of the two strains suggest that having the glutamine transport mutant could be a promising perspective to validate that glutamine is a crucial carbon sources *in planta*.

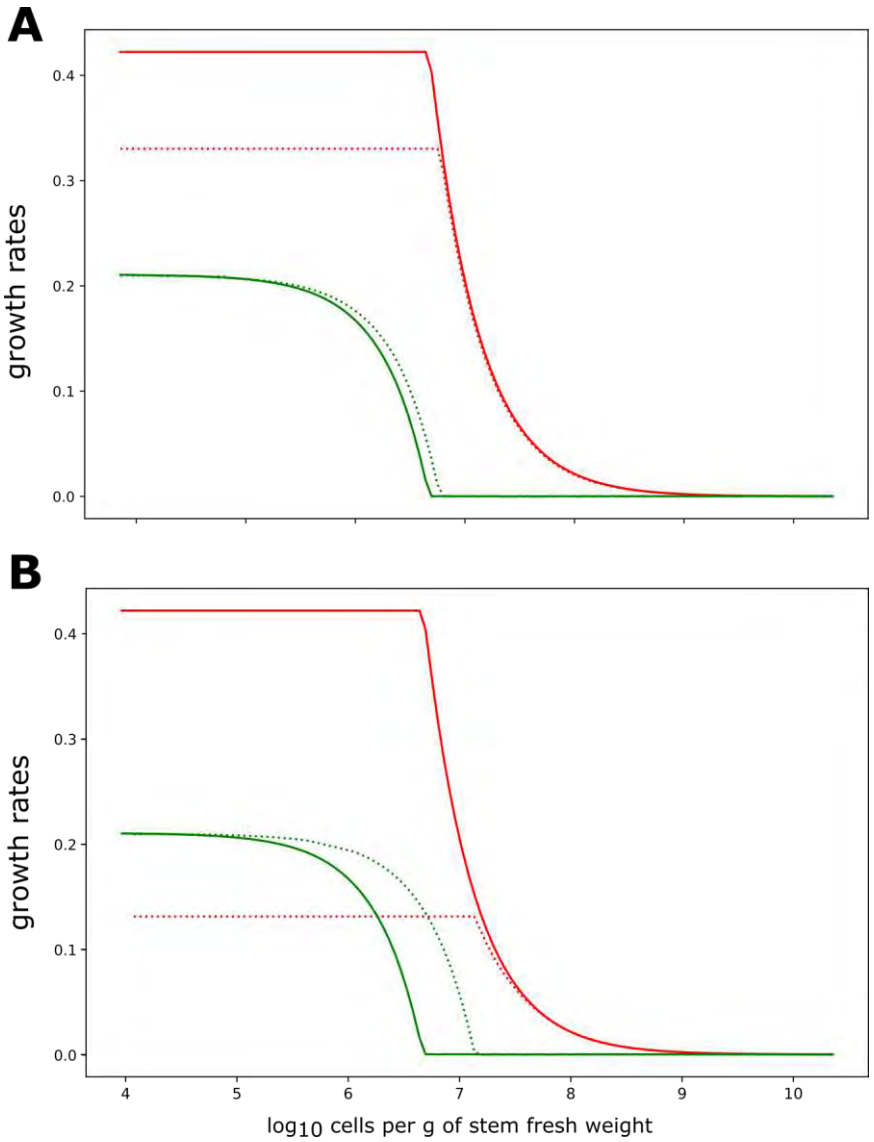



**Figure 5-4: Comparison of growth rates of plant and pathogen with different inabilities to assimilate substrates.**

Full lines: model with a wild-type *R. solanacearum* strain. Dotted lines: model with a *R. solanacearum* unable to assimilate either in A. glucose and sucrose or in B. glutamine. Red: *R. solanacearum* growth rate ($h^{-1}$) from biomass reaction flux. Green: leaf growth rate ($d^{-1}$) from biomass reaction flux.

### 5.3.4 Xylem fluxes of matter sustain the growth of high bacterial densities

The previous results suggested that in our model, the flux values in xylem were not limiting factors in the interaction, contradicting the recurrent view that xylem sap is an environment diluted in carbon sources for vascular pathogen growth.

I modeled (see 5.2.6) the growth of *R. solanacearum* with almost all plant related constraints removed by keeping only xylem fluxes as limiting factors, to determine the

maximal density of the pathogen theoretically achievable (**Figure 5-5**). I found that up to $3.10^{10}$ cells per g of stem FW could grow with the xylem fluxes of matter measured in a healthy plant, which is in the range of the maximal densities that can be found experimentally in plant stems. The growth rate remained high between $10^9$ and $10^{10}$ cells per g of stem FW. The inability to grow above this threshold was due to the coefficient of transpiration reduction (see 5.2.4) which led to almost zero fluxes in xylem. If we also neglect this physiological behavior, the growth of *R. solanacearum* can be sustained up to almost $10^{13}$ cells per g of stem FW (largely above the densities found experimentaly). A strong growth can be reached until $10^{10} – 10^{11}$ cells per g of stem FW.

These results highlights that the straw-man argument that the "xylem is nutrient-poor" should not be considered anymore, for several reasons: i) before reaching high densities (here around $10^7$ cells per g of stem FW), vascular pathogen population represents an almost negligeable weight proportion compared to the plant, and optimal growth can be achievable without quantitatively interfering with plant resources, ii) xylem fluxes can theoritically support the fast growth of a bacterial density similar to the highest densities observed in planta ($10^9 – 10^{10}$ cells per g of stem FW), iii) other factors will be limiting the pathogen growth when $10^7$ cells per g of stem are reached (**Figure 5-2**) before the xylem content becomes limiting at $10^9$ cells per g of stem FW (**Figure 5-3**). Also, our model was built and calibrated with *R. solanacearum*, which is likely one of the xylem-colonizing bacteria with the highest growth rate. Since growth/uptake capacities of other pathogens such as *X. fastidiosa* and *Xcc* or other vascular pathogens (such as fungi) are lower (see Chapter 1, section 1.5.3), if xylem concentrations are not limiting for *R. solanacearum*, it should be the same for these organisms.

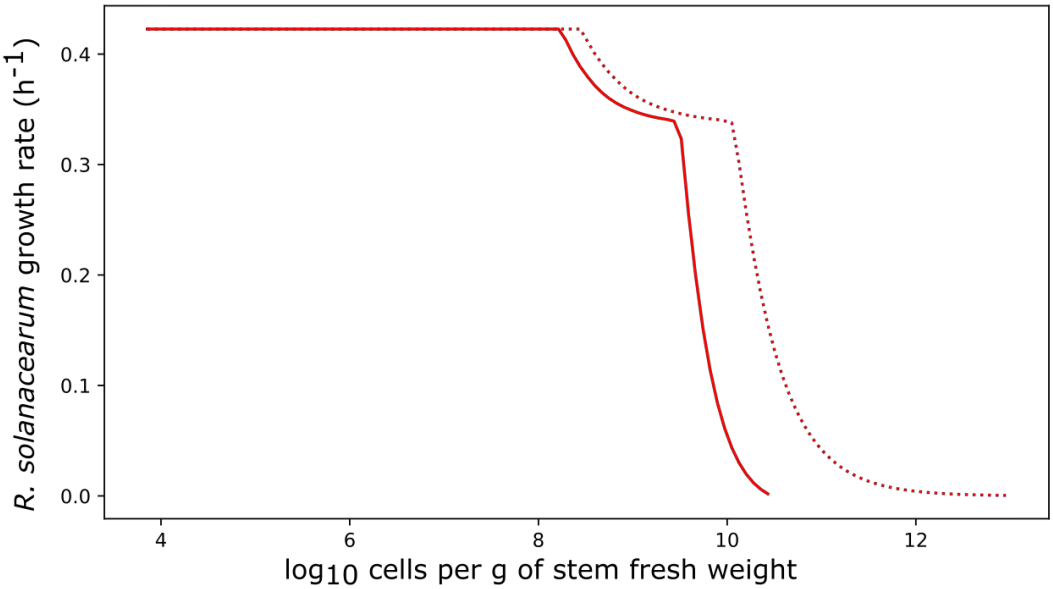



**Figure 5-5: Maximal colonization and growth capacities of *R. solanacearum* under xylem fluxes.**

Growth rates (through biomass reaction flux) were determined by modeling the interaction without plant growth constraints, excepted xylem fluxes upper bounds. Full lines: transpiration taken into account through reduction coefficient. Dotted lines: transpiration limitation through reduction coefficient not taken into account.

### 5.3.5 Toward a dynamic metabolic model of the interaction

Only a static model of plant – pathogen interaction was developed so far. An interesting perspective would be to develop a dynamic model of the interaction. However, dynamical modeling imposes additional methodological considerations. Usual dynamic modeling techniques such as dFBA (Mahadevan et al., 2002) rely on representing the biological system using a set of ODEs (Ordinary Differential Equations), whose dynamics evolves by solving successive FBAs at each time intervals. This could be easily done to simulate growth on xylem-mimicking medium, either with limited resources as in a batch culture, or considering continuous supply of nutrients (fed-batch culture) or supply plus washout of the medium (chemostat culture), as either fed-batch or chemostat cultures should better mimic what is happening *in planta* (see Chapter 4, section 4.4.3). Dynamic modeling of the plant is more complex since the growth depends on light uptake and water/soil nutrients assimilation (autotrophy) difficult to measure. Thus, we could, for example, simulate the growth of *R. solanacearum* at different time points through a dynamic model, and at each time point, subsequently optimize plant growth with the limitation imposed by *R.*

*solanacearum*. Hence, the dynamics on the system will be imposed by *R. solanacearum* growth and substrate exhaustion, and plant metabolism will be subsequently modeled according to it.

Once these methodological issues are addressed, dynamic metabolic model would be of great interest for a deeper understanding of the interaction. Simulating a time-kinetics of *R. solanacearum* – tomato interaction would allow an easier comparison with plant phenotyping experiments, since following the infection over time is more straightforward than depending on the bacterial density. We could compare the tipping point found experimentally (Chapter 4) with the dynamic model, as well as the delay of the infection with a mutant deprived of sugar or glutamine transporters.

# Chapter 6

# Genome-scale metabolic model of *Xylella fastidiosa* and metabolic diversity among xylem-colonizing bacteria

## 6.1 Deciphering the lifestyle-metabolism relationships of xylem-colonizing bacteria

In the last years, the metabolism of *R. solanacearum* (strain GMI1000) was studied increasingly. My work and previous studies from my team gathered metabolomics, physiology and modeling of the bacteria alone (Peyraud et al., 2016, 2018, Baroukh et al., accepted), its host (Chapter 3) and the interaction (Chapters 4 and 5). This topic is also increasingly studied in other laboratories (Hamilton et al., 2021; Jyoti et al., 2020; Khokhani et al., 2017; Lowe-Power, Hendrich, et al., 2018; G. Wang et al., 2019). This combination of studies gives a broader representation of the biology of a plant pathogen and its interaction with the host, enriching and completing the studies done on the molecular aspects of the interaction.

While there is still a lot of promising perspectives to pursue on *R. solanacearum*, both on experiments or modeling, comparing the metabolism of *R. solanacearum* with other xylem-colonizing bacteria would give us a broader view of metabolism in the interaction. One can expect that differences in the lifestyle (e.g. environments the bacteria encounter, virulence strategies) would leave an imprint on the metabolic potential of the organisms, which can be studied using genome-scale metabolic modeling. In Chapter 1 (section 1.5.3), I presented three xylem-colonizing bacteria that have contrasted lifestyles and intrinsic properties such as genome size or growth capability. A perspective would be to perform similar studies than the ones performed

on *R. solanacearum* on the two other organisms: *Xcc* and *X. fastidiosa*. I believe that simple metabolic reconstruction and physiological studies could already reveal traits related to their life cycles, with a similarity for these three bacteria in the type of environment they colonize (the vascular system of plants) but distinct in many other aspects.

In collaboration with IRHS (Institut de Recherche en Horticulture et Semences) as part of the European project XF-ACTORS (Xylella Fastidiosa Active Containment Through a multidisciplinary-Oriented Research Strategy), I reconstructed and modeled the metabolism of *X. fastidiosa*. Collaborators from IRHS were able to characterize the carbon assimilation capacities of *X. fastidiosa*, which were integrated in the model. One of our central question was to determine if the very slow growth of *X. fastidiosa* (quite different to *R. solanacearum)* was related to specificities in its metabolism.

## 6.2 Genome-scale metabolic model reconstruction of *X. fastidiosa* [article]

**Article**

Genome-scale investigation of the metabolic determinants generating bacterial fastidious growth

**Journal**

*mSystems* (published online 31 March 2020)

**Contributions**

- Metabolic reconstruction and curation (with L. Cottret)
- Simulation and analysis of the model (with C. Baroukh)
- Analysis of the Biolog experiment (carbon substrate phenotyping) from the raw experimental data
- Writing the complete draft article and designing the figures
- Editing the article (with the co-authors)

**Supplementary Materials** available in **Appendix VII**.

**Summary**

The metabolic model of *X. fastidiosa* was reconstructed by orthology, using five high-quality genome-scale metabolic models, and manually curated. Experimental phenotyping of carbon usage was performed and integrated in the model. The final model was analyzed and compared with other bacteria such as *R. solanacearum*, *E. coli* or *B. pertussis* using gene essentiality, enzyme redundancy, FVA and estimation of metabolic yields. Results obtained revealed that the metabolism of *X. fastidiosa* was reduced and fragile in comparison with other bacteria, at the pathway, reaction and metabolic gene scales. This loss of robustness affected some parts of the metabolism that were not essential but important for an optimal growth, indicating that several losses in the metabolism could be (at least partially) responsible of the observed slow growth.

AMERICAN SOCIETY FOR MICROBIOLOGY | mSystems®

RESEARCH ARTICLE
Host-Microbe Biology

# Genome-Scale Investigation of the Metabolic Determinants Generating Bacterial Fastidious Growth

**Léo Gerlin,[a] Ludovic Cottret,[a] Sophie Cesbron,[b] Géraldine Taghouti,[b] Marie-Agnès Jacques,[b] Stéphane Genin,[a] Caroline Baroukh[a]**

[a]LIPM, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France
[b]IRHS, INRAE, AGROCAMPUS-Ouest, Université d'Angers, Beaucouzé, France

**ABSTRACT** High proliferation rate and robustness are vital characteristics of bacterial pathogens that successfully colonize their hosts. The observation of drastically slow growth in some pathogens is thus paradoxical and remains unexplained. In this study, we sought to understand the slow (fastidious) growth of the plant pathogen *Xylella fastidiosa*. Using genome-scale metabolic network reconstruction, modeling, and experimental validation, we explored its metabolic capabilities. Despite genome reduction and slow growth, the pathogen's metabolic network is complete but strikingly minimalist and lacking in robustness. Most alternative reactions were missing, especially those favoring fast growth, and were replaced by less efficient paths. We also found that the production of some virulence factors imposes a heavy burden on growth. Interestingly, some specific determinants of fastidious growth were also found in other slow-growing pathogens, enriching the view that these metabolic peculiarities are a pathogenicity strategy to remain at a low population level.

**IMPORTANCE** *Xylella fastidiosa* is one of the most important threats to plant health worldwide, causing disease in the Americas on a range of agricultural crops and trees, and recently associated with a critical epidemic affecting olive trees in Europe. A main challenge for the detection of the pathogen and the development of physiological studies is its fastidious growth, as the generation time can vary from 10 to 100 h for some strains. This physiological peculiarity is shared with several human pathogens and is poorly understood. We performed an analysis of the metabolic capabilities of *X. fastidiosa* through a genome-scale metabolic model of the bacterium. This model was reconstructed and manually curated using experiments and bibliographical evidence. Our study revealed that fastidious growth most probably results from different metabolic specificities such as the absence of highly efficient enzymes or a global inefficiency in virulence factor production. These results support the idea that the fragility of the metabolic network may have been shaped during evolution to lead to the self-limiting behavior of *X. fastidiosa*.

**KEYWORDS** metabolic network, metabolic pathways, metabolic modeling, robustness, pathogen, growth, *Xylella fastidiosa*, *Xanthomonas*, fastidious

Optimal growth rate and robustness against environmental perturbations are major properties contributing to bacterial fitness (1–3). For example, experimental evolution of the plant pathogen *Ralstonia solanacearum* on several hosts revealed that bacteria acquired both enhanced growth rate and better robustness through enlarged catabolic capacities, showing that these phenotypes are under selective pressure (4, 5). The robustness of virulence functions in this pathogen is also illustrated by the plethora of effectors transiting through the type 3 secretion system, as well as for the regulatory and metabolic networks (6, 7). These traits were shown to be crucial to invade the host and bypass its immune system (7, 8). However, it is also well documented that some

**Citation** Gerlin L, Cottret L, Cesbron S, Taghouti G, Jacques M-A, Genin S, Baroukh C. 2020. Genome-scale investigation of the metabolic determinants generating bacterial fastidious growth. mSystems 5:e00698-19. https://doi.org/10.1128/mSystems.00698-19.

**Editor** Joshua E. Elias, Chan Zuckerberg Biohub

**Copyright** © 2020 Gerlin et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Stéphane Genin, stephane.genin@inrae.fr, or Caroline Baroukh, caroline.baroukh@inrae.fr.

Why is #Xylellafastidiosa "fastidious"? Genome-scale metabolic modeling unraveled specificities that could explain the slow-growing phenotype of the pathogen. This phenotype could be a pathogenicity strategy.

**Received** 23 October 2019
**Accepted** 5 March 2020
**Published** 31 March 2020

Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.

**TABLE 1** Comparison of metabolic networks

| Network characteristic | No. of reactions | No. of metabolites |
|---|---|---|
| *E. coli* | 2,583 | 1,805 |
| *R. solanacearum* | 2,644 | 2,574 |
| *B. pertussis* | 1,203 | 1,143 |
| *X. fastidiosa* | 1,158 | 1,107 |

bacterial pathogens are slow growing, qualified as fastidious, such as *Legionella* species (9), *Brucella* species (9), and *Bordetella pertussis* (10–12). Fastidiousness refers to an arduous laboratory handling, and these difficulties were first hypothesized as a consequence of a lack of knowledge of nutritional and environmental requirements. However, later studies reported that fastidiousness is mostly, or complementarily, related to an intrinsic slow growth (13, 14) which could not be overcome by optimized culture medium (13).

*Xylella fastidiosa* is the infectious agent of several diseases affecting plants of agronomical interest (15). This threatening bacterium is transmitted through xylem-sap-feeding insect vectors (15). Once inside the xylem capillary vessels, *X. fastidiosa* spreads, grows, produces virulence factors like exopolysaccharides (EPS), and forms biofilms (15, 16) before colonizing the foregut of a new insect host (15). *X. fastidiosa*'s fastidious growth was reported both *in vitro* and *in planta* and is responsible for its difficult isolation (17). Design of synthetic growth medium was carried out (18–20), but after nearly 2 decades of medium optimization, the persisting difficulty of growing *X. fastidiosa*, with a doubling time in the range of 100 h for some strains (19, 21), brings out that the slow growth is a property of the organism. As the genome size of the bacterium is severely reduced, it is tempting to speculate that part of the metabolic robustness is lost in *X. fastidiosa*. To understand the metabolic factors at the origin of fastidious growth, and the level of robustness of the bacterium, we reconstructed a high-quality genome-scale metabolic network of *X. fastidiosa*, based on genomic and experimental data. This network allowed us to explore the metabolic capabilities of *X. fastidiosa*. We unraveled a lack of metabolic robustness to sustain growth through flux variability, enzyme redundancy, gene deletion, and trophic capabilities, supported by both modeling and experiments. We then sought to deduce some metabolic determinants of fastidious growth. We found that loss of robustness mostly affected efficient parts of the network, such as enzymes with high catalytic activities, or central reactions which are assumed to play a crucial role in growth. Thereby, we unraveled that *X. fastidiosa* metabolism mostly relies on pathways and reactions which cannot ensure a fast proliferation. We demonstrated that the global network has a weak metabolic yield in producing central virulence factor, thus showing that it is structurally inefficient, which could also severely impair growth.

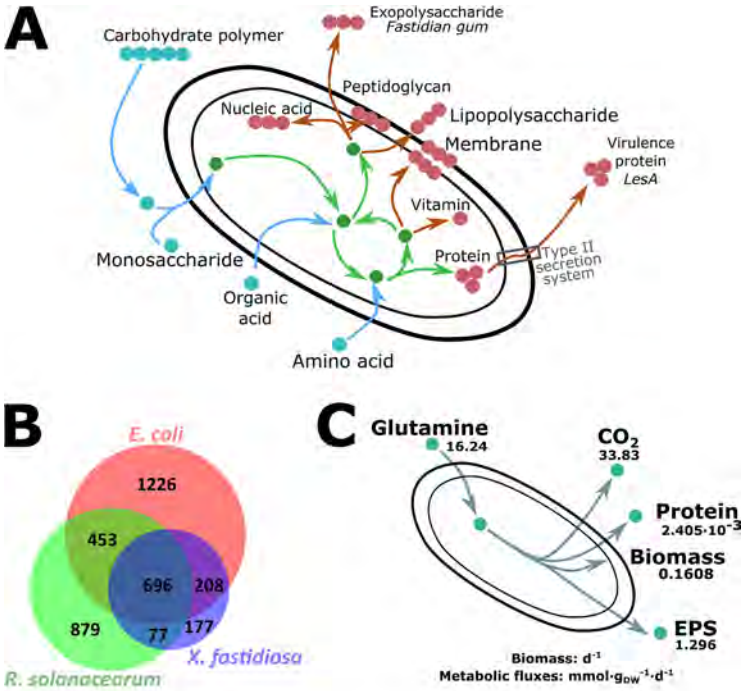
## RESULTS

**The metabolic network of *X. fastidiosa*.** We built a genome-scale metabolic network of *X. fastidiosa* (see Data Set S1, Text S1, and Text S2 in the supplemental material) using in-house automatic reconstruction algorithms (22) from the genomic sequence of *Xylella fastidiosa* subsp. *multiplex* strain CFBP 8418 isolated in 2015 in Corsica (23). A manual curation of each reaction was then performed using databases, literature, and simulations. Novel reactions were written to take into account the biosynthesis of EPS and lipopolysaccharides (LPS) (24–26). To represent the global cost of excreted proteins, which are crucial virulence factors, we included reactions for the biosynthesis and excretion of a proteic virulence factor through the type II secretion system (27). The reconstruction process is summarized in Data Set S2.

The final numbers of reactions and metabolites of the *X. fastidiosa* network (Table 1) are remarkably small compared to *Escherichia coli* and *R. solanacearum*. However, all the core carbon and nitrogen reactions from the central metabolism were found (except one enzyme; see below), as well as the biosynthesis of all vital compounds (Fig. 1A). The



Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.




D

| Enzyme(s) per reaction (%) | 1 | 2 | 3 | 4 | 5 | >5 |
|---|---|---|---|---|---|---|
| *X. fastidiosa* | 87 | 10 | 1 | 0 | 0 | 0 |
| *B. pertussis* | 81 | 9 | 4 | 2 | 1 | 2 |
| *E. coli* | 69 | 14 | 2 | 14 | 0 | 1 |
| *R. solanacearum* | 74 | 16 | 3 | 5 | 1 | 3 |
| *P. aeruginosa* | 70 | 14 | 6 | 3 | 2 | 5 |
| *R. eutropha* | 60 | 17 | 7 | 5 | 3 | 8 |

**FIG 1** General characteristics of the reconstructed metabolic network. (A) Global overview of the reconstructed metabolic pathways in *Xylella fastidiosa*. This schema highlights the metabolic behaviors of *Xylella fastidiosa*. For catabolism (external substrates and catabolic reactions in blue), *X. fastidiosa* is able to degrade polymers and assimilate them or catabolize elements from xylem sap such as amino acids and organic acids. Central metabolites and reactions are schematized in green (glycolysis and TCA cycle). Metabolites/reactions in red/brown highlight anabolic capabilities: the production of biomass macromolecules (membranes, lipopolysaccharides, peptidoglycans, vitamins, proteins, and nucleic acids) and the secretion of virulence factors (virulence protein through the type II secretion system and fastidian gum [EPS]). (B) Venn diagram depicting the number of metabolic reactions in *E. coli*, *R. solanacearum*, and *X. fastidiosa*. The lists of the reactions found in each part of the Venn diagram are available in Data Set S2. (C) Inputs and outputs of an FBA simulation of *X. fastidiosa* growth on glutamine under conditions close to experimental conditions. The objective function of FBA was minimization of glutamine import. Growth and production rates were calculated from previous publications (19, 82). Flux values are in mmol·g (DW)$^{-1}$·day$^{-1}$ for all inputs/outputs except for biomass (day$^{-1}$). See Text S3 for the constraints used for the simulation and Data Set S2 for the FBA solution. Scripts are available on https://github.com/lgerlin/xfas-metabolic-model. (D) Distribution of the number of enzymes for each metabolic reaction in *X. fastidiosa* and metabolic models from other bacteria. Proportions of reactions carried out by 1, 2, 3, 4, 5, and >5 enzymes were computed for each metabolic model. The complete data are available in Data Set S2.

precursors for the specific macromolecules EPS and LPS, which are key virulence determinants (28), can also be synthesized.

We performed genome sequence comparisons between *X. fastidiosa* subsp. *multiplex* strain CFBP 8418, *Xylella fastidiosa* subsp. *pauca* strain 9a5c (isolated from citrus in Brazil) (29), and *Xylella fastidiosa* subsp. *fastidiosa* strain Temecula1 (isolated from grapevine in California) (30) to estimate the metabolic network conservation between the three subspecies (Table S1 and Table S2). Because these differential reactions represented 1% or less of the network, with most of them being not connected to central metabolism, we concluded that the architecture of the metabolic networks is



Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.

probably highly similar between these three *X. fastidiosa* subspecies. Therefore, the conclusions drawn in the following sections on strain CFBP 8418 extend to at least three subspecies of *X. fastidiosa*.

**Predicted metabolic capacities of *X. fastidiosa*.** The comparison of the number of metabolic reactions shared between *X. fastidiosa*, *E. coli*, and *R. solanacearum* revealed that *X. fastidiosa* has undergone a strong metabolic reduction, which is associated with its reduced genome size, and illustrated by the decreased numbers of reactions that are specific to *X. fastidiosa* (Fig. 1B). The complete list of shared or specific reactions is available in Data Set S2. The metabolic core shared between the three organisms was around 700 reactions, which represents 58% of the *X. fastidiosa* model. *E. coli* shares a wide range of reactions with *X. fastidiosa*, including vitamin biosynthesis and transport processes. *X. fastidiosa* and *R. solanacearum* have fewer reactions in common, probably due to their more distant phylogenetic relationship (beta- versus gammaproteobacteria). Among them, we identified several metabolic pathways related to *in planta* lifestyle and pathogenicity such as nitrilase (EC 3.5.5) reactions, which detoxify nitriles (31), and peroxidases (EC 1.11.1) degrading radical oxygen species (ROS) (32), both nitriles and ROS being produced by plants upon infection (32, 33). These two plant pathogens also shared degradation reactions for plant polymers like pectin.

The comparative analysis revealed a large number of reactions (450) absent only in *X. fastidiosa*. Several of them are related to lipid metabolism, such as the pathway to use lipid as carbon source ($\beta$-oxidation) and to alternate carbon metabolism reactions, which enable the use of diverse compounds as carbon source, like carbohydrates (e.g., tagatose and glycogen) or small organic molecules (e.g., benzoate, D-galacturonate, and phenylacetate). These findings highlighted that *X. fastidiosa* metabolism appears limited in terms of metabolic diversity.

Although *X. fastidiosa* possesses all the vital vitamin and cofactor biosynthetic pathways, it lacks several related pathways that are generally found in fast-growing bacteria. For instance, the couple ubiquinone-8/ubiquinol-8 was the only redox couple from the quinone group found to be synthesized. In contrast to *R. solanacearum*, no cobalamin (vitamin $B_{12}$) biosynthesis is possible. Accordingly, *X. fastidiosa* uses only enzymes with no requirement for cobalamin, which cannot be found in plants (34). We observed a similar case for molybdopterin, which forms the molybdenum cofactor.

Previous studies suggested that *X. fastidiosa* was limited in aerobic respiration, because no cytochromes with a high affinity for $O_2$ were found in the genome sequences (35). The absence of cytochrome *c* oxidase was also reported, limiting the aerobic respiration to cytochrome $bo_3$ ubiquinone oxidase (EC 1.10.3.10). It was thus suggested that the bacterium preferentially uses anaerobic respiration (35). Contrary to this hypothesis, no complete anaerobic respiration was found to be functional in the metabolic network, which supports the view of a functional and favored aerobic respiration in *X. fastidiosa*. However, the limitation of aerobic respiration suggests that the bacterium is also limited in its respiratory system.

Succinate dehydrogenase (EC 1.3.99.1), an enzyme from central and respiratory metabolism composed of four subunits, was also found to have lost a membrane subunit in *X. fastidiosa*. This enzymatic complex was reported to be functional with only one membrane domain (36). We thus expect that the metabolic reactions can occur in *X. fastidiosa* but might also be less efficient than the complete form of the enzyme complex.

Our primary analysis of the network structure showed that the core metabolism is complete, and its functionality was confirmed by flux balance analysis (FBA) simulations (Fig. 1C and Data Set S2), accepting the fact that an FBPase activity exists (see below). Simulation (Data Set S2) results show that all biomass components can be produced from the tested carbon source, glutamine, a major component of xylem sap (37–40). The majority of glutamine uptake is converted into glutamate (95% of carbon uptake) and then into $\alpha$-ketoglutarate. $\alpha$-Ketoglutarate is then converted through the citric acid cycle (tricarboxylic acid [TCA] cycle) to malate and oxaloacetate. Yet, the TCA cycle is



Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.

not cyclic. The branch converting oxaloacetate to citric acid is not used, which could be expected for growth in minimal medium with glutamine as sole carbon source. The part of the citric acid cycle used generates central metabolites such as succinyl coenzyme A (CoA) and oxaloacetate. Pyruvate is generated from the citric acid cycle metabolite oxaloacetate, through oxaloacetate decarboxylase (EC 4.1.1.3, 22% of carbon uptake) and malic enzyme (EC 1.1.1.40, 34% of carbon uptake). Finally, these central metabolites are used to generate biomass and virulence factors (EPS and virulence proteins), and the remaining carbon is wasted during metabolic processes as $CO_2$, more than half through the TCA cycle. NADH is also mainly synthesized thanks to the TCA cycle and ATP through oxidative phosphorylation.

The core central metabolism from *X. fastidiosa* appears to be conserved and functional, although it seems to have lost several pathways compared to other bacteria. These losses do not affect the survival of the bacterium but tend to restrict it to a limited diversity of metabolic behaviors, and probably to a slow and inefficient growth.

**Unexpected features in *X. fastidiosa* metabolism.** A deeper analysis of the *X. fastidiosa* metabolism revealed the surprising lack of some enzymes which are generally ubiquitous in bacteria. First, no gene was identified to code for fructose-1,6-bisphosphatase (FBPase) (EC 3.1.3.11), despite intensive homology searches with query sequences from a wide range of organisms. The absence of the corresponding gene in strain CFBP 8418, as well as all the other available genomes of *X. fastidiosa*, suggested that these bacteria are not able to perform gluconeogenesis. We verified *in silico* that in our metabolic network, gluconeogenesis was not functional; a precursor such as glutamine could not be converted into glucose with no FBPase. It would imply that *X. fastidiosa* can grow only in the presence of a hexose or a hexose polymer (41, 42), but that was invalidated experimentally (see below and Table 2). Therefore, conversion of fructose-1,6-bisphosphate into fructose-6-phosphate must be achieved by an alternative and previously uncharacterized route.

Strikingly, the pentose phosphate pathway also lacks two of its enzymes: the 6-phosphogluconate dehydrogenase (EC 1.1.1.44) and the transaldolase (EC 2.2.1.2). The absence of 6-phosphogluconate dehydrogenase was already reported in other prokaryotes (43), and its product ribulose 5-phosphate could be rescued through the Entner-Doudoroff pathway. However, the absence of the transaldolase is more surprising, since the enzyme is well conserved in prokaryotes (43).

**Experimental validation of carbon source usage.** In order to validate the metabolic model predictions, we experimentally assessed the growth of *X. fastidiosa* on 190 potential carbon (C) sources using Biolog phenotype microarrays (PMs); 8 can provide relatively fast growth and 18 can provide relatively slow growth at the scale of *X. fastidiosa* fastidious growth (Table 2). Growth is achievable on diverse xylem fluid components, including five amino acids, citric acid, and D-fructose, even though, surprisingly, eight compounds leading to fast growth are not major components. The number of carbon sources identified using Biolog PMs was 25, which is considerably lower than the number of C sources identified in *R. solanacearum* and *E. coli* (respectively, 36 and 86) (22, 44), confirming our prediction of reduced metabolic diversity. We also used FBA to predict the relative growth rate of *X. fastidiosa* on several substrates. Relative *in silico* growth rates were determined by computing the ratio between the growth rate and the highest growth rate obtained (on glycerol). The discrepancies observed between experimental and simulation results (growth *in vivo* but not predicted *in silico*) were used to correct the network, by adding the missing metabolic reactions for 12 substrates (as specified in the final metabolic network in Data Set S1). The refined metabolic network predicted growth on 24 of the 26 experimentally verified substrates; growth was observed with Biolog PMs on D,L-$\alpha$-glycerol phosphate and 3-*O*-$\beta$-D-galactopyranosyl-D-arabinose but not modeled because there are no assimilation reactions for these substrates in available metabolic models. Based on the FBA results, we classified the different carbon sources (Table 2 and Data Set S3) by defining relatively fast growth *in silico* as $>$75% of the maximal growth rate. Sixteen of



Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.

**TABLE 2** *In silico* and *in vivo* growth assays on different carbon sources using FBA and Biolog PMs

| Substrate | Growth assessment | | Presence of compound in xylem[c] | |
|---|---|---|---|---|
| | *In silico*[a] | Exptl[b] | Grapevine | Olive tree |
| Amino acids | | | | |
| L-Proline | + | ++ | Minor | Minor |
| L-Glutamate | + | + | ND | Minor |
| L-Aspartate | + | + | ND | Minor |
| L-Alanine | + | + | Minor | Minor |
| L-Glutamine | + | + | Major | Major |
| L-Arginine | + | + | Minor | Minor |
| L-Histidine | + | + | Minor | ND |
| L-Ornithine | + | + | ND | ND |
| γ-Aminobutyric acid | + | + | ND | Minor |
| Organic acids | | | | |
| Acetic acid | + | + | ND | Minor |
| Citric acid | + | + | Major | ND |
| Pyruvic acid | + | ++ | ND | ND |
| Monosaccharides | | | | |
| D-Glucose | ++ | ++ | Minor | Major |
| D-Galactose | ++ | ++ | ND | ND |
| D-Trehalose | ++ | ++ | ND | ND |
| D-Fructose | ++ | + | Minor | Major |
| D-Ribose | + | + | ND | ND |
| D-Mannose | ++ | + | ND | ND |
| D-Xylose | + | + | ND | ND |
| Others | | | | |
| Chitin | ++ | ++/+* | ND | ND |
| Dextrin | ++ | + | ND | ND |
| *myo*-Inositol | + | ++ | ND | Minor |
| Glycerol | ++ | ++ | ND | ND |
| D-Glucosamine | ++ | + | ND | ND |
| 3-*O*-β-D-Galactopyranosyl-D-arabinose | Unknown | + | ND | ND |
| DL-α-Glycerol phosphate | Unknown | + | ND | ND |

[a]Relative *in silico* growth rates were determined by computing the ratio between the growth rate and the highest growth rate obtained (on glycerol). If the relative *in silico* growth rate was >75%, growth was assessed as relatively fast (++), and if the relative *in silico* growth rate was ≤75%, growth was assessed as relatively slow (+).

[b]Growth assessment was made using Biolog PM respiration rate: ++ for relatively fast growth, referring to a respiration rate above 0.125 $h^{-1}$, and + for relatively slow growth, referring to a respiration rate under 0.125 $h^{-1}$ (see Materials and Methods and Data Set S3), except for chitin (*), for which growth was previously determined experimentally (45). Growth was observed with Biolog PMs on D,L-α-glycerol phosphate and 3-*O*-β-D-galactopyranosyl-D-arabinose but not modeled, because there are no assimilation reactions for these substrates in available metabolic models. Substrates with indeterminate growth or no growth are listed in Data Set S3.

[c]Xylem fluid composition from grapevine and olive tree was extracted from the work of Andersen and Brodbeck (37) and Montes Borrego et al. (39). Major, concentration ≥ 500 μM; minor, concentration < 500 μM; ND, not detected.

23 (70%) of the predictions were in accordance with the experimental measurements. The highest predicted growth rate was observed with glycerol, and this was confirmed experimentally.

We also tested *in silico* the growth on chitin, a polymer composing the insect foregut wall, not present on Biolog PM plates. A recent study showed the necessity of a chitinase (ChiA) for both insect and plant colonization and suggested that chitin is the favored carbon source in the insect environment (45). Our results (Table 2) support the hypothesis that chitin is a favored substrate since it provides some of the fastest growth (relatively speaking for *X. fastidiosa*).

**Modeling-based evidence for a fragile metabolism.** To first estimate the level of functional redundancy in the *X. fastidiosa* network, we determined the number of enzymes associated with each reaction in comparison with other reference metabolic models (Fig. 1D and Data Set S2). Eighty-seven percent of the metabolic reactions in *X. fastidiosa* are carried out by a unique enzyme, while no other organism exceeds 80% except the also-fastidious *Bordetella pertussis* (11). Alternative enzymes are indicative of functional redundancies in the network and are generally associated with environmen-



Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.

**TABLE 3** Simulation analyses of the metabolic network

| Organism | Proportion (%) of: | | | |
|---|---|---|---|---|
| | Essential metabolic genes (gene essentiality[a]) in: | | Reactions (flux variability[b]) | |
| | Glucose medium | Protein-rich environment | Varying | Bidirectional |
| *E. coli* | 15 | 12 | 60 | 4 |
| *R. solanacearum* | 15 | 13 | 48 | 3 |
| *B. pertussis* | No growth | 40 | | |
| *X. fastidiosa* | 54 | 51 | 32 | 1 |

[a]Proportion of essential metabolic genes according to metabolic modeling, in glucose medium and a protein-rich environment. As *B. pertussis* cannot grow on glucose, the gene essentiality value is given only for a simulated protein-rich growth environment (12). Data for *B. pertussis* and *E. coli* (growth on glucose) were extracted from the work of Fyson et al. (12) and Orth et al. (67). Values for *R. solanacearum* and *X. fastidiosa* in both environments and for *E. coli* in a protein-rich environment were computed for this study. The detailed results are available in Data Set S4.

[b]Assessment of flux variability. The proportion of varying reactions represents the number of reactions with non-null flux variation divided by the total number of metabolic reactions in the network. Bidirectional reactions are reversible reactions that can carry out both positive and negative flux to sustain an optimal growth. The proportion was equally determined by calculating the ratio of these reactions divided by the total number of metabolic reactions in the network. The detailed results are available in Data Set S2.

tal plasticity. This lack in *X. fastidiosa* illustrated further the low flexibility of its metabolism. We then performed a gene essentiality analysis on the network. The results for *X. fastidiosa* and other reference organisms are presented in Table 3 (detailed information in Data Set S4). Strikingly, the predicted proportion of essential genes was 54% on glucose and 51% in a protein-rich environment, which is considerably higher than the other tested bacteria. This outstandingly high proportion of essential genes in *X. fastidiosa* illustrates the fragility of the metabolic network against perturbations.

Fragility and low flexibility of *X. fastidiosa* metabolism were confirmed using flux variability analysis (FVA). FVA is an *in silico* analysis to estimate if alternative flux values, due to an internal or external perturbation, will still sustain an optimal metabolic behavior. We performed FVA using L-glutamine as carbon source and growth as the objective function. Similar analyses were also conducted with *E. coli* and *R. solanacearum* (Table 3 and Data Set S2). As expected from the previous results, the proportion of varying reactions (i.e., reactions with alternative fluxes able to sustain the biological objective) was significantly lower (32%) in *X. fastidiosa* than in *R. solanacearum* (48%) and *E. coli* (60%). Thus, only a limited proportion of the metabolic reactions can still sustain optimal growth if their fluxes vary. This particular trait confirmed a lack of flexibility in *X. fastidiosa* metabolism.

**A higher cost for virulence can compromise *X. fastidiosa* growth.** We finally assessed the efficiency of the network structure for a specific task by determining metabolic yields on proliferation and on virulence factors. Metabolic yields were defined as the maximal proportion of carbon which could be invested in the biosynthesis of a specific macromolecule (virulence factor) or biomass. The obtained efficiency data are presented in Table 4 and compared to another plant pathogen, *R. solanacearum*. This other plant pathogen also has a network including virulence proteins

**TABLE 4** Structural efficiency analysis of *X. fastidiosa* and *R. solanacearum* metabolic models[a]

| Biological objective | Yield (%) for species: | |
|---|---|---|
| | *Xylella fastidiosa* | *Ralstonia solanacearum* |
| Biomass | 54 | 58 |
| Protein | 61 | 64 |
| EPS | 58 | 72 |

[a]Metabolic yields were compared for different biological objectives: (i) production of biomass, (ii) production of virulence proteins, and (iii) production of EPS. The mathematical details on yield determination are given in Data Set S2.



Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.

and EPS secretion in the model, in contrast to other metabolic models including *E. coli*. The comparison was thus limited to these two plant pathogens.

Similar efficiency is observed for biomass production and extracellular protein secretion (Table 4; Data Set S2). However, the structure of the *X. fastidiosa* metabolic network is less efficient for production of EPS than *R. solanacearum* (Table 4 and Data Set S2), with an increase of losses as $CO_2$ (+16%). In other words, biosynthesis of this virulence factor has a higher carbon cost in *X. fastidiosa* and this probably impairs reallocation of carbon fluxes to sustain efficient biomass production.

## DISCUSSION

Fastidious growth is a paradoxical property of some prokaryote pathogens, and we sought to understand the emergence of this phenotype through a metabolic network study in a particularly fastidious organism, the plant pathogen *X. fastidiosa*. This bacterium appears to possess a minimal but fully functional metabolic network, and this was confirmed experimentally. This is in contrast with several other pathogen or symbiont species which have also undergone reductive evolution with the loss of essential metabolic functions (12, 46, 47).

**The reduction of the *X. fastidiosa* metabolic network probably results from its adaptation to a limited number of environments.** The reduction of the metabolic network does not create any auxotrophy for *X. fastidiosa* growth but tends to restrict the pathogen to a limited diversity of metabolic behaviors. *X. fastidiosa* is specifically restricted to two environments (plant vascular vessels and the foregut of xylem-eating insects) (15). These environments have common features as they are both constituted of xylem sap and carbonated polymers (plant cell wall and chitin), and it has been shown that the chitin-degrading enzyme is essential for colonization both in the insect and *in planta* (45). Thereby, *X. fastidiosa* appears to be restricted to a homeostatic environment, i.e., overall constant in composition and not prone to external perturbations. This lifestyle restricted to specific environments probably explains the limitation of metabolic behaviors (particularly the catabolic capacities), which was inferred from our study.

One could argue that fastidiousness of *Xylella* may be due to an environment poor in nutrients. However, *Xylella* has a fastidious growth even in nutrient-rich artificial media (e.g., buffered charcoal yeast extract [BCYE] medium [21]). In addition, other vascular plant pathogens such as *R. solanacearum* have a high proliferation rate in the same environment: 3 days after injection in stems, *R. solanacearum* can reach up to $10^9$ CFU per gram of fresh weight (48). Thereby, slow growth cannot be explained by a nutritional limitation *in planta* but seems to be due to intrinsic specificities of the pathogen.

**The genome-scale metabolic analysis reveals a lack of metabolic robustness.** Robustness is a key feature of biological systems, which allows them to maintain their function(s) when subjected to environmental or internal perturbation (7, 49). Genome-scale analysis of *X. fastidiosa* metabolism was carried out to examine if robustness could have shaped the evolution of its ancestor. By studying gene essentiality and enzyme distribution, we showed that functional redundancy, one of the main sources of robustness, is substantially missing in the *X. fastidiosa* metabolic network. This lack of robustness reduces the possibilities for *X. fastidiosa* to protect itself against perturbations both internal as mutations and external as environmental perturbations. FVA showed that the network lacks flexibility, and as mentioned above, this lack of robustness is also characterized by restricted trophic capabilities. This global fragility certainly makes *X. fastidiosa* very sensitive to deleterious mutations. The high level of homologous recombination observed in the species (50, 51) could serve as a rescue mechanism in this context.

**Metabolic peculiarities responsible for a slow-growth phenotype.** Several elements were raised by our study (Fig. 2).

**(i) *X. fastidiosa* metabolism lacks certain genes/pathways which are not responsible for auxotrophy but are nevertheless assumed to ensure rapid growth.**



Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.

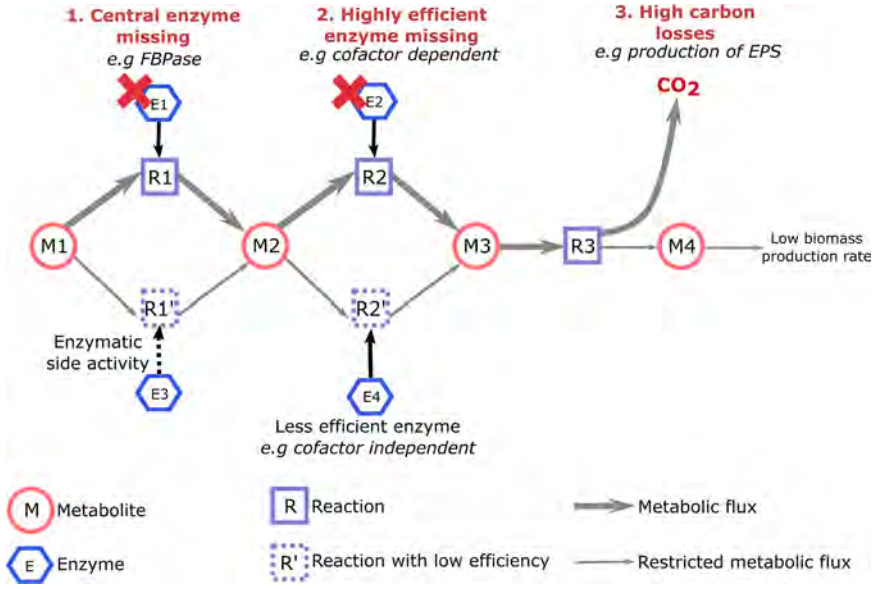



**FIG 2** Metabolic properties contributing to fastidious growth. The three main metabolic properties unraveled by our study and probably contributing to fastidious growth are listed in red at the top of the figure. These three properties limit the metabolic fluxes. For 1 and 2, as the highly efficient enzyme is not available, a less efficient enzyme performs the reaction, which decreases the metabolic fluxes. For 3, the production of virulence factor strongly enhances carbon losses as $CO_2$, which will limit the ability of the network to efficiently convert substrate into biomass. These flux limitations are predicted to strongly reduce growth rate and be responsible, possibly with additional inputs such as regulation, for fastidious growth.

Some peripheral biosynthetic pathways appear to be missing, such as the ones for cobalamin and molybdopterin. It limits the bacterium to cobalamin- and molybdopterin-independent enzymes, usually possessing a lower enzymatic activity (52, 53). Similar observations were made on the respiratory metabolism, which appears to be restricted to a combination of poorly efficient cytochrome and cytochrome oxidase. A central enzyme, succinate dehydrogenase, also involved in respiratory metabolism, surprisingly lacks a membrane subunit, which could also affect the efficiency of central metabolic fluxes.

Furthermore, *X. fastidiosa* lacks two key enzymes: fructose-1,6-bisphosphatase (FBPase) for gluconeogenesis and transaldolase in the pentose phosphate pathway. We found that these lacks were not deleterious, suggesting that for FBPase the reaction was processed by a nonidentified route. The emerging concept of underground metabolism provides a reasonable hypothesis: the missing reaction could be achieved through the side activity of another enzyme, which leads to a reduced enzymatic activity (54).

A reduction in enzymatic activity in a reaction central for growth is assumed to strongly reduce the proliferation rate. To illustrate this phenomenon, we estimated that the experimentally observed *X. fastidiosa* growth rate (0.1608 $day^{-1}$) (19) was recovered *in silico* with a 3.37-fold reduction of the optimal FBPase flux (see Data Set S2 in the supplemental material). This reduction converted the generation time from 1.45 h (typical optimal bacterial growth) to 103 h (fastidious growth), so the growth rate is reduced by 71-fold, while the FBPase flux is reduced by only 3.37-fold. This *in silico* reduction seems reasonable, as higher flux reductions were experimentally observed (see, for example, a 200-fold reduction for an enzymatic activity in reference 54).

In photosynthetic organisms, a double role of sedoheptulose-1,7-bisphosphatase (SBPase) and FBPase activity in the same enzyme was reported, reinforcing the view that FBPase activity could be achieved by an alternative enzyme, with a lower efficiency (55–57). No SBPase exists in nonphotosynthetic organisms, but we hypothesize that another enzyme, *myo*-inositol 1-phosphatase (IMPase), could have a bisphosphatase



Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.

activity on fructose-1,6-bisphosphate. This IMPase enzyme (EC 3.1.3.25) was identified in the *X. fastidiosa* metabolic network. Although *myo*-inositol is not a hexose, it is a 6-carbon cyclic molecule with hydroxyl groups, which makes its chemical structure very similar to fructose. An enzyme with dual IMPase/FBPase activity was found in archaea (58) and also in the slow-growing pathogen *Mycobacterium tuberculosis* (59). It was shown that the archaeal IMPase had a catalytic activity on fructose-1,6-bisphosphate 3-fold lower than a classical FBPase (21 $s^{-1}$ versus 7 $s^{-1}$) (58). Remarkably, the difference of activity measured in that study (3-fold) is similar to the loss of enzymatic efficiency we predicted to match fastidious growth (reduction of flux: 3.37-fold).

The loss of FBPase is observed among all *X. fastidiosa* strains sequenced to date. This observation suggests that this enzyme was already lost in the *Xylella* ancestor, whereas this enzyme (EC 3.1.3.11) is broadly conserved in living organisms (60). Looking more closely at the *Xanthomonadaceae* phylum to which *Xylella* belongs, it appears that FBPase is conserved in all *Xanthomonas* species except *Xanthomonas albilineans*. Interestingly, *X. albilineans* is also a xylem-limited and slow-growing plant pathogen that has undergone reductive evolution, although less extensively than *X. fastidiosa* (61). Future investigations should address whether the loss of FBPase is linked to a specialization in terms of environment or lifestyle.

**(ii) *X. fastidiosa* metabolism is structurally inefficient.** The decreased metabolic efficiency (in terms of yield) was particularly obvious in comparison with another vascular plant pathogen, *R. solanacearum*, when we determined the cost for the production and excretion of EPS, a major virulence factor of these bacteria (24, 62). In *R. solanacearum*, it was shown that EPS production has a significant metabolic cost that can impact biomass production (loss of approximately 25% of the growth rate due to EPS production) (22). Our simulations (Table 4) predict that the cost for EPS production is an even higher burden for *X. fastidiosa*.

Fastidious growth probably results from a combination of two factors. First is the lack of reactions catalyzed by efficient enzymes, cobalamin/molybdopterin-dependent enzymes, complete succinate dehydrogenase complex, efficient respiratory chain, and, strikingly, the absence of a central enzyme in gluconeogenesis. Second is a global inefficiency in producing virulence factors. We cannot formally exclude that enzymatic efficiency could be globally lower in *Xylella* species than in other bacteria. Surprisingly, the predictions from the efficiency analysis indicate that the potential for biomass production by *X. fastidiosa* is virtually similar to that of *R. solanacearum* (Table 4). This observation suggests that constraints imposed on metabolic efficiency both by virulence factor production and by regulation patterns specifically limit the growth of *X. fastidiosa*. Supporting this hypothesis, it was reported that an *X. fastidiosa* mutant deficient for the production of a secreted and virulence-related protein had a growth rate significantly increased in comparison to the wild-type strain (63). In such a scenario, the fastidious behavior could be viewed as a developmental strategy of the pathogen to remain at a low population level inside the host and avoid detection by its immune system. This hypothesis is in agreement with the view of *Xylella* as a "self-limiting" organism, which has many features of a plant commensal but accidentally provokes epidemics due to an exaggerated and late plant response (64), probably linked to specific environmental conditions.

**Common metabolic peculiarities in fastidious pathogens.** It is interesting to put these findings in relation to other fastidious organisms. The pathogen *B. pertussis*, similarly to *X. fastidiosa*, experienced a strong genome reduction, which implied a metabolic network reduction, leading to a number of reactions and metabolites close to *X. fastidiosa* (Table 1). The fastidious growth of *B. pertussis* could thus also be related to a lack of functional redundancy favoring less efficient pathways, and it is also conceivable that this reduction could affect the network efficiency and increase the cost of virulence factor production. Furthermore, another core metabolic enzyme is missing in *B. pertussis*, making its glycolysis unachievable (12). In contrast to *X. fastidiosa*, this absence led to an auxotrophy for amino acids, but one can consider that this



Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.

absence also constrains the metabolic fluxes and provokes fastidious growth. Another fastidious pathogen, *M. tuberculosis*, has an interesting similarity to *X. fastidiosa*; the lack of standard fructose-1,6-bisphosphatase was also reported in the pathogen, suggesting the use of an enzyme side activity (59) or an atypical fructose-1,6-bisphosphatase enzyme (65). These common metabolic particularities between slow-growing pathogens might indicate that in an environment where fast growth could activate the host immune system, similar evolutionary processes of growth reduction, at different degrees and strengths, might favor the emergence of fastidious phenotypes.

In conclusion, based on genetic and phenotypic data and computational approaches, we were able to unravel several candidate genes responsible for bacterial fastidious growth. A perspective of this study is to validate these candidates by functional genetic approaches to establish their involvement in the fastidious growth phenotype.

## MATERIALS AND METHODS

**Bacterial strain.** Analyses were conducted with the strain CFBP 8418, an *X. fastidiosa* subsp. *multiplex* ST6 (sequence type 6) strain isolated from *Spartium junceum* in Alata (France) (23) and provided by the French Collection for Plant-associated Bacteria (CIRM-CFBP).

**Carbon substrate phenotyping.** Phenotyping was performed using Biolog phenotype microarray plates PM1 and PM2A according to the manufacturer's protocol, except that Tween 40 was replaced by Tween 80. An initial optical density (OD) of 0.18 (600 nm) was used for inoculation. Incubation time was 72.5 h. Three independent replicates were performed.

Growth was assumed proportional to respiration and was assessed by calculating the area under the curve (AUC) and the slope of the logarithm on the exponential phase (respiration rate). Growth was considered effective if AUC was $>AUC_{negative\ control}$ + threshold (see Data Set S3). If the respiration rate was above 0.125 $h^{-1}$, the growth was categorized as fast (++); otherwise, it was categorized as slow (+). If the results were nonreproducible, the growth was assessed to be undetermined. For six substrates (out of 190), for which growth was assessed unambiguously as positive for two out of the three replicates, growth was referred to (+). The complete results are available in Data Set S3.

**Genome-scale metabolic network reconstruction.** The genome-scale metabolic network reconstruction was performed from the genomic sequences of the strain CFBP 8418 (GenBank accession no. LUYA00000000.1) (23). The protocol to generate high-quality genome-scale metabolic network proposed by Thiele and Palsson (66) was followed.

For the automatic reconstruction, the genomic sequences were compared, by priority order, to the five following metabolic models: *Escherichia coli* K-12 MG1655 (iJO1366 [67]), *Ralstonia solanacearum* GMI1000 (iRP1476 [22]), *Pseudomonas aeruginosa* PAO1 (iMO1086 [68]), *Ralstonia eutropha* H16 (RehMBEL1391 [69]), and *Bacillus subtilis* 168 (iYO844 [70]). *X. fastidiosa* sequences were compared by orthology to the sequences from each metabolic model, generating five draft models, thanks to the Autograph method (71). The SAMIR tool (22) was used to reconcile the identifiers between the models. Finally, the five metabolic models were merged into one draft model expressed with BiGG identifiers (72), following the priority order stated above and depending on the orthology quality.

The draft metabolic model was manually curated, evaluating the reactions one by one. The accuracy of each reaction was checked using KEGG (60), BiGG (72), BioCyc (73), MetaCyc (73), and bibliographical searches. Simulations (see below) were performed pathway by pathway to check if the curated pathway was functional. A manual gap-filling step was processed for nonfunctional pathways, looking for evidence of the missing reactions. Each reaction was scored to assess its reliability with a specific score that was more adapted to the network than the confidence score proposed by Thiele and Palsson (66). The meaning of this specific score and the complete manual curation protocol are available in Text S2. The two scores are given for each reaction in the metabolic network (Data Set S1).

As proposed by Thiele and Palsson (66), a biomass objective function was defined. General characterizations of the *X. fastidiosa* genome (29) and specificities of the studied strain (23) were both used. For missing information, data from *Xanthomonas campestris* pv. campestris (74) and *E. coli* (67) were used (Data Set S2). After curation, the tool CarveMe (75) was used to find other potential reactions. The generated reactions were manually curated and added to the network.

For metabolic network comparisons between subspecies, draft metabolic networks of *Xylella fastidiosa* subsp. *pauca* 9a5c (GenBank accession no. GCA_000006725.1) (29) and *X. fastidiosa* subsp. *fastidiosa* Temecula1 (GenBank accession no. GCA_000007245.1) (30) were generated with the same reconstruction process.

**Comparative analysis.** The reaction identifiers were given as input of the comparative analysis between *E. coli*, *R. solanacearum*, and *X. fastidiosa*. The metabolic model considered for *E. coli* was the model iJO1366 (67). For *R. solanacearum*, the model iRP1476 (22) was used, taking into account only the metabolic module. The Venn diagram was generated using the online tool BioVenn (76). The list of reactions in each part of the diagram is available in Data Set S2.

**Computational simulations.** To test the functionality of pathways and of the global network, simulations were performed on the model. To access intracellular fluxes and growth rate, the flux balance analysis (FBA) methodology was used (77). The constraints used to model the system are detailed in Text S3. To access the variability of each flux, flux variability analysis (FVA), a methodology based upon the



Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.

same principles, was then performed (78). A deviation of 1% from the optimality was allowed. *In silico* gene essentiality analysis was performed through multiple FBAs on the network. For each FBA, a metabolic gene of the network is deleted to test if its deletion has an impact on the organism growth. The *in silico* gene essentiality analysis in a protein-rich environment was inspired by a *B. pertussis* publication (12): L-glutamate as the major carbon source and an availability 10 times smaller in number of carbon for each other proteinogenic amino acid. For FVA and *in silico* gene essentiality, protein and EPS constraints were removed.

The structural efficiency of the network was tested in *X. fastidiosa* and, as a comparison, in *R. solanacearum*. To avoid bias, all the constraints were removed, and a generic L-glutamine uptake rate of 1 mmol•$h^{-1}$•g (dry weight [DW])$^{-1}$ was used. Each of the biological processes (biomass and virulence) was studied separately, and the maximization of their synthesis was used as the objective. Each simulation result gave access to the proportion of carbon lost in the form of $CO_2$, with the exported $CO_2$ rate (R_EX_CO2_e). This allowed us to determine the maximal amount of carbon which could be dedicated to a virulence factor or biomass production. Metabolic yields, and consequently cost of virulence factors, were computed as the ratio between maximal production rate and the uptake rate, normalized by the number of carbons.

Text S4 provides a detailed description of mathematical modeling of metabolism used in the study.

Simulations were performed with the open source tool FlexFlux (79) and are reproducible using the guide provided in Text S4. The linear programming solver CPLEX, developed by IBM, was used to solve the system and get solutions. Scripts and command lines are available online on GitHub: https://github.com/lgerlin/xfas-metabolic-model.

**Data availability.** All data generated or analyzed during this study are included in figures or tables or in the supplemental material. In particular, raw data from Biolog phenotype microarray plates (for carbon substrate phenotyping) are available in Data Set S3.

**Model and code availability.** The *X. fastidiosa* subsp. *multiplex* strain CFBP 8418 metabolic model was named Xfm1158. The metabolic model was deposited in BioModels (80) and assigned the identifier MODEL2003100001. Exploration, omics mapping, and basic flux analyses can be performed on the model in MetExplore (81) (https://metexplore.toulouse.inrae.fr/metexplore2/?idBioSource=5822). The main scripts and command lines used for the study are available on GitHub: https://github.com/lgerlin/xfas-metabolic-model.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**TEXT S1**, TXT file, 2.5 MB.

**TEXT S2**, PDF file, 0.4 MB.

**TEXT S3**, PDF file, 0.3 MB.

**TEXT S4**, PDF file, 0.9 MB.

**TABLE S1**, PDF file, 0.6 MB.

**TABLE S2**, PDF file, 0.6 MB.

**DATA SET S1**, XLSX file, 0.2 MB.

**DATA SET S2**, XLSX file, 0.2 MB.

**DATA SET S3**, XLSX file, 0.9 MB.

**DATA SET S4**, XLSX file, 0.3 MB.

## ACKNOWLEDGMENTS

L.G. was funded by a grant from the French Ministry of National Education and Research. The study was funded by the European Union's Horizon 2020 research and innovation program through grant agreement XF-ACTORS (727987) and the "Laboratoire d'Excellence (LABEX)" TULIP (ANR-10-LABX-41). The views expressed are purely those of the writers, and the EU funding agency is not responsible for any use that may be made of the information it contains. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

We acknowledge CIRM-CFBP (Beaucouzé, INRAE, France; http://www6.inrae.fr/cirm_eng/CFBP-Plant-Associated-Bacteria) for strain preservation and supply and access to the Omnilog station.

The authors declare that they have no competing interests.

Conceptualization: L.C., S.G., C.B. Metabolic reconstruction and curation: L.G., L.C. Simulations and analysis: L.G., C.B. Experiments: G.T., S.C. Funding acquisition: M.-A.J., S.G. Supervision: C.B. Writing—original draft: L.G. Writing—review & editing: L.G., C.B., L.C., G.T., S.C., M.-A.J., S.G.



Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.

## REFERENCES


1. Ibarra RU, Edwards JS, Palsson BO. 2002. *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. Nature 420:186–123. https://doi.org/10.1038/nature01149.
2. Edwards JS, Palsson BO. 1999. Systems properties of the *Haemophilus influenzae* Rd metabolic genotype. J Biol Chem 274:17410–17416. https://doi.org/10.1074/jbc.274.25.17410.
3. Lewis NE, Hixson KK, Conrad TM, Lerman JA, Charusanti P, Polpitiya AD, Adkins JN, Schramm G, Purvine SO, Lopez-Ferrer D, Weitz KK, Eils R, König R, Smith RD, Palsson B. 2010. Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. Mol Syst Biol 6:390. https://doi.org/10.1038/msb.2010.47.
4. Guidot A, Jiang W, Ferdy JB, Thébaud C, Barberis P, Gouzy J, Genin S. 2014. Multihost experimental evolution of the pathogen *Ralstonia solanacearum* unveils genes involved in adaptation to plants. Mol Biol Evol 31:2913–2928. https://doi.org/10.1093/molbev/msu229.
5. Perrier A, Peyraud R, Rengel D, Barlet X, Lucasson E, Gouzy J, Peeters N, Genin S, Guidot A. 2016. Enhanced in planta fitness through adaptive mutations in EfpR, a dual regulator of virulence and metabolic functions in the plant pathogen *Ralstonia solanacearum*. PLoS Pathog 12: e1006044-23. https://doi.org/10.1371/journal.ppat.1006044.
6. Deslandes L, Genin S. 2014. Opening the *Ralstonia solanacearum* type III effector tool box: insights into host cell subversion mechanisms. Curr Opin Plant Biol 20:110–117. https://doi.org/10.1016/j.pbi.2014.05.002.
7. Peyraud R, Cottret L, Marmiesse L, Genin S. 2018. Control of primary metabolism by a virulence regulatory network promotes robustness in a plant pathogen. Nat Commun 9:418. https://doi.org/10.1038/s41467-017-02660-4.
8. Ribet D, Cossart P. 2015. How bacterial pathogens colonize their hosts and invade deeper tissues. Microbes Infect 17:173–183. https://doi.org/10.1016/j.micinf.2015.01.004.
9. Doern GV. 2000. Detection of selected fastidious bacteria. Clin Infect Dis 30:166–173. https://doi.org/10.1086/313586.
10. Thalen M, van den IJssel J, Jiskoot W, Zomer B, Roholl P, de Gooijer C, Beuvery C, Tramper J. 1999. Rational medium design for *Bordetella pertussis*: basic metabolism. J Biotechnol 75:147–159. https://doi.org/10.1016/s0168-1656(99)00155-8.
11. Kilgore PE, Salim AM, Zervos MJ, Schmitt H. 2016. Pertussis: microbiology, disease, treatment, and prevention. Clin Microbiol Rev 29:449–486. https://doi.org/10.1128/CMR.00083-15.
12. Fyson N, King J, Belcher T, Preston A, Colijn C. 2017. A curated genome-scale metabolic model of *Bordetella pertussis* metabolism. PLoS Comput Biol 13:e1005639-17. https://doi.org/10.1371/journal.pcbi.1005639.
13. James BW, Williams A, Marsh PD. 2000. The physiology and pathogenicity of *Mycobacterium tuberculosis* grown under controlled conditions in a defined medium. J Appl Microbiol 88:669–677. https://doi.org/10.1046/j.1365-2672.2000.01020.x.
14. Edelstein PH. 2000. Detection of selected fastidious bacteria. Clin Infect Dis 31:846. https://doi.org/10.1086/314002.
15. Sicard A, Zeilinger AR, Vanhove M, Schartel TE, Beal DJ, Daugherty MP, Almeida R. 2018. *Xylella fastidiosa*: insights into an emerging plant pathogen. Annu Rev Phytopathol 56:181–202. https://doi.org/10.1146/annurev-phyto-080417-045849.
16. Rapicavoli J, Ingel B, Blanco-Ulate B, Cantu D, Roper C. 2018. *Xylella fastidiosa*: an examination of a re-emerging plant pathogen. Mol Plant Pathol 19:786–800. https://doi.org/10.1111/mpp.12585.
17. European and Mediterranean Plant Protection Organization. 2016. PM 7/24 (2) *Xylella fastidiosa*. EPPO Bull 46:463–500. https://doi.org/10.1111/epp.12327.
18. Lemos E, Alves LMC, Campanharo JC. 2003. Genomics-based design of defined growth media for the plant pathogen *Xylella fastidiosa*. FEMS Microbiol Lett 219:39–45. https://doi.org/10.1016/S0378-1097(02)01189-8.
19. Leite B, Andersen PC, Ishida ML. 2004. Colony aggregation and biofilm formation in xylem chemistry-based media for *Xylella fastidiosa*. FEMS Microbiol Lett 230:283–290. https://doi.org/10.1016/S0378-1097(03)00917-0.
20. Almeida RPP, Mann R, Purcell AH. 2004. *Xylella fastidiosa* cultivation on a minimal solid defined medium. Curr Microbiol 48:368–372. https://doi.org/10.1007/s00284-003-4219-x.
21. Shriner AD, Andersen PC. 2014. Effect of oxygen on the growth and biofilm formation of *Xylella fastidiosa* in liquid media. Curr Microbiol 69:866–873. https://doi.org/10.1007/s00284-014-0660-2.
22. Peyraud R, Cottret L, Marmiesse L, Gouzy J, Genin S. 2016. A resource allocation trade-off between virulence and proliferation drives metabolic versatility in the plant pathogen *Ralstonia solanacearum*. PLoS Pathog 12:e1005939. https://doi.org/10.1371/journal.ppat.1005939.
23. Denancé N, Legendre B, Briand M, Olivier V, de Boisseson C, Poliakoff F, Jacques MA. 2017. Several subspecies and sequence types are associated with the emergence of *Xylella fastidiosa* in natural settings in France. Plant Pathol 66:1054–1064. https://doi.org/10.1111/ppa.12695.
24. Da Silva FR, Vettore AL, Kemper EL, Leite A, Arruda P. 2001. Fastidian gum: the *Xylella fastidiosa* exopolysaccharide possibly involved in bacterial pathogenicity. FEMS Microbiol Lett 203:165–171. https://doi.org/10.1111/j.1574-6968.2001.tb10836.x.
25. Clifford JC, Rapicavoli JN, Roper MC. 2013. A rhamnose-rich O-antigen mediates adhesion, virulence, and host colonization for the xylem-limited phytopathogen *Xylella fastidiosa*. Mol Plant Microbe Interact 26:676–685. https://doi.org/10.1094/MPMI-12-12-0283-R.
26. Rapicavoli JN, Blanco-Ulate B, Muszyński A, Figueroa-Balderas R, Morales-Cruz A, Azadi P, Dobruchowska JM, Castro C, Cantu D, Roper MC. 2018. Lipopolysaccharide O-antigen delays plant innate immune recognition of *Xylella fastidiosa*. Nat Commun 9:390. https://doi.org/10.1038/s41467-018-02861-5.
27. Nascimento R, Gouran H, Chakraborty S, Gillespie HW, Almeida-Souza HO, Tu A, Rao BJ, Feldstein PA, Bruening G, Goulart LR, Dandekar AM. 2016. The type II secreted lipase/esterase LesA is a key virulence factor required for *Xylella fastidiosa* pathogenesis in grapevines. Sci Rep 6:18598. https://doi.org/10.1038/srep18598.
28. Bucci EM. 2018. *Xylella fastidiosa*, a new plant pathogen that threatens global farming: ecology, molecular biology, search for remedies. Biochem Biophys Res Commun 502:173–182. https://doi.org/10.1016/j.bbrc.2018.05.073.
29. Simpson AJG, Reinach FC, Arruda P, Abreu FA, Acencio M, Alvarenga R, Alves LMC, Araya JE, Baia GS, Baptista CS, Barros MH, Bonaccorsi ED, Bordin S, Bové JM, Briones MRS, Bueno MRP, Camargo AA, Camargo LEA, Carraro DM, Carrer H, Colauto NB, Colombo C, Costa FF, Costa MCR, Costa-Neto CM, Coutinho LL, Cristofani M, Dias-Neto E, Docena C, El-Dorry H, Facincani AP, Ferreira AJS, Ferreira VCA, Ferro JA, Fraga JS, França SC, Franco MC, Frohme M, Furlan LR, Garnier M, Goldman GH, Goldman MHS, Gomes SL, Gruber A, Ho PL, Hoheisel JD, Junqueira ML, Kemper EL, Kitajima JP, Krieger JE, Kuramae EE, Laigret F, Lambais MR, Leite LCC, Lemos E, Lemos MVF, Lopes SA, Lopes CR, Machado JA, Machado MA, Madeira A, Madeira HMF, Marino CL, Marques MV, Martins EAL, Martins EMF, Matsukuma AY, Menck CFM, Miracca EC, Miyaki CY, Monteiro-Vitorello CB, Moon DH, Nagai MA, Nascimento A, Netto LES, Nhani A, Nobrega FG, Nunes LR, Oliveira MA, de Oliveira MC, de Oliveira RC, Palmieri DA, Paris A, Peixoto BR, Pereira GAG, Pereira HA, Pesquero JB, Quaggio RB, Roberto PG, Rodrigues V, de M, Rosa AJ, de Rosa VE, de Sá RG, Santelli RV, Sawasaki HE, da Silva ACR, da Silva AM, da Silva FR, Silva WA, da Silveira JF, Silvestri MLZ, Siqueira WJ, de Souza AA, de Souza AP, Terenzi MF, Truffi D, Tsai SM, Tsuhako MH, Vallada H, Van Sluys MA, Verjovski-Almeida S, Vettore AL, Zago MA, Zatz M, Meidanis J, Setubal JC. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature 406:151–157. https://doi.org/10.1038/35018003.
30. Van Sluys MA, de Oliveira MC, Monteiro-Vitorello CB, Miyaki CY, Furlan LR, Camargo LEA, da Silva ACR, Moon DH, Takita MA, Lemos EGM, Machado MA, Ferro MIT, da Silva FR, Goldman MHS, Goldman GH, Lemos MVF, El-Dorry H, Tsai SM, Carrer H, Carraro DM, de Oliveira RC, Nunes LR, Siqueira WJ, Coutinho LL, Kimura ET, Ferro ES, Harakava R, Kuramae EE, Marino CL, Giglioti E, Abreu IL, Alves LMC, do Amaral AM, Baia GS, Blanco SR, Brito MS, Cannavan FS, Celestino AV, da Cunha AF, Fenille RC, Ferro JA, Formighieri EF, Kishi LT, Leoni SG, Oliveira AR, Rosa VE, Sassaki FT, Sena JAD, de Souza AA, Truffi D, Tsukumo F, Yanai GM, Zaros LG, Civerolo EL, Simpson AJG, Almeida NF, Setubal JC, Kitajima JP. 2003. Comparative analyses of the complete genome sequences of Pierce's disease and citrus variegated chlorosis strains of *Xylella fastidiosa*. J Bacteriol 185:1018–1026. https://doi.org/10.1128/jb.185.3.1018-1026.2003.
31. O'Reilly C, Turner PD. 2003. The nitrilase family of CN hydrolysing enzymes—a comparative study. J Appl Microbiol 95:1161–1174. https://doi.org/10.1046/j.1365-2672.2003.02123.x.
32. Camejo D, Guzmán-Cedeño Á, Moreno A. 2016. Reactive oxygen species, essential molecules, during plant-pathogen interactions. Plant Physiol Biochem 103:10–23. https://doi.org/10.1016/j.plaphy.2016.02.035.
33. Howden AJM, Preston GM. 2009. Nitrilase enzymes and their role in




Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.

plant-microbe interactions. Microb Biotechnol 2:441–451. https://doi.org/10.1111/j.1751-7915.2009.00111.x.
34. Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS. 2003. Comparative genomics of the vitamin $B_{12}$ metabolism and regulation in prokaryotes. J Biol Chem 278:41148–41159. https://doi.org/10.1074/jbc.M305837200.
35. Bhattacharyya A, Stilwagen S, Reznik G, Feil H, Feil WS, Anderson I, Bernal A, D'Souza M, Ivanova N, Kapatral V, Larsen N, Los T, Lykidis A, Selkov E, Walunas TL, Purcell A, Edwards RA, Hawkins T, Haselkorn R, Overbeek R, Kyrpides NC, Predki PF. 2002. Draft sequencing and comparative genomics of *Xylella fastidiosa* strains reveal novel biological insights. Genome Res 12:1556–1563. https://doi.org/10.1101/gr.370702.
36. Hederstedt L, Heden LO. 1989. New properties of *Bacillus subtilis* succinate dehydrogenase altered at the active site. The apparent active site thiol of succinate oxidoreductases is dispensable for succinate oxidation. Biochem J 260:491–497. https://doi.org/10.1042/bj2600491.
37. Andersen PC, Brodbeck BV. 1989. Diurnal and temporal changes in the chemical profile of xylem exudate from *Vitis rotundifolia*. Physiol Plant 75:63–70. https://doi.org/10.1111/j.1399-3054.1989.tb02064.x.
38. Andersen PC, Brodbeck BV, Oden S, Shriner A, Leite B. 2007. Influence of xylem fluid chemistry on planktonic growth, biofilm formation and aggregation of *Xylella fastidiosa*. FEMS Microbiol Lett 274:210–217. https://doi.org/10.1111/j.1574-6968.2007.00827.x.
39. Montes Borrego M, Jiménez-Díaz RM, Trapero Casas JL, Navas Cortés JA, Haro C, Rivas JC, de la Fuente L, Landa BB. 2017. Metabolomic characterization of xylem sap of different olive cultivars growing in Spain. Eur Conf Xylella fastidiosa, poster 2.6.
40. Zuluaga AP, Puigvert M, Valls M. 2013. Novel plant inputs influencing *Ralstonia solanacearum* during infection. Front Microbiol 4:349. https://doi.org/10.3389/fmicb.2013.00349.
41. Sedivy JM, Fraenkel DG. 1985. Fructose bisphosphatase of *Saccharomyces cerevisiae*. Cloning, disruption and regulation of the FBP1 structural gene. J Mol Biol 186:307–319. https://doi.org/10.1016/0022-2836(85)90107-X.
42. Rittmann D, Schaffer S, Wendisch VF, Sahm H. 2003. Fructose-1,6-bisphosphatase from *Corynebacterium glutamicum*: expression and deletion of the fbp gene and biochemical characterization of the enzyme. Arch Microbiol 180:285–292. https://doi.org/10.1007/s00203-003-0588-6.
43. Stincone A, Prigione A, Cramer T, Wamelink MMC, Campbell K, Cheung E, Olin-Sandoval V, Grüning NM, Krüger A, Tauqeer Alam M, Keller MA, Breitenbach M, Brindle KM, Rabinowitz JD, Ralser M. 2015. The return of metabolism: biochemistry and physiology of the pentose phosphate pathway. Biol Rev Camb Philos Soc 90:927–963. https://doi.org/10.1111/brv.12140.
44. Yoon SH, Han MJ, Jeong H, Lee CH, Xia XX, Lee DH, Shim JH, Lee SY, Oh TK, Kim JF. 2012. Comparative multi-omics systems analysis of *Escherichia coli* strains B and K-12. Genome Biol 13:R37. https://doi.org/10.1186/gb-2012-13-5-r37.
45. Labroussaa F, Ionescu M, Zeilinger AR, Lindow SE, Almeida R. 2017. A chitinase is required for *Xylella fastidiosa* colonization of its insect and plant hosts. Microbiology 163:502–509. https://doi.org/10.1099/mic.0.000438.
46. Fenn K, Blaxter M. 2006. *Wolbachia* genomes: revealing the biology of parasitism and mutualism. Trends Parasitol 22:60–65. https://doi.org/10.1016/j.pt.2005.12.012.
47. Thomas GH, Zucker J, Macdonald SJ, Sorokin A, Goryanin I, Douglas AE. 2009. A fragile metabolic network adapted for cooperation in the symbiotic bacterium *Buchnera aphidicola*. BMC Syst Biol 3:24. https://doi.org/10.1186/1752-0509-3-24.
48. Perrier A, Barlet X, Rengel D, Prior P, Poussier S, Genin S, Guidot A. 2019. Spontaneous mutations in a regulatory gene induce phenotypic heterogeneity and adaptation of *Ralstonia solanacearum* to changing environments. Environ Microbiol 21:3140–3152. https://doi.org/10.1111/1462-2920.14717.
49. Félix M-A, Barkoulas M. 2015. Pervasive robustness in biological systems. Nat Rev Genet 16:483–496. https://doi.org/10.1038/nrg3949.
50. Nunney L, Yuan X, Bromley RE, Stouthamer R. 2012. Detecting genetic introgression: high levels of intersubspecific recombination found in *Xylella fastidiosa* in Brazil. Appl Environ Microbiol 78:4702–4714. https://doi.org/10.1128/AEM.01126-12.
51. Coletta-Filho HD, Francisco CS, Lopes JRS, Muller C, Almeida R. 2017. Homologous recombination and *Xylella fastidiosa* host-pathogen associations in South America. Phytopathology 107:305–312. https://doi.org/10.1094/PHYTO-09-16-0321-R.
52. Schomburg I, Chang A, Ebeling C, Gremse M, Heldt C, Huhn G, Schomburg D. 2004. BRENDA, the enzyme database: updates and major new developments. Nucleic Acids Res 32:D431–D433. https://doi.org/10.1093/nar/gkh081.
53. Plener L, Boistard P, González A, Boucher C, Genin S. 2012. Metabolic adaptation of *Ralstonia solanacearum* during plant infection: a methionine biosynthesis case study. PLoS One 7:e36877. https://doi.org/10.1371/journal.pone.0036877.
54. Notebaart RA, Szappanos B, Kintses B, Pál F, Györkei Á, Bogos B, Lázár V, Spohn R, Csörgő B, Wagner A, Ruppin E, Pál C, Papp B. 2014. Network-level architecture and the evolutionary potential of underground metabolism. Proc Natl Acad Sci U S A 111:11762–11767. https://doi.org/10.1073/pnas.1406102111.
55. Ding F, Wang M, Zhang S, Ai X. 2016. Changes in SBPase activity influence photosynthetic capacity, growth, and tolerance to chilling stress in transgenic tomato plants. Sci Rep 6:32741. https://doi.org/10.1038/srep32741.
56. Gerbling K-P, Steup M, Latzko E. 1986. Fructose 1,6-bisphosphatase form B from *Synechococcus leopoliensis* hydrolyzes both fructose and sedoheptulose bisphosphate. Plant Physiol 80:716–720. https://doi.org/10.1104/pp.80.3.716.
57. Jiang Y-H, Wang D-Y, Wen J-F. 2012. The independent prokaryotic origins of eukaryotic fructose-1,6-bisphosphatase and sedoheptulose-1,7-bisphosphatase and the implications of their origins for the evolution of eukaryotic Calvin cycle. BMC Evol Biol 12:208. https://doi.org/10.1186/1471-2148-12-208.
58. Stec B, Yang H, Johnson KA, Chen L, Roberts MF. 2000. MJ0109 is an enzyme that is both an inositol monophosphatase and the "missing" archaeal fructose-1,6-bisphosphatase. Nat Struct Biol 7:1046–1050. https://doi.org/10.1038/80968.
59. Gu X, Chen M, Shen H, Jiang X, Huang Y, Wang H. 2006. Rv2131c gene product: an unconventional enzyme that is both inositol monophosphatase and fructose-1,6-bisphosphatase. Biochem Biophys Res Commun 339:897–904. https://doi.org/10.1016/j.bbrc.2005.11.088.
60. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 45:D353–D361. https://doi.org/10.1093/nar/gkw1092.
61. Pieretti I, Royer M, Barbe V, Carrere S, Koebnik R, Cociancich S, Couloux A, Darrasse A, Gouzy J, Jacques M-A, Lauber E, Manceau C, Mangenot S, Poussier S, Segurens B, Szurek B, Verdier V, Arlat M, Rott P. 2009. The complete genome sequence of *Xanthomonas albilineans* provides new insights into the reductive genome evolution of the xylem-limited Xanthomonadaceae. BMC Genomics 10:616. https://doi.org/10.1186/1471-2164-10-616.
62. Lowe-Power TM, Khokhani D, Allen C. 2018. How *Ralstonia solanacearum* exploits and thrives in the flowing plant xylem environment. Trends Microbiol 26:929–942. https://doi.org/10.1016/j.tim.2018.06.002.
63. Gouran H, Gillespie H, Nascimento R, Chakraborty S, Zaini PA, Jacobson A, Phinney BS, Dolan D, Durbin-Johnson BP, Antonova ES, Lindow SE, Mellema MS, Goulart LR, Dandekar AM. 2016. The secreted protease PrtA controls cell growth, biofilm formation and pathogenicity in *Xylella fastidiosa*. Sci Rep 6:31098. https://doi.org/10.1038/srep31098.
64. Roper C, Castro C, Ingel B. 2019. *Xylella fastidiosa*: bacterial parasitism with hallmarks of commensalism. Curr Opin Plant Biol 50:140–147. https://doi.org/10.1016/j.pbi.2019.05.005.
65. Ganapathy U, Marrero J, Calhoun S, Eoh H, De Carvalho LPS, Rhee K, Ehrt S. 2015. Two enzymes with redundant fructose bisphosphatase activity sustain gluconeogenesis and virulence in *Mycobacterium tuberculosis*. Nat Commun 6:7912. https://doi.org/10.1038/ncomms8912.
66. Thiele I, Palsson BØ. 2010. A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat Protoc 5:93–121. https://doi.org/10.1038/nprot.2009.203.
67. Orth JD, Conrad TM, Na J, Lerman JA, Nam H, Feist AM, Palsson BØ. 2011. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism–2011. Mol Syst Biol 7:535. https://doi.org/10.1038/msb.2011.65.
68. Oberhardt MA, Puchałka J, Martins dos Santos VAP, Papin JA. 2011. Reconciliation of genome-scale metabolic reconstructions for comparative systems analysis. PLoS Comput Biol 7:e1001116. https://doi.org/10.1371/journal.pcbi.1001116.
69. Park J, Kim T, Lee S. 2011. Genome-scale reconstruction and in silico analysis of the *Ralstonia eutropha* H16 for polyhydroxyalkanoate synthesis, lithoautotrophic growth, and 2-methyl citric acid production. BMC Syst Biol 5:101. https://doi.org/10.1186/1752-0509-5-101.
70. Oh Y-K, Palsson BO, Park SM, Schilling CH, Mahadevan R. 2007. Genome-scale reconstruction of metabolic network in *Bacillus subtilis* based on



Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.

high-throughput phenotyping and gene essentiality data. J Biol Chem 282:28791–28799. https://doi.org/10.1074/jbc.M703759200.
71. Notebaart RA, van Enckevort FH, Francke C, Siezen RJ, Teusink B. 2006. Accelerating the reconstruction of genome-scale metabolic networks. BMC Bioinformatics 7:296. https://doi.org/10.1186/1471-2105-7-296.
72. King ZA, Lu J, Dräger A, Miller P, Federowicz S, Lerman JA, Ebrahim A, Palsson BO, Lewis NE. 2016. BiGG Models: a platform for integrating, standardizing and sharing genome-scale models. Nucleic Acids Res 44:D515–D522. https://doi.org/10.1093/nar/gkv1049.
73. Caspi R, Billington R, Ferrer L, Foerster H, Fulcher CA, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Mueller LA, Ong Q, Paley S, Subhraveti P, Weaver DS, Karp PD. 2016. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res 44:D471–D480. https://doi.org/10.1093/nar/gkv1164.
74. Schatschneider S, Persicke M, Watt SA, Hublik G, Pühler A, Niehaus K, Vorhölter FJ. 2013. Establishment, in silico analysis, and experimental verification of a large-scale metabolic network of the xanthan producing *Xanthomonas campestris* pv. campestris strain B100. J Biotechnol 167:123–134. https://doi.org/10.1016/j.jbiotec.2013.01.023.
75. Machado D, Andrejev S, Tramontano M, Patil KR. 2018. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. Nucleic Acids Res 46:7542–7553. https://doi.org/10.1093/nar/gky537.
76. Hulsen T, de Vlieg J, Alkema W. 2008. BioVenn—a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC Genomics 9:488. https://doi.org/10.1186/1471-2164-9-488.
77. Orth JD, Thiele I, Palsson BØ. 2010. What is flux balance analysis? Nat Biotechnol 28:245–248. https://doi.org/10.1038/nbt.1614.
78. Mahadevan R, Schilling CH. 2003. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab Eng 5:264–276. https://doi.org/10.1016/j.ymben.2003.09.002.
79. Marmiesse L, Peyraud R, Cottret L. 2015. FlexFlux: combining metabolic flux and regulatory network analyses. BMC Syst Biol 9:93. https://doi.org/10.1186/s12918-015-0238-z.
80. Glont M, Nguyen TVN, Graesslin M, Hälke R, Ali R, Schramm J, Wimalaratne SM, Kothamachu VB, Rodriguez N, Swat MJ, Eils J, Eils R, Laibe C, Malik-Sheriff RS, Chelliah V, Le Novère N, Hermjakob H. 2018. BioModels: expanding horizons to include more modelling approaches and formats. Nucleic Acids Res 46:D1248–D1253. https://doi.org/10.1093/nar/gkx1023.
81. Cottret L, Frainay C, Chazalviel M, Cabanettes F, Gloaguen Y, Camenen E, Merlet B, Heux S, Portais J-C, Poupin N, Vinson F, Jourdan F. 2018. MetExplore: collaborative edition and exploration of metabolic networks. Nucleic Acids Res 46:W495–W502. https://doi.org/10.1093/nar/gky301.
82. Navarrete F, De La Fuente L. 2014. Response of *Xylella fastidiosa* to zinc: decreased culturability, increased exopolysaccharide production, and formation of resilient biofilms under flow conditions. Appl Environ Microbiol 80:1097–1107. https://doi.org/10.1128/AEM.02998-13.

Downloaded from https://journals.asm.org/journal/msystems on 06 August 2021 by 147.100.179.233.

## 6.3 *Xanthomonas* metabolic capacities, in between *Xylella* and *R. solanacearum*?

After comparing *X. fastidiosa* and *R. solanacearum* metabolism using modeling, generating a metabolic model of *Xcc* appears as the next step. *Xcc* has an in between size of genome and growth rate. As it can be found at the plant surface (epiphytic lifestyle) or on plant debris and seeds, *Xcc* is more versatile in environment colonization than *X. fastidiosa* that encounters "homeostatic" environment (see the first subsection of the article Discussion), but less than *R. solanacearum* that survives in waters and soils (saprophytic lifestyle) (see Chapter 1, section 1.5.3). To this day, only a core metabolic model of *Xcc* exists, that has been developed in the context of xanthan gum production (Schatschneider et al., 2013). To study the lifestyle-metabolism relationship and compare the metabolic capabilities with *X. fastidiosa* and *R. solanacearum*, a genome-scale metabolic model should be reconstructed.

During my second and third years of PhD, I supervised a master student, Antoine Guillemin, to start the reconstruction and curation of *Xcc* genome-scale metabolic model using a similar reconstruction protocol than described in the article. The model is not finished yet, but central pathways have already been intensively curated with a large amount of bibliographical data and experiments on carbon substrate phenotyping. Also, the model is already functional for FBA. I used this provisory model to draw a comparison with *R. solanacearum* and *X. fastidiosa*. As the reconstruction is not final yet, the results are still preliminary, thus the final number of reactions will probably change. I analyzed here global trends, which should be indicative of what should be obtained with the final model.

### 6.3.1 Central metabolism in *R. solanacearum*, *Xcc* and *X. fastidiosa*

Several metabolic losses in *X. fastidiosa* metabolic model reconstruction were found in central metabolic pathways. Central metabolism comprises many interconnected pathways and even though most of the reactions are widely conserved, there are recurrent losses or inactivations even in organisms not described as fastidious. **Figure 6-1** summarizes the losses or inactivations found in *X. fastidiosa*, but also in *R. solanacearum* according to the metabolic model of Peyraud et al. (2016) and in *Xcc* according to the model reconstruction started during my thesis.

Interestingly *R. solanacearum* lacks PFK (phosphofructokinase, part of Embden-Meyerhof-Parnas glycolysis) and, as *X. fastidiosa*, 6PGDH (6-phosphogluconate dehydrogenase, oxidative branch of pentose phosphate pathway) even though it has a very robust and versatile metabolism. These lacks were confirmed by a recent experimental assay with labeled glucose (Jyoti et al., 2020). The Entner-Doudoroff pathway allows the catabolism of glucose-6-phosphate, through EDD (6-phosphogluconate dehydratase) reaction.

Several losses in the central metabolism of *X. fastidiosa* such as transaldolase (TAL), fructose-1,6-bisphosphatase (FBPase), subunit of succinate dehydrogenase (SDH) remain extremely rare among living organisms and the growth of the bacteria with no auxotrophy is still an open question. To better understand their potential contribution to fastidious growth, an interesting perspective would be to perform directed mutagenesis on *Xcc* targeting one of several of the missing genes, taking advantage of its taxonomic proximity with *X. fastidiosa*. Reduction of growth on the obtained *Xcc* mutant would indicate that these losses favor a slow growing phenotype without being lethal. An alternative would be to complement *X. fastidiosa* with the genes from *Xcc*, and potentially improve growth.

*Xcc* central metabolism appears more complete than *X. fastidiosa* and *R. solanacearum*, but one enzyme, the PFK, appears to be poorly active in an assay with labeled glucose (Schatschneider et al., 2014). Hence, glucose-6-phosphate is majorly assimilated through Entner-Doudoroff pathway, and growth assay with an *edd* mutant revealed that growth on glucose was severely impacted (Schatschneider et al., 2013) while the *pfk* mutant was not (**Figure 6-1**). This suggested that *Xcc* PFK is either strongly repressed or not very effective. PFK was purified and characterized by Frese and collaborators (2014). They demonstrated that:

i) PFK is pyrophosphate (PPi) dependent, using PPi instead of ATP as cosubstrate, leading to the reaction:

**Fructose-6-Phosphate + PPi -> Fructose-1,6-bisphosphate + Pi**

instead of the classical:

**Fructose-6-Phosphate + ATP -> Fructose-1,6-bisphosphate + Pi + ADP**

ii) the reaction is reversible, contrary to classical PFK reaction,

iii) *pfk* orthologs of *Xanthomonas* and *Xylella* strains had a very conserved sequence, including the residues specific for PPi-dependent PFK.

It can be hypothesized that PFK is able to perform gluconeogenesis in *Xcc* and in *X. fastidiosa*, and compensate the lack of FBPase in the latter. However, as the enzyme seems poorly efficient or repressed in *Xcc*, gluconeogenesis through PPi-dependent PFK is certainly less efficient, and may contribute to the fastidious growth of *X. fastidiosa*.

Considering xylem colonization, the phase during which the three bacteria are the most metabolically active, it must be reminded that sugars are in very low amount and act only as accessory carbon sources (at least for *R. solanacearum*, see Chapters 4 and 5). Thus, this could explain the simplification of hexose catabolism to a unique metabolic pathway (Entner-Doudoroff). The potential limitation of SDH from tricarboxylic acid cycle and the absence of FBPase from gluconeogenesis (or the use of weakly-efficient PFK), observed in *X. fastidiosa*, are probably more growth impacting *in planta* as they would strongly impact the growth on amino acid such as glutamine.

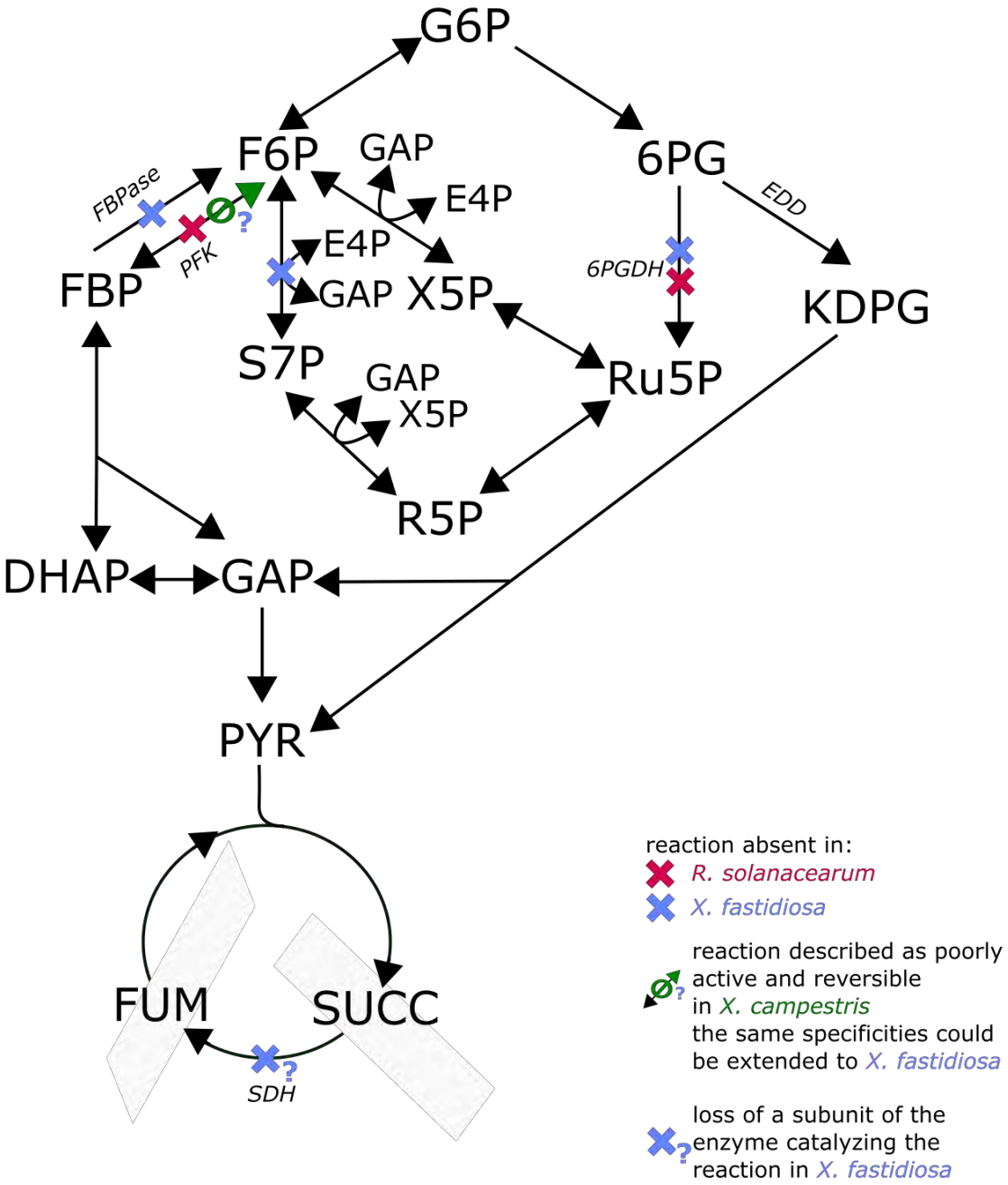



**Figure 6-1: Simplified representation of central metabolism and the losses or inactivations found in *R. solanacearum*, *Xcc* and *X. fastidiosa*.**

Abbreviations for metabolites: G6P: glucose-6-phosphate, 6PG: 6-phosphogluconate, KDPG: 2-keto-3-deoxy-6-phosphogluconate, F6P: fructose-6-phosphate, GAP: glyceraldehyde-3-phosphate, E4P: erythrose-4-phosphate, X5P: xylulose-5-phosphate, Ru5P: ribulose-5-phosphate, R5P: ribose-5-phosphate, S7P: sedopheptulose-7-phosphate, FBP: fructose bisphosphate, DHAP: dihydroxyacetone phosphate, PYR: pyruvate, SUCC: succinate, FUM: fumarate. Abbreviations for reactions: FBPase: fructose-1,6-bisphosphatase, PFK: phosphofructokinase, 6PGDH: 6-phosphogluconate dehydrogenase, EDD: 6-phosphogluconate dehydratase. Only relevant metabolites and reactions were represented. One arrow can represent several reactions.

### 6.3.2 Contrasted metabolic versatilities among xylem-colonizing bacteria and among γ-Proteobacteria

Using the reconstructed metabolic model of *Xcc*, I determined the reactions shared or unique between the three xylem-colonizing bacteria (**Figure 6-2A**). The Venn diagram obtained highlights that *Xcc* has a metabolic versatility in between *R. solanacearum* and *X. fastidiosa*, with a total of 1877 reactions versus 2105 on *R. solanacearum* and 1158 on *X. fastidiosa*. The amount of reaction unique for each organism, which could represent specificities in the lifestyle, also similarly decreases: 965 in *R. solanacearum* versus 638 in *Xcc* and 179 in *X. fastidiosa*. This could be linked to the progressive environmental restriction, from *R. solanacearum* that survives in very different ones, to *Xcc* that is limited to "plant" environments and *X. fastidiosa* that is limited to a homeostatic environment. In line with this environmental reduction, the 367 reactions shared between *R. solanacearum* and *Xcc* (but absent in *X. fastidiosa*) include catabolism of alternative carbon sources (alcohols, sugar alcohols, ferulate, aromatic compounds) and nitrite and sulfur assimilation. These shared reactions also include the biosynthesis of some secondary metabolites that could interact with plants such as ethylene and auxin, which could be seen as pathogenicity determinants (see Chapter 1, section 1.2.5) absent in *X. fastidiosa* that only interacts with dead cells in plant. The two pathogens have apparently intact molybdopterin and cobalamin biosynthetic pathways, which could allow an increased growth rate compared to *X. fastidiosa*, as reported in the article. *X. fastidiosa* shares more reactions with *Xcc* than with *R. solanacearum* (107 *R. solanacearum* & *X. fastidiosa* but not *Xcc* versus 206 *Xcc* & *X. fastidiosa* but not *R. solanacearum*), and this could be explained by their closer phylogenetic relationship.

*E. coli*, *Xcc* and *X. fastidiosa* are all γ-Proteobacteria, this phylogenetic proximity is apparent on the Venn diagram of the three bacteria (**Figure 6-2B**). Indeed, the three circles overlap more than with *R. solanacearum* (**Figure 6-2A**): 804 versus 666 reactions shared between the three, and 663 shared between *Xcc* and *E. coli* versus 367 between *Xcc* and *R. solanacearum*. *E. coli* is a bacterium with a robust and versatile metabolism, similarly to *R. solanacearum*. The progressive metabolic reduction from *E. coli* to *X. fastidiosa* is even more clearly observable graphically than with *R. solanacearum*.

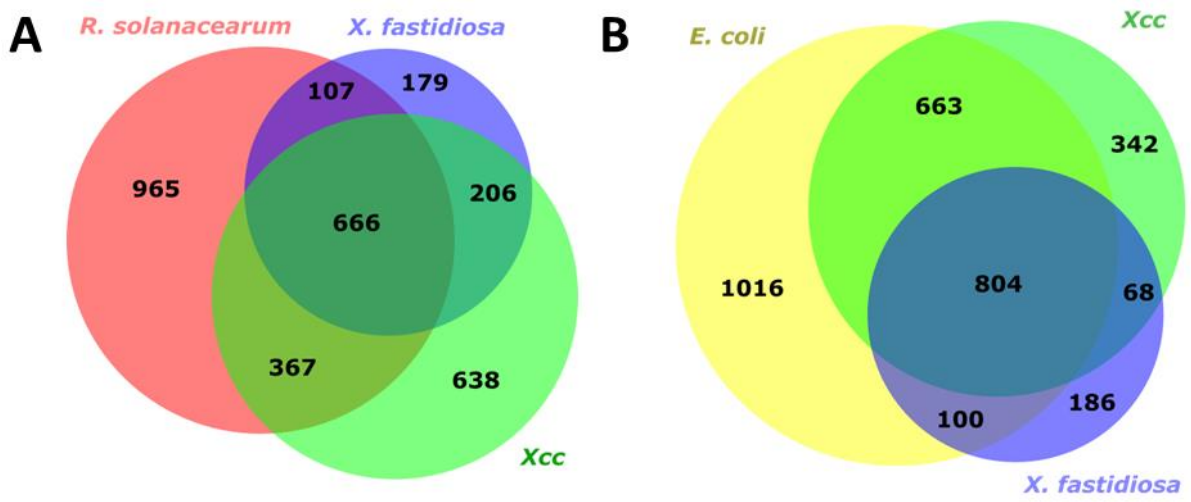



**Figure 6-2: Venn diagram depicting the number of metabolic reactions in *R. solanacearum*, *Xcc* and *X. fastidiosa* (A) and *E. coli*, *Xcc* and *X. fastidiosa* (B).**

Reaction lists are from Orth et al. (2011) for *E. coli* and from Peyraud et al. (2016) for *R. solanacearum* (metabolic module of the model only).

### 6.3.3 Metabolic specificities of xylem-colonizing bacteria

Combining the four datasets of metabolic reactions, I was also able to list the reactions shared between the three xylem-colonizing pathogens but absent in *E. coli*. Only 29 reactions are specific to *Xcc*, *X. fastidiosa* and *R. solanacearum*, and half of them come from unlinked pathways. Six of the shared reactions are nitrilases or peroxidases, that can detoxify the environment from defense molecules secreted by the plant, as reported in the article. Eight of them are reactions from amino acid catabolism, potentially linked to the predominance of amino acids in xylem sap (see Chapters 3 and 4). Extended amino acid catabolism could bring a selective advantage in xylem environment, and would explain the presence of additional reactions compared with *E. coli*, which may be less dependent on amino acids in its natural environment. In addition, two reactions are involved in 3-hydroxybutyric acid metabolism. Curiously, this metabolite is transitory overrepresented in xylem sap during the infection by *R. solanacearum* (see Chapter 4) and could be excreted by the plant and assimilated by the bacteria, according to the reactions found in the plant – pathogen metabolic model (Chapter 5). This conservation could be an indicator of its importance in xylem sap colonization. However, the role of this metabolite in the interaction is not yet elucidated.

## 6.4 Evolutionary trajectories toward metabolic reduction and fastidious growth

### 6.4.1 Loss of FBPase among the *Xanthomonadaceae*

A major and intriguing result from the metabolic reconstruction of *X. fastidiosa* is the lack of gene coding for FBPase, as well as the lack of gene for TAL, two conserved enzymes of central carbon metabolism (**Figure 6-1**). To estimate if these losses were ancestral and shared among phylogenetically close bacteria, I searched for orthologs of FBPase and TAL genes among the *Xanthomonadaceae*. I found that TAL is conserved among all the species with available genomes of the *Xanthomonadaceae* family, but that FBPase is missing in one unique species, *X. albilineans*. The structure of the phylogenetic tree (**Figure 6-3**) suggests that loss of FBPase occurred independently twice among the *Xanthomonadaceae*, which makes this particularity even more intriguing, keeping in mind that it is extremely rare among all living organisms (as evoked in the Discussion section of the article).

Interestingly, *X. albilineans* shares some features with *X. fastidiosa*. It is also described as a relatively slow-growing bacterium (Birch, 2001) and has a relatively reduced genome (3.7 - 3.9 Mb). In addition, it belongs to the members of the *Xanthomonadaceae* that have no Hrp (Hypersensitive response and pathogenicity) T3SS, as *X. fastidiosa*. The other members, such as some strains of *X. arboricola*, are for most of them described as commensal bacteria (Merda et al., 2017). It is tempting to hypothesize that as *X. fastidiosa*, *X. albilineans* results from an evolution toward a self-limiting behavior, with a reduced metabolic potential limiting their ability to grow fast, taking advantage of a latent and limited detection by the plant as it remains at low population level. The loss of Hrp T3SS, key pathogenicity determinant of several bacterial plant pathogens, could point toward a less aggressive behavior, and as mentioned for *X. fastidiosa,* an ongoing evolution toward commensalism (Roper et al., 2019). *X. albilineans* also lacks the operon for EPS (xanthan gum) synthesis, another important pathogenicity determinant shared among plant pathogens (Chapter 1, 1.2.2). The diseases provoked by the two pathogens, and their sometime dramatic agronomical consequences could be due, as already described in *X. fastidiosa* (Sun et al., 2013) to the extensive production of tylose by the plant at the detection of the pathogen. As *X. albilineans* transmission is mostly reported to be due to infected

cuttings (Birch, 2001), environmental restriction to homeostatic environment could shape these similar evolutionary trajectories.

Overall, the evolutionary trajectory appears to be less marked in *X. albilineans* than in *X. fastidiosa*, as the TAL gene is still present and the genome reduction is less drastic. We must also keep in mind that *X. albineans* was detected in non-vascular tissues (Mensi et al., 2013) and seems thus more versatile in environment colonization than *X. fastidiosa*. Also, the acquisition (almost unique among the *Xanthomonadaceae*) of SPI-1 (*Salmonella* pathogenicity island 1) T3SS (Marguerettaz et al., 2011) could indicate that this trajectory have been reversed at some time point. SPI-1 T3SS does not appear to be essential for virulence (Marguerettaz et al., 2011), but it could still play a role in its lifestyle and enhance its capacity to survive outside xylem vessels. Also, the losses of Hrp T3SS and EPS associated genes could be counterbalanced by the acquisition of other major virulence determinants such as albicidin toxin, that is able to block DNA replication and plastid development and to disrupt host response (Birch, 2001). *X. albilineans* seems to have an intermediate and oscillating lifestyle between a self-limiting, xylem-restricted and almost commensal behavior as *X. fastidiosa*, and a still invasive and fast colonizing behavior allowed by larger metabolic potential and T3SS.

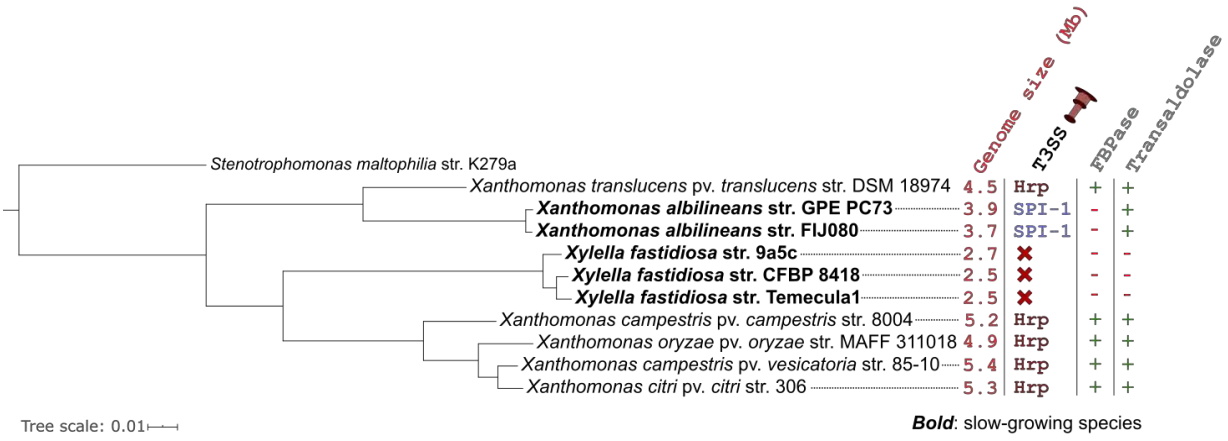



**Figure 6-3: Phylogenetic tree of strains from the *Xanthomonadaceae* family and some of their properties.**

T3SS: Type-Three Secretion System. FBPase: fructose-1,6-bisphosphatase. Hrp: Hypersensitive response and pathogenicity. SPI-1: *Salmonella* pathogenicity island 1. The species tree (generated by L. Cottret) has been inferred from proteomes by Orthofinder 2.2.7 with the following options: -S diamond -M msa (Emms & Kelly, 2015).

### 6.4.2 Fastidious and xylem-limited bacteria in RSSC

As presented in Chapter 1, *R. solanacearum* is a complex of species (RSSC) with an important diversity between strains, that are organized in four phylotypes or three species, according to phylogenetic classification. Among this wide range of species, two subspecies from the phylotype IV have a lifestyle rather different than the one previously presented: *R. syzygii* subsp. *celebesensis* and *R. syzygii* subsp. *syzygii* (**Table 6-1**). These two species are insect-transmitted, appear to be restricted to a unique host, and are described as fastidious and xylem-limited. While it is tempting to imagine that they have undergone global genome reduction as *X. fastidiosa* and *X. albilineans*, there is no drastic size reduction, although they have the smallest genomes among the RSSC (Remenant et al., 2011). Contrary to what is observed in *X. fastidiosa* and *X. albilineans*, these two *R. syzygii* subspecies do not lack genes coding for central enzymes such as FBPase or TAL. They also have an intact Hrp T3SS, although they have the two smallest T3E repertories of RSSC (Sabbagh et al., 2019).

Even though some peculiarities are not observed in these *R. syzygii* subspecies, the similarities of these specializations in growth, lifestyle and host specificities in very distant families (*Xanthomonadaceae* and RSSC) encourages pursuing the reconstruction and analysis of other metabolic models in the two families. A work is ongoing in my team to reconstruct metabolic models representative of RSSC, and could unravel metabolic specificities or fragilities in these fastidious strains.

**Table 6-1: Properties of two *R. syzygii* subspecies.**

Informations presented are from Remenant et al. (2011), excepted the T3E repertory size obtained from https://iant.toulouse.inra.fr/T3E. Other strains of *R. solanacearum* usually have a genome size around 5.7 – 5.8 Mb and a T3E repertory size around 65 - 75.

| **Species** | ***R. syzygii* subsp. *celebesensis*** | ***R. syzygii* subsp. *syzygii*** |
|---|---|---|
| Previous denomination | *R. solanacearum* BDB | *R. syzygii* |
| Transmission | Pollinating insects | Cercopoid insects |
| Host range | Banana tree only | Clove tree only |
| Specificities | Fastidious and xylem-limited | Fastidious and xylem-limited |
| Genome size (Mb) | 5.2 | 5.4 |
| T3E repertory size | 42 | 45 |

# Chapter 7
# General discussion and perspectives

## 7.1 Systems biology on metabolism

### 7.1.1 The virtuous cycle between data and modelling

The work presented in the previous chapters did not required new constraint-based metabolic modeling methodologies, but rather used and adapted existing ones (model reconstruction, multi-compartment modeling) to untangle the complexity of metabolism in a plant – pathogen system, and acquired the data necessary for rigorous calibration of the models generated. It illustrates that besides developing and improving modeling frameworks, relevant models often require important or complex experimental data. For example, diverse host – microbe metabolic models exist (see Dunphy & Papin (2018) for review with human as host and see Gerlin et al. (2020) for review with plant as host). However, predictions and interpretations could be rapidly biased, as several parameters are complex to determine, such as the assimilation capacities of the microbe in the mixture of substrates encountered *in vivo*. In addition, it has been very difficult to calibrate the metabolic model of *Xylella fastidiosa* as it is still impossible to grow the bacterium in minimal medium, and this limited the range of possible interpretations. By building the plant metabolic model and analyzing the previous work done on plant model reconstructions, I realized that knowledge on plant primary metabolism was still sparse and fragmentary, due to the complexity of its compartmentation and with less literature available on biochemical aspects than for mammals, yeast or bacteria. These strong requirements on experimental data could create a gap between the advancements on modeling techniques on one side, and the possible use on an organism of interest on the other side. However, these requirements also create virtuous circles, as they open biological questions previously eluded. For example, our physiological and metabolic characterization of bacterial wilt dynamics

(Chapter 4), guided by modeling needs, provided interesting insights on the interaction even before integrating them into a model.

### 7.1.2 Modeling perspectives

In this work, genome-scale metabolic modeling methodology was used as a tool to understand on one side the interactions in multi-organ/multi-organism systems, and on the other side to compare different organisms. A direct perspective of this work would be to pursue on this work, by building additional metabolic models, such as other members of the RSSC (in particular *R. syzygii*) and of the *Xanthomonadaceae* (in particular *X. albilineans*, after validating *Xcc* model) as well as to model the interaction of other organisms (such as *X. fastidiosa*) with plant hosts.

While these perspectives are very promising, there are also important additional work which can be performed on the models already built. The numerous specificities of the genome-scale metabolic model of *X. fastidiosa* are a gold mine to explore metabolic "minimalism" and pathway/enzyme efficiencies. For example, the limitations imposed by cofactor-independent enzymes (such as the cobalamin-independent ones) or by a gluconeogenesis mediated by a poorly efficient PFK (Chapter 6, section 6.3.1) open a way toward modeling how they restrict growth. To this end, several constraint-based techniques integrating enzyme synthesis cost and enzyme efficiency were developed (Goelzer et al., 2015; Waldherr et al., 2015), and could be adapted to *X. fastidiosa*.

Having generated a plant metabolic model in its vegetative growth also encourages us to add a tomato fruit metabolic model, based on the literature available on tomato fruit metabolism (Bénard et al., 2015; Biais et al., 2014; Colombié et al., 2015, 2017). This work would require additional experimental data such as the tomato fruit relative weight in the plant, the transpiration rate and the growth rate constraining the model. In addition, the metabolic model of *R. solanacearum* includes the biosynthesis and excretion of numerous macromolecules used for virulence (e.g. T3Es, cell-wall degrading enzymes) constituting a "macromolecule module" (Peyraud et al., 2016). This allows to include their excretion fluxes in the simulations in addition to biomass production, as they require additional energy and carbon and thus impact growth. To mirror this growth/virulence trade-off on the pathogen side, it would be relevant to represent the growth/immunity trade-off of the plant, compiling the knowledge on defense macromolecules. The main challenge on this part would be to quantify their production, which is complex to establish in a whole plant model.

## 7.2 Nutrition and lifestyles of xylem-colonizing bacteria

### 7.2.1 Looking for carbon sources *in planta*

Different observations point toward the importance of glutamine in the nutrition of xylem-colonizing bacterial nutrition, as for example for *R. solanacearum* in a tomato plant. First, glutamine is the principal organic carbon molecule of xylem sap (between 75% and 80% of organic carbon) (Chapter 3, Chapter 4). Secondly, *R. solanacearum* grows rapidly on glutamine, particularly when associated with other carbon sources such as asparagine (Baroukh et al., accepted). Thirdly, we predict an important delay in disease progression in the absence of glutamine assimilation (Chapter 6). Fourthly, glutamine is strongly depleted during *R. solanacearum* colonization, and reached zero on some infected samples (Chapter 4), which is striking compared to his relative abundance in healthy xylem. Fifthly, prevalence of glutamine in xylem sap is very well conserved among plant species, while other xylem nutrients vary according to botanical families, varieties or culture conditions (Chapter 3). Thus, glutamine appears as a reliable source of carbon for a pathogen such as *R. solanacearum* able to invade plants from diverse botanical families.

Curiously, after decades of work on *R. solanacearum*'s biology worldwide, the literature did not pinpoint the major role of glutamine in the plant-pathogen interaction before our work (Gerlin et al., 2021). In contrast, some studies pointed toward the utilization of sugars, and in particular sucrose (Hamilton et al., 2021; Jacobs et al., 2012). Sugar level in plant – pathogen interaction can be modulated by action of T3Es (see Poueymiro et al., (2014) for trehalose phosphate synthase in *R. solanacearum*; see Timilsina et al. (2020) for review in *Xanthomonas* spp.) and it has often been proposed that pathogen nutrition is supported by a sugar release induced by the pathogen. However, this scenario seems implausible for *R. solanacearum* if we consider that: i) sugars, and in particular sucrose, do not sustain a high growth rate (Peyraud et al., (2016) for preliminary data on Biolog and Baroukh et al., (accepted) for growth on minimal medium), ii) we found very scarce proportion of sugars in xylem sap, even in a preliminary metabolomic analysis of xylem sap infected by a sugar transporter mutant (not published), iii) sugar concentration seems to be condition- and plant-dependent as the concentrations reported in the literature are extremely variable (Chapter 3, Chapter 4) and not so abundant. There is an increasing literature

presenting sugars as signal molecules regulating the plant defenses and physiology (reviewed by Kanwar & Jha, 2018), so the role of effectors modulating plant sugars content could be a way to disrupt plant immune responses. Investing sophisticated virulence mechanisms to release sugars in xylem while this environment is already rich-enough for bacterial proliferation (Chapter 4, Chapter 5) does not appear justified in terms of cost rationalization: “Who breaks a butterfly upon a wheel?”

The importance of glutamine in xylem sap is not so surprising: glutamine is the primary way to store nitrogen uptaken, through glutamine synthetase (EC 6.3.1.2), before subsequently converting it into other amino acids, so it is a way to convey nitrogen through the plant without the potential toxicity of an excess of inorganic nitrogen (Britto & Kronzucker, 2002). Similarly, in animals, glutamine is the most abundant amino acid in plasma and is used for interorgan amino acid transport (Brosnan, 2003). It worth also noticing that glutamine is highly metabolized during the fast proliferation of cancer cells, as nitrogen but also carbon source (DeBerardinis et al., 2007; Shlomi et al., 2011).

Metabolomics studies of plant xylem sap remains sparse in the literature, being sometimes incomplete or performed in very different conditions. It would be of great interest to perform a large-scale quantitative analysis of plant xylem fluids, by cultivating different plant species in similar environmental conditions, extracting xylem sap for plants of similar age or weight, and finally measuring concentrations using the same analytical technique. We could potentially relate the concentration of glutamine, and the presence of other complementary carbon sources, to the occurrence and severity of bacterial wilt in the different species, or more broadly to the susceptibility to xylem-colonizing organisms.

Apprehending glutamine level as a way to store inorganic nitrogen and to convey nitrogen in a multi-organ organism fits with the relationship between ammonia concentration in the soil and glutamine concentration in xylem (Chapter 3, Bialczyk et al. (2004)). As reducing ammonia in the soil also decreases the susceptibility of tomato to *R. solanacearum* (Kelman, 1950), we can hypothesize that modulating glutamine in xylem sap, notably through plant nutrition, could help controlling bacterial wilt on crops. In areas subjected to the bacterial wilt disease, the effect of nitrogen fertilizers on the concentration of glutamine in the xylem flow and thus, in turn, on the proliferation of *R. solanacearum* could be an important parameter to take into

consideration. This could restrain the pathogen to a population density that is not affecting plant physiology (under $10^6$ cells per g of stem FW in our data, see Chapter 4 and Chapter 5) and limit the suddenness and severity of an outbreak in the field. Modulating disease severity through attenuating the pathogen access to nutrients would also favor a coexistence of the bacteria and the plant together as the selective pressure would be less strong than using harsh treatments or the development of fully resistant crops that could be overpassed by the evolution of pathogen populations.

### 7.2.2 Xylem, the same battlefield but for different purposes?

In the previous chapter, I reconstructed and analyzed the metabolism of *Xylella fastidiosa*. As *R. solanacearum, X. fastidiosa* is a "vascular wilt pathogen", meaning that both invade xylem and provoke relatively close wilting symptoms. However, *X. fastidiosa* and *R. solanacearum* have physiological properties diametrically opposed. *R. solanacearum* impresses by its very rapid proliferation *in planta* and the bacterial density reached, while the replication time of *X. fastidiosa*, among the lowest in the bacterial kingdom, raises questions on the biological bases of its capacity to provoke diseases and notably spectacular outbreaks.

We modeled that, at low densities, the amount of plant resources exploited by *R. solanacearum* is negligible for the plant (Chapter 5). The pathogen behaves as a commensal from a nutritional point of view, even though the density is still above the required one to activate quorum sensing-dependent virulence factors (Chapter 4). This "transitory commensalism" would explain the important amount of non-symptomatic hosts of *X. fastidiosa*: intrinsic properties of the bacterium, such as its metabolism, evolved toward a self-limiting behavior (Roper et al., 2019) that restrain the switch toward disease that appears irrevocable above a specific bacterial density in our plant – pathogen model. *X. fastidiosa* is thus able to maintain the plant in a healthy state. This is more advantageous to spread to another plant through insects, as insects would not feed on a heavily diseased plant. The fast growth of *R. solanacearum* could be seen as more profitable for it than for *X. fastidiosa*: its capacity to spread among different plants in a field depends on its ability to generate a heavy bacterial load in the plant, to afterward thrive into the soil and water and eventually be still numerous enough to invade another plant.

According to this, the "vascular wilts" provoked by *R. solanacearum* and *X. fastidiosa* have different "mechanical" explanations. For *X. fastidiosa*, it has been shown that it

is mainly due to the plant overproduction of tyloses in response to the detection of *X. fastidiosa* (Sun et al., 2013), while in *R. solanacearum*, it has been very recently demonstrated that the bacterial load by itself obstructs the xylem vessels (Ingel et al., 2021). Thus, the disease could be seen as "accidental" for *X. fastidiosa*, whose expansion across plants depends on the propagation in healthy plant, while the expansion of *R. solanacearum* is dependent on yielding to the highest bacterial density from plant materials, with fewer advantages on keeping it alieve. In Chapter 6, I often described the metabolic capacities of *X. fastidiosa* as "limited", "restricted" or "weakly robust", which could be perceived by the reader as a weakness of the pathogen. However, in a broader view, these specificities appear in fact as a sophisticated way to switch toward a more profitable lifestyle than killing its host.

As suggested at the end of Chapter 6, using mutagenesis or complementation to synthetically change the metabolic capacities of a plant pathogen would be of interest. We could imagine that a *R. solanacearum* deprived of some metabolic parts sustaining fast growth could remain tolerated by the plant a longer time, while a faster growing *X. fastidiosa* could reach high population densities and induce disease in plant that usually remain asymptomatic. However, this approach will be challenged by the tight regulation that controls the activation of metabolic pathways and could prevent the possibility to reshape bacterial metabolism.

To conclude, even though *X. fastidiosa* and *R. solanacearum* colonize the same tissue, analyzing their respective genome-scale metabolic models revealed that their metabolisms have been tailored toward opposite manners of "being a pathogen".

# References


Agrios, G. N. (2005). *Plant Pathology* (Fifth Edit). Elsevier Inc. https://doi.org/https://doi.org/10.1016/C2009-0-02037-6

Akimoto-Tomiyama, C., Furutani, A., & Ochiai, H. (2014). Real time live imaging of phytopathogenic bacteria xanthomonas campestrispv. campestrisMAFF106712 in "plant sweet home." *PLoS ONE*, *9*(4). https://doi.org/10.1371/journal.pone.0094386

Arnold, A., & Nikoloski, Z. (2014). Bottom-up Metabolic Reconstruction of Arabidopsis and Its Application to Determining the Metabolic Costs of Enzyme Production. *Plant Physiology*, *165*(3), 1380–1391. https://doi.org/10.1104/PP.114.235358

Baroukh, C., Muñoz-tamayo, R., Steyer, J., & Bernard, O. (2013). DRUM : A New Framework for Metabolic Modeling under Non- Balanced Growth. Application to the Carbon Metabolism of Unicellular Microalgae. *PloS One*, *9*(8), 113–118. https://doi.org/10.1371/journal.pone.0104499

Baroukh, C., Zemouri, M., & Genin, S. (n.d.). Trophic preferences of the pathogen Ralstonia solanacearum and consequences on its growth in xylem sap. *MicrobiologyOpen*, *accepted*.

Bénard, C., Bernillon, S., Biais, B., Osorio, S., Maucourt, M., Ballias, P., Deborde, C., Colombié, S., Cabasson, C., Jacob, D., Vercambre, G., Gautier, H., Rolin, D., Génard, M., Fernie, A. R., Gibon, Y., & Moing, A. (2015). Metabolomic profiling in tomato reveals diel compositional changes in fruit affected by source-sink relationships. *Journal of Experimental Botany*, *66*(11), 3391–3404. https://doi.org/10.1093/jxb/erv151

Berger, S., Sinha, A. K., & Roitsch, T. (2007). Plant physiology meets phytopathology: plant primary metabolism and plant–pathogen interactions. *Journal of Experimental Botany*, *58*(15–16), 4019–4026. https://doi.org/10.1093/JXB/ERM298

Biais, B., Bénard, C., Beauvoit, B., Colombié, S., Prodhomme, D., Ménard, G., Bernillon, S., Gehl, B., Gautier, H., Ballias, P., Mazat, J. P., Sweetlove, L., Génard, M., & Gibon, Y. (2014). Remarkable reproducibility of enzyme activity profiles in tomato fruits grown under contrasting environments provides a roadmap for studies of fruitmetabolism. *Plant Physiology*, *164*(3), 1204–1221. https://doi.org/10.1104/pp.113.231241

Bialczyk, J., Lechowski, Z., & Dziga, D. (2004). Composition of the xylem sap of tomato seedlings cultivated on media with HCO 3 – and nitrogen source as NO 3 – or NH 4 +. *Plant and Soil 2004 263:1*, *263*(1), 265–272. https://doi.org/10.1023/B:PLSO.0000047739.11698.CA

Birch, R. G. (2001). Xanthomonas albilineans and the antipathogenesis approach to disease control. *Molecular Plant Pathology*, *2*(1), 1–11. https://doi.org/10.1046/J.1364-3703.2001.00046.X

Bordbar, A., Lewis, N. E., Schellenberger, J., Palsson, B., & Jamshidi, N. (2010). Insight into human alveolar macrophage and M. tuberculosis interactions via metabolic reconstructions. *Molecular Systems Biology*, *6*(422). https://doi.org/10.1038/msb.2010.68

Boshou, L., Zhihui, S., Yong, L., Yujun, T., Dong, L., & Naixiong, D. (1998). Reaction to latent infection by Ralstonia solanacearum in groundnut. *Zhongguo You Liao Zuo Wu Xue Bao = Chinese Journal of Oil Crop Sciences*, *20*(4), 61–65. https://europepmc.org/article/cba/320283

Botero, D., Monk, J., Rodríguez Cubillos, M. J., Rodríguez Cubillos, A., Restrepo, M., Bernal-Galeano, V., Reyes, A., González Barrios, A., Palsson, B. Ø., Restrepo, S., & Bernal, A. (2020). Genome-Scale Metabolic Model of Xanthomonas phaseoli pv. manihotis: An Approach to Elucidate Pathogenicity at the Metabolic Level. *Frontiers in Genetics*, *0*, 837. https://doi.org/10.3389/FGENE.2020.00837

Botero, K., Restrepo, S., & Pinzón, A. (2018). A genome-scale metabolic model of potato late blight suggests a photosynthesis suppression mechanism 06 Biological Sciences 0607 Plant Biology. *BMC Genomics*, *19*(8), 31–44. https://doi.org/10.1186/s12864-018-5192-x

Britto, D. T., & Kronzucker, H. J. (2002). NH4+ toxicity in higher plants: a critical review. *Journal of Plant Physiology*, *159*(6), 567–584. https://doi.org/10.1078/0176-1617-0774

Brosnan, J. T. (2003). Interorgan Amino Acid Transport and its Regulation. *The Journal of Nutrition*, *133*(6), 2068S-2072S. https://doi.org/10.1093/JN/133.6.2068S

Brunings, A. M., & Gabriel, D. W. (2003). Xanthomonas citri: breaking the surface. *Molecular Plant Pathology*, *4*(3), 141–157. https://doi.org/10.1046/J.1364-3703.2003.00163.X

Burgard, A. P., Nikolaev, E. V., Schilling, C. H., & Maranas, C. D. (2004). Flux coupling analysis of genome-scale metabolic network reconstructions. *Genome Research*, *14*(2), 301–312. https://doi.org/10.1101/gr.1926504

Büttner, D., & Bonas, U. (2010). Regulation and secretion of *Xanthomonas* virulence factors. *FEMS Microbiology Reviews*, *34*(2), 107–133. https://doi.org/10.1111/j.1574-6976.2009.00192.x

Caspi, R., Billington, R., Ferrer, L., Foerster, H., Fulcher, C. A., Keseler, I. M., Kothari, A., Krummenacker, M., Latendresse, M., Mueller, L. A., Ong, Q., Paley, S., Subhraveti, P., Weaver, D. S., & Karp, P. D. (2016). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Research*, *44*(D1), D471–D480. https://doi.org/10.1093/nar/gkv1164

Chatterjee, S., Almeida, R. P. P., & Lindow, S. (2008). Living in two Worlds: The Plant and Insect Lifestyles of Xylella fastidiosa. *Annual Review of Phytopathology*, *46*(1), 243–271. https://doi.org/10.1146/annurev.phyto.45.062806.094342

Chou, F. L., Chou, H. C., Lin, Y. S., Yang, B. Y., Lin, N. T., Weng, S. F., & Tseng, Y. H. (1997). The Xanthomonas campestris gumD gene required for synthesis of xanthan gum is involved in normal pigmentation and virulence in causing black rot. *Biochemical and Biophysical Research Communications*, *233*(1), 265–269. https://doi.org/10.1006/bbrc.1997.6365

Clough, S. J., Lee, K. E., Schell, M. A., & Denny, T. P. (1997). A two-component system in Ralstonia (Pseudomonas) solanacearum modulates production of PhcA-regulated virulence factors in response to 3-hydroxypalmitic acid methyl ester. *Journal of Bacteriology*, *179*(11), 3639–3648. https://doi.org/10.1128/jb.179.11.3639-3648.1997

Colombié, S., Beauvoit, B., Nazaret, C., Bénard, C., Vercambre, G., Le Gall, S., Biais, B., Cabasson, C., Maucourt, M., Bernillon, S., Moing, A., Dieuaide-Noubhani, M., Mazat, J. P., & Gibon, Y. (2017). Respiration climacteric in tomato fruits elucidated by constraint-based modelling. *New Phytologist*, *213*(4), 1726–1739. https://doi.org/10.1111/nph.14301

Colombié, S., Nazaret, C., Bénard, C., Biais, B., Mengin, V., Solé, M., Fouillen, L., Dieuaide-Noubhani, M., Mazat, J. P., Beauvoit, B., & Gibon, Y. (2015). Modelling central metabolic fluxes by constraint-based optimization reveals metabolic reprogramming of developing Solanum lycopersicum (tomato) fruit. *Plant Journal*, *81*(1), 24–39. https://doi.org/10.1111/tpj.12685

Corral, J., Sebastià, P., Coll, N. S., Barbé, J., Aranda, J., & Valls, M. (2020). Twitching and Swimming Motility Play a Role in Ralstonia solanacearum Pathogenicity. *MSphere*, *5*(2). https://doi.org/10.1128/msphere.00740-19

Cottret, L., Frainay, C., Chazalviel, M., Cabanettes, F., Gloaguen, Y., Camenen, E., Merlet, B., Heux, S., Portais, J.-C., Poupin, N., Vinson, F., & Jourdan, F. (2018). MetExplore: collaborative edition and exploration of metabolic networks. *Nucleic Acids Research*, *46*(W1), W495–W502. https://doi.org/10.1093/nar/gky301

da Silva, A. C. R., Ferro, J. A., Reinach, F. C., Farah, C. S., Furlan, L. R., Quaggio, R. B., Monteiro-Vitorello, C. B.,

Sluys, M. A. Van, Almeida, N. F., Alves, L. M. C., do Amaral, A. M., Bertolini, M. C., Camargo, L. E. A., Camarotte, G., Cannavan, F., Cardozo, J., Chambergo, F., Ciapina, L. P., Cicarelli, R. M. B., … Kitajima, J. P. (2002). Comparison of the genomes of two Xanthomonas pathogens with differing host specificities. *Nature 2002 417:6887*, *417*(6887), 459–463. https://doi.org/10.1038/417459a

Da Silva, F. R., Vettore, A. L., Kemper, E. L., Leite, A., & Arruda, P. (2001). Fastidian gum: The Xylella fastidiosa exopolysaccharide possibly involved in bacterial pathogenicity. *FEMS Microbiology Letters*, *203*(2), 165–171. https://doi.org/10.1016/S0378-1097(01)00348-2

Darsonval, A., Darrasse, A., Durand, K., Bureau, C., Cesbron, S., & Jacques, M. A. (2009). Adhesion and fitness in the bean phyllosphere and transmission to seed of xanthomonas fuscans subsp. Fuscans. *Molecular Plant-Microbe Interactions*, *22*(6), 747–757. https://doi.org/10.1094/MPMI-22-6-0747

Das, A., Rangaraj, N., & Sonti, R. V. (2009). Multiple adhesin-like functions of Xanthomonas oryzae pv. oryzae are involved in promoting leaf attachment, entry, and virulence on rice. *Molecular Plant-Microbe Interactions*, *22*(1), 73–85. https://doi.org/10.1094/MPMI-22-1-0073

David, L., Marashi, S. A., Larhlimi, A., Mieth, B., & Bockmayr, A. (2011). FFCA: A feasibility-based method for flux coupling analysis of metabolic networks. *BMC Bioinformatics*, *12*(1), 1–7. https://doi.org/10.1186/1471-2105-12-236

De Oliveira Dal'Molin, C. G., Orellana, C., Gebbie, L., Steen, J., Hodson, M. P., Chrysanthopoulos, P., Plan, M. R., McQualter, R., Palfreyman, R. W., & Nielsen, L. K. (2016). Metabolic reconstruction of setaria italica: A systems biology approach for integrating tissue-specific omics and pathway analysis of bioenergy grasses. *Frontiers in Plant Science*, *7*(AUG2016), 1138. https://doi.org/10.3389/fpls.2016.01138

de Oliveira Dal'Molin, C. G., Quek, L.-E., Saa, P. A., & Nielsen, L. K. (2015). A multi-tissue genome-scale metabolic modeling framework for the analysis of whole plant systems. *Frontiers in Plant Science*, *6*, 4. https://doi.org/10.3389/fpls.2015.00004

De Oliveira Dal'Molin, C. G., Quek, L. E., Palfreyman, R. W., Brumbley, S. M., & Nielsen, L. K. (2010). AraGEM, a genome-scale reconstruction of the primary metabolic network in Arabidopsis. *Plant Physiology*, *152*(2), 579–589. https://doi.org/10.1104/pp.109.148817

DeBerardinis, R. J., Mancuso, A., Daikhin, E., Nissim, I., Yudkoff, M., Wehrli, S., & Thompson, C. B. (2007). Beyond aerobic glycolysis: Transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *Proceedings of the National Academy of Sciences*, *104*(49), 19345–19350. https://doi.org/10.1073/PNAS.0709747104

Denny, T. P. (1995). Involvement of bacterial polysaccharides in plant pathogenesis. In *Annual Review of Phytopathology* (Vol. 33, pp. 173–197). Annual Reviews Inc. https://doi.org/10.1146/annurev.py.33.090195.001133

Desaint, H., Aoun, N., Deslandes, L., Vailleau, F., Roux, F., & Berthomé, R. (2021). Fight hard or die trying: when plants face pathogens under heat stress. In *New Phytologist* (Vol. 229, Issue 2, pp. 712–734). Blackwell Publishing Ltd. https://doi.org/10.1111/nph.16965

Deslandes, L., & Genin, S. (2014). Opening the Ralstonia solanacearum type III effector tool box: Insights into host cell subversion mechanisms. *Current Opinion in Plant Biology*, *20*, 110–117. https://doi.org/10.1016/j.pbi.2014.05.002

diCenzo, G. C., Benedict, A. B., Fondi, M., Walker, G. C., Finan, T. M., Mengoni, A., & Griffitts, J. S. (2018). Robustness encoded across essential and accessory replicons of the ecologically versatile bacterium Sinorhizobium meliloti. In *PLoS Genetics* (Vol. 14, Issue 4). https://doi.org/10.1371/journal.pgen.1007357

diCenzo, G. C., Checcucci, A., Bazzicalupo, M., Mengoni, A., Viti, C., Dziewit, L., Finan, T. M., Galardini, M., & Fondi, M. (2016). Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in Sinorhizobium meliloti. *Nature Communications*, *7*. https://doi.org/10.1038/ncomms12219

diCenzo, G. C., Tesi, M., Pfau, T., Mengoni, A., & Fondi, M. (2020). Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. *Nature Communications*, *11*(1), 1–11. https://doi.org/10.1038/s41467-020-16484-2

Dunphy, L. J., & Papin, J. A. (2018). Biomedical applications of genome-scale metabolic network reconstructions of human pathogens. *Current Opinion in Biotechnology*, *51*, 70–79. https://doi.org/10.1016/j.copbio.2017.11.014

Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, *16*(1), 1–14. https://doi.org/10.1186/s13059-015-0721-2

Escobar, M. A., & Dandekar, A. M. (2003). Agrobacterium tumefaciens as an agent of disease. In *Trends in Plant Science* (Vol. 8, Issue 8, pp. 380–386). Elsevier Ltd. https://doi.org/10.1016/S1360-1385(03)00162-6

Fagard, M., Launay, A., Clément, G., Courtial, J., Dellagi, A., Farjad, M., Krapp, A., Soulié, M. C., & Masclaux-Daubresse, C. (2014). Nitrogen metabolism meets phytopathology. *Journal of Experimental Botany*, *65*(19), 5643–5656. https://doi.org/10.1093/jxb/eru323

FAO, IFAD, UNICEF, WFP, & WHO. (2020). In Brief to The State of Food Security and Nutrition in the World 2020. Transforming food systems for affordable healthy diets. In *In Brief to The State of Food Security and Nutrition in the World 2020*. FAO, WHO, IFAD, WFP, UNICEF. https://doi.org/10.4060/ca9699en

Fargier, E., & Manceau, C. (2007). Pathogenicity assays restrict the species Xanthomonas campestris into three pathovars and reveal nine races within X. campestris pv. campestris. *Plant Pathology*, *56*(5), 805–818. https://doi.org/10.1111/J.1365-3059.2007.01648.X

Fegan, M., & Prior, P. (2005). How complex is the "Ralstonia solanacearum species complex"? *Bacterial Wilt Disease and the Ralstonia Solanacearum Species Complex*, *July*, 449–461.

Feist, A. M., Herrgård, M. J., Thiele, I., Reed, J. L., & Palsson, B. (2009). Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology*, *7*(2), 129–143. https://doi.org/10.1038/nrmicro1949

Feist, A. M., & Palsson, B. O. (2010). The biomass objective function. *Current Opinion in Microbiology*, *13*(3), 344–349. https://doi.org/10.1016/j.mib.2010.03.003

Ferguson, A. R., & Johnston, J. S. (1980). Phaseolotoxin : chlorosis, ornithine accumulation and inhibition of ornithine carbamoyltransferase in different plants. *Physiological Plant Pathology*, *16*(2), 269–275. https://doi.org/10.1016/0048-4059(80)90041-7

Frese, M., Schatschneider, S., Voss, J., Vorhölter, F. J., & Niehaus, K. (2014). Characterization of the pyrophosphate-dependent 6-phosphofructokinase from Xanthomonas campestris pv. campestris. *Archives of Biochemistry and Biophysics*, *546*, 53–63. https://doi.org/10.1016/J.ABB.2014.01.023

Genin, S. (2010). Molecular traits controlling host range and adaptation to plants in Ralstonia solanacearum. *New Phytologist*, *187*(4), 920–928. https://doi.org/10.1111/J.1469-8137.2010.03397.X

Genin, S., & Denny, T. P. (2012). Pathogenomics of the *Ralstonia solanacearum* Species Complex. *Annual Review of Phytopathology*, *50*(1), 67–89. https://doi.org/10.1146/annurev-phyto-081211-173000

Gerlin, L., Frainay, C., Jourdan, F., Baroukh, C., & Prigent, S. (2020). Plant genome-scale metabolic networks. In *Advances in Botanical Research*. Academic Press Inc. https://doi.org/10.1016/bs.abr.2020.09.021

Giampetruzzi, A., Chiumenti, M., Saponari, M., Donvito, G., Italiano, A., Loconsole, G., Boscia, D., Cariddi, C., Martelli, G. P., & Saldarelli, P. (2015). Draft genome sequence of the Xylella fastidiosa CoDiRO strain. *Genome Announcements*, *3*(1), 0–1. https://doi.org/10.1128/genomeA.01538-14

Glont, M., Nguyen, T. V. N., Graesslin, M., Hälke, R., Ali, R., Schramm, J., Wimalaratne, S. M., Kothamachu, V. B., Rodriguez, N., Swat, M. J., Eils, J., Eils, R., Laibe, C., Malik-Sheriff, R. S., Chelliah, V., Le Novère, N., & Hermjakob, H. (2018). BioModels: expanding horizons to include more modelling approaches and formats. *Nucleic Acids Research*, *46*(D1), D1248–D1253. https://doi.org/10.1093/nar/gkx1023

Goelzer, A., Muntel, J., Chubukov, V., Jules, M., Prestel, E., Nölker, R., Mariadassou, M., Aymerich, S., Hecker, M., Noirot, P., Becher, D., & Fromion, V. (2015). Quantitative prediction of genome-wide resource allocation in bacteria. *Metabolic Engineering*, *32*, 232–243. https://doi.org/10.1016/j.ymben.2015.10.003

Gottig, N., Garavaglia, B. S., Garofalo, C. G., Orellano, E. G., & Ottado, J. (2009). A filamentous hemagglutinin-like protein of Xanthomonas axonopodis pv. citri, the phytopathogen responsible for citrus canker, is involved in bacterial virulence. *PLoS ONE*, *4*(2), e4358. https://doi.org/10.1371/journal.pone.0004358

Gough, C. L., Dow, J. M., Barber, C. E., & Daniels, M. J. (1988). Cloning of Two Endoglucanase Genes of Xanthomonas campestris pv. campestris : Analysis of the Role of the Major Endoglucanase in Pathogenesis. *Molecular Plant-Microbe Interactions*, *1*(7), 275. https://doi.org/10.1094/mpmi-1-275

Grafahrend-Belau, E., Junker, A., Eschenroder, A., Muller, J., Schreiber, F., & Junker, B. H. (2013). Multiscale Metabolic Modeling: Dynamic Flux Balance Analysis on a Whole-Plant Scale. *Plant Physiology*, *163*(2), 637–647. https://doi.org/10.1104/pp.113.224006

Hamilton, C. D., Steidl, O. R., MacIntyre, A. M., Hendrich, C. G., & Allen, C. (2021). Ralstonia solanacearum depends on catabolism of myo-inositol, sucrose, and trehalose for virulence in an infection stage-dependent manner. *Molecular Plant-Microbe Interactions*®, MPMI-10-20-0298-R. https://doi.org/10.1094/MPMI-10-20-0298-R

Hayward, A. C. (1991). Biology and Epidemiology of Bacterial Wilt Caused by Pseudomonas Solanacearum. *Annual Review of Phytopathology*, *29*(1), 65–87. https://doi.org/10.1146/annurev.py.29.090191.000433

Henry, C. S., Dejongh, M., Best, A. A., Frybarger, P. M., Linsay, B., & Stevens, R. L. (2010). High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nature Biotechnology*, *28*(9), 977–982. https://doi.org/10.1038/nbt.1672

Herms, D. A., & Mattson, W. J. (1992). The dilemma of plants: To grow or defend. *Quarterly Review of Biology*, *67*(3), 283–335. https://doi.org/10.1086/417659

Hopkins, D. L., & Purcell, A. H. (2002). Xylella fastidiosa: Cause of Pierce's Disease of Grapevine and Other Emergent Diseases. *Plant Disease*, *86*(10), 1056–1066. https://doi.org/10.1094/PDIS.2002.86.10.1056

Hugouvieux, V., Barber, C. E., & Daniels, M. J. (2007). Entry of Xanthomonas campestris pv. campestris into Hydathodes of Arabidopsis thaliana Leaves: A System for Studying Early Infection Events in Bacterial Pathogenesis. *Http://Dx.Doi.Org/10.1094/MPMI.1998.11.6.537*, *11*(6), 537–543. https://doi.org/10.1094/MPMI.1998.11.6.537

Ingel, B., Caldwell, D., Duong, F., Parkinson, D., McCulloh, K., Iyer-Pascuzzi, A. S., McElrone, A., & Lowe Power, T. (2021). Revisiting the source of wilt symptoms: X-ray microcomputed tomography provides direct evidence that Ralstonia biomass clogs xylem vessels. *PhytoFrontiers™*. https://doi.org/10.1094/PHYTOFR-06-21-0041-R

Ishiga, Y. (2017). Studies on mode of action of phytotoxin coronatine produced by Pseudomonas syringae pv. tomato. In *Journal of General Plant Pathology* (Vol. 83, Issue 6, pp. 424–426). Springer Tokyo.

https://doi.org/10.1007/s10327-017-0748-1

Jacobs, J. M., Babujee, L., Meng, F., Milling, A., & Allen, C. (2012). The in planta transcriptome of Ralstonia solanacearum: conserved physiological and virulence strategies during bacterial wilt of tomato. *MBio*, *3*(4), e00114-12. https://doi.org/10.1128/mBio.00114-12

Jha, G., Rajeshwari, R., & Sonti, R. V. (2007). Functional interplay between two Xanthomonas oryzae pv. oryzae secretion systems in modulating virulence on rice. *Molecular Plant-Microbe Interactions*, *20*(1), 31–40. https://doi.org/10.1094/MPMI-20-0031

Jyoti, P., Shree, M., Joshi, C., Prakash, T., Ray, S. K., Satapathy, S. S., & Masakapalli, S. K. (2020). The Entner-Doudoroff and Nonoxidative Pentose Phosphate Pathways Bypass Glycolysis and the Oxidative Pentose Phosphate Pathway in Ralstonia solanacearum. *MSystems*, *5*(2). https://doi.org/10.1128/msystems.00091-20

Kai, K., Ohnishi, H., Shimatani, M., Ishikawa, S., Mori, Y., Kiba, A., Ohnishi, K., Tabuchi, M., & Hikichi, Y. (2015). Methyl 3-Hydroxymyristate, a Diffusible Signal Mediating *phc* Quorum Sensing in *Ralstonia solanacearum*. *ChemBioChem*, *16*(16), 2309–2318. https://doi.org/10.1002/cbic.201500456

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research*, *45*(D1), D353–D361. https://doi.org/10.1093/nar/gkw1092

Kanwar, P., & Jha, G. (2018). Alterations in plant sugar metabolism: signatory of pathogen attack. *Planta*, *2018*. https://doi.org/10.1007/s00425-018-3018-3

Kelman, A. (1950). Influence of nitrogen nutrition on the development of bacterial wilt in tomato and tobacco. *Phytopathology*, *40*, 14.

Khokhani, D., Lowe-Power, T. M., Tran, T. M., & Allen, C. (2017). A Single Regulator Mediates Strategic Switching between Attachment/Spread and Growth/Virulence in the Plant Pathogen Ralstonia solanacearum. *MBio*, *8*(5), e00895-17. https://doi.org/10.1128/mBio.00895-17

Killiny, N., & Almeida, R. P. P. (2014). Factors affecting the initial adhesion and retention of the plant pathogen xylella fastidiosa in the foregut of an insect vector. *Applied and Environmental Microbiology*, *80*(1), 420–426. https://doi.org/10.1128/AEM.03156-13

King, Z. A., Lu, J., Dräger, A., Miller, P., Federowicz, S., Lerman, J. A., Ebrahim, A., Palsson, B. O., & Lewis, N. E. (2016). BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Research*, *44*(D1), D515–D522. https://doi.org/10.1093/nar/gkv1049

Koduru, L., Kim, H. Y., Lakshmanan, M., Mohanty, B., Lee, Y. Q., Lee, C. H., & Lee, D. (2020). Genome-scale metabolic reconstruction and in silico analysis of the rice leaf blight pathogen, Xanthomonas oryzae. *Molecular Plant Pathology*, mpp.12914. https://doi.org/10.1111/mpp.12914

Kwaik, Y. A., & Bumann, D. (2013). Microbial quest for food in vivo: 'Nutritional virulence' as an emerging paradigm. *Cellular Microbiology*, *15*(6), 882–890. https://doi.org/10.1111/CMI.12138

Lakshmanan, M., Lim, S. H., Mohanty, B., Kim, J. K., Ha, S. H., & Lee, D. Y. (2015). Unraveling the light-specific metabolic and regulatory signatures of rice through combined in silico modeling and multiomics analysis. *Plant Physiology*, *169*(4), 3002–3020. https://doi.org/10.1104/pp.15.01379

Landry, D., González-Fuente, M., Deslandes, L., & Peeters, N. (2020). The large, diverse, and robust arsenal of Ralstonia solanacearum type III effectors and their in planta functions. *Molecular Plant Pathology*, *April*, 1–12. https://doi.org/10.1111/mpp.12977

Larhlimi, A., David, L., Selbig, J., & Bockmayr, A. (2012). F2C2: a fast tool for the computation of flux coupling in

genome-scale metabolic networks. *BMC Bioinformatics*, *13*(1), 1–9. https://doi.org/10.1186/1471-2105-13-57

Leite, B., Andersen, P. C., & Ishida, M. L. (2004). Colony aggregation and biofilm formation in xylem chemistry-based media for Xylella fastidiosa. *FEMS Microbiology Letters*, *230*(2), 283–290. https://doi.org/10.1016/S0378-1097(03)00917-0

Liu, H., Zhang, S., Schell, M. A., & Denny, T. P. (2005). Pyramiding unmarked deletions in Ralstonia solanacearum shows that secreted proteins in addition to plant cell-wall-degrading enzymes contribute to virulence. *Molecular Plant-Microbe Interactions*, *18*(12), 1296–1305. https://doi.org/10.1094/MPMI-18-1296

Lombardot, T., Morgat, A., Axelsen, K. B., Aimo, L., Hyka-Nouspikel, N., Niknejad, A., Ignatchenko, A., Xenarios, I., Coudert, E., Redaschi, N., & Bridge, A. (2019). Updates in Rhea: SPARQLing biochemical reaction data. *Nucleic Acids Research*, *47*(D1), D596–D600. https://doi.org/10.1093/nar/gky876

Lowe-Power, T. M., Hendrich, C. G., von Roepenack-Lahaye, E., Li, B., Wu, D., Mitra, R., Dalsing, B. L., Ricca, P., Naidoo, J., Cook, D., Jancewicz, A., Masson, P., Thomma, B., Lahaye, T., Michael, A. J., & Allen, C. (2018). Metabolomics of tomato xylem sap during bacterial wilt reveals *Ralstonia solanacearum* produces abundant putrescine, a metabolite that accelerates wilt disease. *Environmental Microbiology*, *20*(4), 1330–1349. https://doi.org/10.1111/1462-2920.14020

Lowe-Power, T. M., Khokhani, D., & Allen, C. (2018). How Ralstonia solanacearum Exploits and Thrives in the Flowing Plant Xylem Environment. *Trends in Microbiology*, *26*(11), 929–942. https://doi.org/10.1016/j.tim.2018.06.002

Lucas, J. A. (2020). *Plant Pathology and Plant Pathogens* (4th Editio). John Wiley & Sons, Ltd.

Mahadevan, R., & Schilling, C. H. (2003). The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metabolic Engineering*, *5*(4), 264–276. https://doi.org/10.1016/J.YMBEN.2003.09.002

Mahadevan, Radhakrishnan, Edwards, J. S., & Doyle, F. J. (2002). Dynamic Flux Balance Analysis of diauxic growth in Escherichia coli. *Biophysical Journal*, *83*(3), 1331–1340. https://doi.org/10.1016/S0006-3495(02)73903-9

Mahout, M., Carlson, R. P., & Peres, S. (2020). Answer Set Programming for Computing Constraints-Based Elementary Flux Modes: Application to Escherichia coli Core Metabolism. *Processes 2020, Vol. 8, Page 1649*, *8*(12), 1649. https://doi.org/10.3390/PR8121649

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M. A. X., Verdier, V., Beer, S. V, Machado, M. A., Toth, I. A. N., Salmond, G., & Foster, G. D. (2012). Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular Plant Pathology*, *13*(6), 614–629. https://doi.org/10.1111/J.1364-3703.2012.00804.X

Marcelletti, S., & Scortichini, M. (2016). Genome-wide comparison and taxonomic relatedness of multiple Xylella fastidiosa strains reveal the occurrence of three subspecies and a new Xylella species. *Archives of Microbiology*, *198*(8), 803–812. https://doi.org/10.1007/s00203-016-1245-1

Marguerettaz, M., Pieretti, I., Gayral, P., Puig, J., Brin, C., Cociancich, S., Poussier, S., Rott, P., & Royer, M. (2011). Genomic and Evolutionary Features of the SPI-1 Type III Secretion System That Is Present in Xanthomonas albilineans but Is Not Essential for Xylem Colonization and Symptom Development of Sugarcane Leaf Scald. *Http://Dx.Doi.Org/10.1094/MPMI-08-10-0188*, *24*(2), 246–259. https://doi.org/10.1094/MPMI-08-10-0188

Meng, Y., Li, Y., Galvani, C. D., Hao, G., Turner, J. N., Burr, T. J., & Hoch, H. C. (2005). Upstream migration of Xylella fastidiosa via pilus-driven twitching motility. *Journal of Bacteriology*, *187*(16), 5560–5567.

https://doi.org/10.1128/JB.187.16.5560-5567.2005

Mensi, I., Vernerey, M.-S., Gargani, D., Nicole, M., & Rott, P. (2013). Breaking dogmas: the plant vascular pathogen Xanthomonas albilineans is able to invade non-vascular tissues despite its reduced genome. *Open Biology*, *4*(FEB). https://doi.org/10.1098/RSOB.130116

Merda, D., Briand, M., Bosis, E., Rousseau, C., Portier, P., Barret, M., Jacques, M. A., & Fischer-Le Saux, M. (2017). Ancestral acquisitions, gene flow and multiple evolutionary trajectories of the type three secretion system and effectors in Xanthomonas plant pathogens. *Molecular Ecology*, *26*(21), 5939–5952. https://doi.org/10.1111/MEC.14343

Méthot, P. O., & Alizon, S. (2014). What is a pathogen? Toward a process view of host-parasite interactions. *Virulence*, *5*(8), 775–785. https://doi.org/10.4161/21505594.2014.960726

Mintz-Oron, S., Meir, S., Malitsky, S., Ruppin, E., Aharoni, A., & Shlomi, T. (2012). Reconstruction of Arabidopsis metabolic network models accounting for subcellular compartmentalization and tissue-specificity. *Proceedings of the National Academy of Sciences*, *109*(1), 339–344. https://doi.org/10.1073/pnas.1100358109

Monk, J. M., Lloyd, C. J., Brunk, E., Mih, N., Sastry, A., King, Z., Takeuchi, R., Nomura, W., Zhang, Z., Mori, H., Feist, A. M., & Palsson, B. O. (2017). i ML1515, a knowledgebase that computes Escherichia coli traits. *Nature Biotechnology 2017 35:10*, *35*(10), 904–908. https://doi.org/10.1038/nbt.3956

Moreira, T. B., Shaw, R., Luo, X., Ganguly, O., Kim, H. S., Coelho, L. G. F., Cheung, C. Y. M., & Williams, T. C. R. (2019). A genome-scale metabolic model of soybean (Glycine max) highlights metabolic fluxes in seedlings. *Plant Physiology*, *180*(4), 1912–1929. https://doi.org/10.1104/pp.19.00122

Moretti, S., Tran, V. D. T., Mehl, F., Ibberson, M., & Pagni, M. (2021). MetaNetX/MNXref: Unified namespace for metabolites and biochemical reactions in the context of metabolic models. *Nucleic Acids Research*, *49*(D1), D570–D574. https://doi.org/10.1093/nar/gkaa992

Myburg, A. A., Lev-Yadun, S., & Sederoff, R. R. (2013). Xylem Structure and Function. In *eLS*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470015902.a0001302.pub2

Nakaho, K., Inoue, H., Takayama, T., & Miyagawa, H. (2004). Distribution and multiplication of Ralstonia solanacearum in tomato plants with resistance derived from different origins. *Journal of General Plant Pathology*, *70*(2), 115–119. https://doi.org/10.1007/s10327-003-0097-0

Nascimento, R., Gouran, H., Chakraborty, S., Gillespie, H. W., Almeida-Souza, H. O., Tu, A., Rao, B. J., Feldstein, P. A., Bruening, G., Goulart, L. R., & Dandekar, A. M. (2016). The Type II Secreted Lipase/Esterase LesA is a Key Virulence Factor Required for Xylella fastidiosa Pathogenesis in Grapevines. *Scientific Reports*, *6*(January), 1–17. https://doi.org/10.1038/srep18598

Newman, K. L., Almeida, R. P. P., Purcell, A. H., & Lindow, S. E. (2003). Use of a Green Fluorescent Strain for Analysis of *Xylella fastidiosa* Colonization of *Vitis vinifera*. *Applied and Environmental Microbiology*, *69*(12), 7319–7327. https://doi.org/10.1128/AEM.69.12.7319-7327.2003

Niño-Liu, D. O., Ronald, P. C., & Bogdanove, A. J. (2006). Xanthomonas oryzae pathovars: Model pathogens of a model crop. In *Molecular Plant Pathology* (Vol. 7, Issue 5, pp. 303–324). John Wiley & Sons, Ltd. https://doi.org/10.1111/j.1364-3703.2006.00344.x

Orth, J. D., Conrad, T. M., Na, J., Lerman, J. A., Nam, H., Feist, A. M., & Palsson, B. Ø. (2011). A comprehensive genome-scale reconstruction of Escherichia coli metabolism--2011. *Molecular Systems Biology*, *7*(1), 535. https://doi.org/10.1038/msb.2011.65

Orth, J. D., Thiele, I., & Palsson, B. Ø. (2010). What is flux balance analysis? *Nat Biotechnol*, *28*(3), 245–248.

https://doi.org/10.1038/nbt.1614.What

Pál, C., Papp, B., & Lercher, M. J. (2005). Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nature Genetics*, *37*(12), 1372–1375. https://doi.org/10.1038/ng1686

Peyraud, R., Cottret, L., Marmiesse, L., & Genin, S. (2018). Control of primary metabolism by a virulence regulatory network promotes robustness in a plant pathogen. *Nature Communications*, *9*(1), 418. https://doi.org/10.1038/s41467-017-02660-4

Peyraud, R., Cottret, L., Marmiesse, L., Gouzy, J., & Genin, S. (2016). A Resource Allocation Trade-Off between Virulence and Proliferation Drives Metabolic Versatility in the Plant Pathogen Ralstonia solanacearum. *PLoS Pathogens*, *12*(10), e1005939. https://doi.org/10.1371/journal.ppat.1005939

Peyraud, R., Mbengue, M., Barbacci, A., & Raffaele, S. (2019). Intercellular cooperation in a fungal plant pathogen facilitates host colonization. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(8), 3193–3201. https://doi.org/10.1073/pnas.1811267116

Pfau, T., Christian, N., Masakapalli, S. K., Sweetlove, L. J., Poolman, M. G., & Ebenhöh, O. (2018). The intertwined metabolism during symbiotic nitrogen fixation elucidated by metabolic modelling. *Scientific Reports*, *8*(1), 12504. https://doi.org/10.1038/s41598-018-30884-x

Pieterse, C. M. J., Van Der Does, D., Zamioudis, C., Leon-Reyes, A., & Van Wees, S. C. M. (2012). Hormonal modulation of plant immunity. *Annual Review of Cell and Developmental Biology*, *28*, 489–521. https://doi.org/10.1146/annurev-cellbio-092910-154055

Plener, L., Boistard, P., González, A., Boucher, C., & Genin, S. (2012). Metabolic adaptation of Ralstonia solanacearum during plant infection: A methionine biosynthesis case study. *PLoS ONE*, *7*(5), e36877. https://doi.org/10.1371/journal.pone.0036877

PM 7/24 (2) *Xylella fastidiosa*. (2016). *EPPO Bulletin*, *46*(3), 463–500. https://doi.org/10.1111/epp.12327

Ponzio, C., Weldegergis, B. T., Dicke, M., & Gols, R. (2016). Compatible and incompatible pathogen–plant interactions differentially affect plant volatile emissions and the attraction of parasitoid wasps. *Functional Ecology*, *30*(11), 1779–1789. https://doi.org/10.1111/1365-2435.12689

Poolman, M. G., Kundu, S., Shaw, R., & Fell, D. A. (2013). Responses to light intensity in a genome-scale model of rice metabolism. *Plant Physiology*, *162*(2), 1060–1072. https://doi.org/10.1104/pp.113.216762

Poolman, M. G., Miguet, L., Sweetlove, L. J., & Fell, D. A. (2009). A genome-scale metabolic model of Arabidopsis and some of its properties. *Plant Physiology*, *151*(3), 1570–1581. https://doi.org/10.1104/pp.109.141267

Poueymiro, M., Cazalé, A. C., François, J. M., Parrou, J. L., Peeters, N., & Genin, S. (2014). A Ralstonia solanacearum type III effector directs the production of the plant signal metabolite trehalose-6-phosphate. *MBio*, *5*(6), 1–9. https://doi.org/10.1128/mBio.02065-14

Rapicavoli, J., Ingel, B., Blanco-Ulate, B., Cantu, D., & Roper, C. (2018). Xylella fastidiosa: an examination of a re-emerging plant pathogen. *Molecular Plant Pathology*, *19*(4), 786–800. https://doi.org/10.1111/mpp.12585

Rapicavoli, J. N., Blanco-Ulate, B., Muszyński, A., Figueroa-Balderas, R., Morales-Cruz, A., Azadi, P., Dobruchowska, J. M., Castro, C., Cantu, D., & Roper, M. C. (2018). Lipopolysaccharide O-antigen delays plant innate immune recognition of Xylella fastidiosa. *Nature Communications*, *9*, 390. https://doi.org/10.1038/s41467-018-02861-5

Ray, S. K., Rajeshwari, R., Sharma, Y., & Sonti, R. V. (2002). A high-molecular-weight outer membrane protein of Xanthomonas oryzae pv.oryzae exhibits similarity to non-fimbrial adhesins of animal pathogenic bacteria and is required for optimum virulence. *Molecular Microbiology*, *46*(3), 637–647. https://doi.org/10.1046/j.1365-2958.2002.03188.x

Remenant, B., Cambiaire, J.-C. de, Cellier, G., Jacobs, J. M., Mangenot, S., Barbe, V., Lajus, A., Vallenet, D., Medigue, C., Fegan, M., Allen, C., & Prior, P. (2011). Ralstonia syzygii, the Blood Disease Bacterium and Some Asian R. solanacearum Strains Form a Single Genomic Species Despite Divergent Lifestyles. *PLOS ONE*, *6*(9), e24356. https://doi.org/10.1371/JOURNAL.PONE.0024356

Rodenburg, S. Y. A., Seidl, M. F., de Ridder, D., & Govers, F. (2018). Genome-wide characterization of Phytophthora infestans metabolism: a systems biology approach. *Molecular Plant Pathology*, *19*(6), 1403–1413. https://doi.org/10.1111/mpp.12623

Rodenburg, S. Y. A., Seidl, M. F., Judelson, H. S., Vu, A. L., Govers, F., & De Ridder, D. (2019). Metabolic model of the Phytophthora infestans-tomato interaction reveals metabolic switches during host colonization. *MBio*, *10*(4). https://doi.org/10.1128/mBio.00454-19

Roper, C., Castro, C., & Ingel, B. (2019). Xylella fastidiosa: bacterial parasitism with hallmarks of commensalism. *Current Opinion in Plant Biology*, *50*, 140–147. https://doi.org/10.1016/j.pbi.2019.05.005

Roux, B., Bolot, S., Guy, E., Denancé, N., Lautier, M., Jardinaud, M.-F., Fischer-Le Saux, M., Portier, P., Jacques, M.-A., Gagnevin, L., Pruvost, O., Lauber, E., Arlat, M., Carrère, S., Koebnik, R., & Noël, L. D. (2015). Genomics and transcriptomics of Xanthomonas campestris species challenge the concept of core type III effectome. *BMC Genomics 2015 16:1*, *16*(1), 1–19. https://doi.org/10.1186/S12864-015-2190-0

Ryan, R. P., Vorhölter, F.-J., Potnis, N., Jones, J. B., Van Sluys, M.-A., Bogdanove, A. J., & Dow, J. M. (2011). Pathogenomics of Xanthomonas: understanding bacterium–plant interactions. *Nature Reviews Microbiology 2011 9:5*, *9*(5), 344–355. https://doi.org/10.1038/nrmicro2558

Sabbagh, C. R. R., Carrere, S., Lonjon, F., Vailleau, F., Macho, A. P., Genin, S., & Peeters, N. (2019). Pangenomic type III effector database of the plant pathogenic Ralstonia spp. *PeerJ*, *7*(8), e7346. https://doi.org/10.7717/PEERJ.7346

Safni, I., Cleenwerck, I., De Vos, P., Fegan, M., Sly, L., & Kappler, U. (2014). Polyphasic taxonomic revision of the Ralstonia solanacearum species complex: Proposal to emend the descriptions of Ralstonia solanacearum and Ralstonia syzygii and reclassify current R. syzygii strains as Ralstonia syzygii subsp. syzygii subsp. nov., R. s. *International Journal of Systematic and Evolutionary Microbiology*, *64*, 3087–3103. https://doi.org/10.1099/ijs.0.066712-0

Salanoubat, M., Genin, S., Artiguenave, F., Gouzy, J., Mangenot, S., Arlat, M., Billault, A., Brottiert, P., Camus, J. C., Cattolico, L., Chandler, M., Choisne, N., Claudel-Renard, C., Cunnac, S., Demange, N., Gaspin, C., Lavie, M., Moisan, A., Robert, C., … Boucher, C. A. (2002). Genome sequence of the plant pathogen Ralstonia solanacearum. *Nature*, *415*(6871), 497–502. https://doi.org/10.1038/415497a

Salimi, F., Zhuang, K., & Mahadevan, R. (2010). Genome-scale metabolic modeling of a clostridial co-culture for consolidated bioprocessing. *Biotechnology Journal*, *5*(7), 726–738. https://doi.org/10.1002/biot.201000159

Samal, B., & Chatterjee, S. (2019). New insight into bacterial social communication in natural host: Evidence for interplay of heterogeneous and unison quorum response. *PLOS Genetics*, *15*(9), e1008395. https://doi.org/10.1371/JOURNAL.PGEN.1008395

Scally, M., Schuenzel, E. L., Stouthamer, R., & Nunney, L. (2005). Multilocus Sequence Type System for the Plant Pathogen *Xylella fastidiosa* and Relative Contributions of Recombination and Point Mutation to Clonal Diversity. *Applied and Environmental Microbiology*, *71*(12), 8491–8499. https://doi.org/10.1128/AEM.71.12.8491-8499.2005

Schatschneider, S., Huber, C., Neuweger, H., Watt, T. F., Pühler, A., Eisenreich, W., Wittmann, C., Niehaus, K., & Vorhölter, F. J. (2014). Metabolic flux pattern of glucose utilization by Xanthomonas campestris pv.

campestris: Prevalent role of the Entner-Doudoroff pathway and minor fluxes through the pentose phosphate pathway and glycolysis. *Molecular BioSystems*, *10*(10), 2663–2676. https://doi.org/10.1039/c4mb00198b

Schatschneider, S., Persicke, M., Watt, S. A., Hublik, G., Pühler, A., Niehaus, K., & Vorhölter, F. J. (2013). Establishment, in silico analysis, and experimental verification of a large-scale metabolic network of the xanthan producing Xanthomonas campestris pv. campestris strain B100. *Journal of Biotechnology*, *167*(2), 123–134. https://doi.org/10.1016/j.jbiotec.2013.01.023

Scheunemann, M., Brady, S. M., & Nikoloski, Z. (2018). Integration of large-scale data for extraction of integrated Arabidopsis root cell-type specific models. *Scientific Reports*, *8*(1), 1–15. https://doi.org/10.1038/s41598-018-26232-8

Schuenzel, E. L., Scally, M., Stouthamer, R., & Nunney, L. (2005). A Multigene Phylogenetic Study of Clonal Diversity and Divergence in North American Strains of the Plant Pathogen *Xylella fastidiosa*. *Applied and Environmental Microbiology*, *71*(7), 3832–3839. https://doi.org/10.1128/AEM.71.7.3832-3839.2005

Schultz, T. (1986). Xanthomonas campestris pv. campestris in Western Washington Crucifer Seed Fields: Occurrence and Survival . *Phytopathology*, *76*(12), 1306. https://doi.org/10.1094/PHYTO-76-1306

Schumann, G. L., & D'Arcy, C. J. (2017). CHAPTER 1: The Irish Potato Famine: The Birth of Plant Pathology. In *Hungry Planet: Stories of Plant Diseases* (pp. 1–19). The American Phytopathological Society. https://doi.org/10.1094/9780890544907.001

Shaw, R., & Cheung, C. Y. M. (2018). A Dynamic Multi-Tissue Flux Balance Model Captures Carbon and Nitrogen Metabolism and Optimal Resource Partitioning During Arabidopsis Growth. *Frontiers in Plant Science*, *9*, 884. https://doi.org/10.3389/fpls.2018.00884

Shaw, R., & Cheung, C. Y. M. (2019). A mass and charge balanced metabolic model of Setaria viridis revealed mechanisms of proton balancing in C4 plants. *BMC Bioinformatics*, *20*(1), 1–11. https://doi.org/10.1186/s12859-019-2941-z

Shlomi, T., Benyamini, T., Gottlieb, E., Sharan, R., & Ruppin, E. (2011). Genome-Scale Metabolic Modeling Elucidates the Role of Proliferative Adaptation in Causing the Warburg Effect. *PLOS Computational Biology*, *7*(3), e1002018. https://doi.org/10.1371/JOURNAL.PCBI.1002018

Simpson, A. J. G., Reinach, F. C., Arruda, P., Abreu, F. A., Acencio, M., Alvarenga, R., Alves, L. M. C., Araya, J. E., Baia, G. S., Baptista, C. S., Barros, M. H., Bonaccorsi, E. D., Bordin, S., Bové, J. M., Briones, M. R. S., Bueno, M. R. P., Camargo, A. A., Camargo, L. E. A., Carraro, D. M., … Setubal, J. C. (2000). The genome sequence of the plant pathogen Xylella fastidiosa. *Nature*, *406*, 151–157. https://doi.org/10.1038/35018003

Song, H. S., & Ramkrishna, D. (2009). When is the quasi-steady-state approximation admissible in metabolic modeling? When admissible, what models are desirable? *Industrial and Engineering Chemistry Research*, *48*(17), 7976–7985. https://doi.org/10.1021/ie900075f

Song, H. S., & Ramkrishna, D. (2010). Prediction of metabolic function from limited data: Lumped Hybrid Cybernetic Modeling (L-HCM). *Biotechnology and Bioengineering*, *106*(2), 271–284. https://doi.org/10.1002/bit.22692

Stolyar, S., Van Dien, S., Hillesland, K. L., Pinel, N., Lie, T. J., Leigh, J. A., & Stahl, D. A. (2007). Metabolic modeling of a mutualistic microbial community. *Molecular Systems Biology*, *3*(92), 1–14. https://doi.org/10.1038/msb4100131

Sun, Q., Sun, Y., Walker, M. A., & Labavitch, J. M. (2013). Vascular Occlusions in Grapevines with Pierce's Disease Make Disease Symptom Development Worse. *Plant Physiology*, *161*(3), 1529–1541. https://doi.org/10.1104/PP.112.208157

Taiz, L., & Zeiger, E. (2006). *Plant physiology*. Sinauer Associates.

Thiele, I., & Palsson, B. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols*, *5*(1), 93–121. https://doi.org/10.1038/nprot.2009.203

Thomma, B. P. H. J., Nürnberger, T., & Joosten, M. H. A. J. (2011). Of PAMPs and effectors: The blurred PTI-ETI dichotomy. *Plant Cell*, *23*(1), 4–15. https://doi.org/10.1105/tpc.110.082602

Timilsina, S., Potnis, N., Newberry, E. A., Liyanapathiranage, P., Iruegas-Bocardo, F., White, F. F., Goss, E. M., & Jones, J. B. (2020). Xanthomonas diversity, virulence and plant–pathogen interactions. In *Nature Reviews Microbiology* (Vol. 18, Issue 8, pp. 415–427). Nature Research. https://doi.org/10.1038/s41579-020-0361-8

Trinh, C. T., Li, J., Blanch, H. W., & Clark, D. S. (2011). Redesigning Escherichia coli metabolism for anaerobic production of isobutanol. *Applied and Environmental Microbiology*, *77*(14), 4894–4904. https://doi.org/10.1128/AEM.00382-11

Turner, J. G., & Taha, R. R. (1984). Contribution of tabtoxin to the pathogenicity of Pseudomonas syringae pv. tabaci. *Physiological Plant Pathology*, *25*(1), 55–69. https://doi.org/10.1016/0048-4059(84)90017-1

Tzfira, T., & Citovsky, V. (2002). Partners-in-infection: Host proteins involved in the transformation of plant cells by Agrobacterium. In *Trends in Cell Biology* (Vol. 12, Issue 3, pp. 121–129). Elsevier Current Trends. https://doi.org/10.1016/S0962-8924(01)02229-2

Uchytil, T. F., & Durbin, R. D. (1980). Hydrolysis of tabtoxins by plant and bacterial enzymes. *Experientia*, *36*(3), 301–302. https://doi.org/10.1007/BF01952288

Unrean, P., Trinh, C. T., & Srienc, F. (2010). Rational design and construction of an efficient E. coli for production of diapolycopendioic acid. *Metabolic Engineering*, *12*(2), 112–122. https://doi.org/10.1016/J.YMBEN.2009.11.002

Valls, M., Genin, S., & Boucher, C. (2006). Integrated regulation of the type III secretion system and other virulence determinants in Ralstonia solanacearum. *PLoS Pathogens*, *2*(8), 0798–0807. https://doi.org/10.1371/journal.ppat.0020082

Vasse, J., Frey, P., & Trigalet, A. (1995). Microscopic studies of intercellular infection and protoxylem invasion of tomato roots by Pseudomonas solanacearum. *Molecular Plant-Microbe Interactions*, *8*(2), 241–251. https://doi.org/10.1094/MPMI-8-0241

Vicente, J. G., & Holub, E. B. (2013). Xanthomonas campestris pv. campestris (cause of black rot of crucifers) in the genomic era is still a worldwide threat to brassica crops. *Molecular Plant Pathology*, *14*(1), 2–18. https://doi.org/10.1111/J.1364-3703.2012.00833.X

Waldherr, S., Oyarzún, D. A., & Bockmayr, A. (2015). Dynamic optimization of metabolic networks coupled with gene expression. *Journal of Theoretical Biology*, *365*, 469–485. https://doi.org/10.1016/J.JTBI.2014.10.035

Walters, D., & Heil, M. (2007). Costs and trade-offs associated with induced resistance. In *Physiological and Molecular Plant Pathology* (Vol. 71, Issues 1–3, pp. 3–17). Academic Press. https://doi.org/10.1016/j.pmpp.2007.09.008

Wang, C., Deng, Z.-L., Xie, Z.-M., Chu, X.-Y., Chang, J.-W., Kong, D.-X., Li, B.-J., Zhang, H.-Y., & Chen, L.-L. (2015). Construction of a genome-scale metabolic network of the plant pathogen Pectobacterium carotovorum provides new strategies for bactericide discovery. *FEBS Letters*, *589*(3), 285–294. https://doi.org/10.1016/J.FEBSLET.2014.12.010

Wang, G., Kong, J., Cui, D., Zhao, H., Niu, Y., Xu, M., Jiang, G., Zhao, Y., & Wang, W. (2019). Resistance against Ralstonia solanacearum in tomato depends on the methionine cycle and the γ-aminobutyric acid metabolic pathway. *The Plant Journal*, *97*(6), 1032–1047. https://doi.org/10.1111/tpj.14175

Wang, Y., Zhao, A., Morcillo, R. J. L., Yu, G., Xue, H., Rufian, J. S., Sang, Y., & Macho, A. P. (2021). A bacterial effector protein uncovers a plant metabolic pathway involved in tolerance to bacterial wilt disease. *Molecular Plant*. https://doi.org/10.1016/j.molp.2021.04.014

Wells, J. M., Raju, B. C., Hung, H., Weisburg, W. G., Mandelco-paul, L., & Brenner, D. O. N. J. (1987). *Xylella fastidiosa* gen. nov., sp. nov: Gram-Negative, Xylem-Limited, Fastidious Plant Bacteria Related to *Xanthomonas* spp. *International Journal of Systematic Bacteriology*, *37*(2), 136–143.

Windt, C. W., Vergeldt, F. J., De Jager, P. A., & Van As, H. (2006). MRI of long-distance water transport: a comparison of the phloem and xylem flow characteristics and dynamics in poplar, castor bean, tomato and tobacco. *Plant, Cell and Environment*, *29*(9), 1715–1729. https://doi.org/10.1111/j.1365-3040.2006.01544.x

Wu, D., von Roepenack-Lahaye, E., Buntru, M., de Lange, O., Schandry, N., Pérez-Quintero, A. L., Weinberg, Z., Lowe-Power, T. M., Szurek, B., Michael, A. J., Allen, C., Schillberg, S., & Lahaye, T. (2019). A Plant Pathogen Type III Effector Protein Subverts Translational Regulation to Boost Host Polyamine Levels. *Cell Host and Microbe*, *26*(5), 638-649.e5. https://doi.org/10.1016/j.chom.2019.09.014

Xian, L., Yu, G., Wei, Y., Rufian, J. S., Li, Y., Zhuang, H., Xue, H., Morcillo, R. J. L., & Macho, A. P. (2020). A Bacterial Effector Protein Hijacks Plant Metabolism to Support Pathogen Nutrition. *Cell Host & Microbe*, 1–10. https://doi.org/10.1016/j.chom.2020.07.003

Yadeta, K. A., & Thomma, B. P. H. J. (2013). The xylem as battleground for plant hosts and vascular wilt pathogens. In *Frontiers in Plant Science* (Vol. 4, Issue APR, p. 97). Frontiers Research Foundation. https://doi.org/10.3389/fpls.2013.00097

Yu, Q., Alvarez, A. M., Moore, P. H., Zee, F., Kim, M. S., Silva, A. de, Hepperly, P. R., & Ming, R. (2007). Molecular Diversity of Ralstonia solanacearum Isolated from Ginger in Hawaii. *Http://Dx.Doi.Org/10.1094/PHYTO.2003.93.9.1124*, *93*(9), 1124–1130. https://doi.org/10.1094/PHYTO.2003.93.9.1124

Yuan, H., Cheung, C. Y. M., Poolman, M. G., Hilbers, P. A. J., & van Riel, N. A. W. (2016). A genome-scale metabolic network reconstruction of tomato (Solanum lycopersicum L.) and its application to photorespiratory metabolism. *The Plant Journal*, *85*(2), 289–304. https://doi.org/10.1111/tpj.13075

Yuan, M., Ngou, B. P. M., Ding, P., & Xin, X. F. (2021). PTI-ETI crosstalk: an integrative view of plant immunity. *Current Opinion in Plant Biology*, *62*, 102030. https://doi.org/10.1016/j.pbi.2021.102030

Zanghellini, J., Ruckerbauer, D. E., Hanscho, M., & Jungreuthmayer, C. (2013). Elementary flux modes in a nutshell: Properties, calculation and applications. In *Biotechnology Journal* (Vol. 8, Issue 9, pp. 1009–1016). John Wiley & Sons, Ltd. https://doi.org/10.1002/biot.201200269

Zoledowska, S., Presta, L., Fondi, M., Decorosi, F., Giovannetti, L., Mengoni, A., & Lojkowska, E. (2019). Metabolic Modeling of Pectobacterium parmentieri SCC3193 Provides Insights into Metabolic Pathways of Plant Pathogenic Bacteria. *Microorganisms*, *7*(4), 101. https://doi.org/10.3390/microorganisms7040101

# Appendices

# Appendix I - Plant genome-scale metabolic networks

**Article**

Plant genome-scale metabolic networks

**Journal**

*Advances in Botanical Research* (published online 27 October 2020)

**Contributions**

- Writing the first draft of the section “Metabolic modeling of plant-microorganism interactions”
- Designing Figure 1
- Reviewing and editing the complete article with the co-authors

CHAPTER EIGHT

# Plant genome-scale metabolic networks

**Léo Gerlin[a], Clément Frainay[b], Fabien Jourdan[b], Caroline Baroukh[a], and Sylvain Prigent[c,*]**

[a]LIPM, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France
[b]Toxalim (Research Center in Food Toxicology), Université de Toulouse, INRAE, ENVT, INP-Purpan, UPS, Toulouse, France
[c]UMR1332, Université de Bordeaux, INRAE, Bordeaux Metabolome, MetaboHUB, PHENOME, Villenave d'Ornon, France
*Corresponding author: e-mail address: sylvain.prigent@inrae.fr

## Contents



## Abstract

With the development of plant metabolomics, projects like the 10,000 plants genome sequencing project and the growth of other omics, the amount of data describing plant metabolism has never been so big. Genome-scale metabolic models (GEMs) are widely used to integrate and study the available information, and to better understand global responses to metabolic changes.

In the first part of this chapter, we will focus on the existing plant GEMs, their history, and the biological questions associated with their reconstruction. Those GEMs were initially reconstructed based on cell suspension systems, but current reconstructions are widely focused on multiorgans and multitissues models, requiring some model integration.

Model integration is the focus of the second part of this chapter, with the study of biotic interactions with the integration of GEMs coming from plant-interacting microorganisms and their host plants, modeling either pathogenic and symbiotic interactions. Emphasis will be placed on modeling quantitative interaction between plants and microorganisms.

*Advances in Botanical Research*, Volume 98
ISSN 0065-2296
https://doi.org/10.1016/bs.abr.2020.09.021

© 2021 Elsevier Ltd
All rights reserved.

Finally, the third part of this chapter describes how the visualization of networks could be used to improve the understanding of the models and the integration of different omics data, and how good representation of metabolic networks can help the user to perform diverse tasks and put knowledge in context.

## 1. An introduction on existing genome-scale metabolic models

Genome-scale metabolic networks provide a representation of all metabolic actors present in the cell, by describing gene-protein-reaction associations, based on genome annotations and experimental data (Thiele & Palsson, 2010). Such networks contain sets of metabolic reactions, ideally mass-balanced, that can be transformed into mathematical models, the genome-scale metabolic models (GEMs) that allow for modeling of complex behavior. Hence, GEMs are of great interest to study the metabolism of species possessing complex metabolic pathways, and plants are such species. It is therefore not surprising that plant GEMs are being developed quite extensively in recent years.

This development is driven by the rise of the omics era in plants, making it easier to obtain genomic, transcriptomic, proteomic and metabolomic data during the past few years. The genome sequence of *Arabidopsis thaliana* was released in 2000 (The Arabidopsis Genome Initiative, 2000), and 3 years later was published the first version of AraCyc, a database containing more than 170 biochemical pathways linked with *A. thaliana*. It was not before 2009 that was published the first curated plant GEM (Poolman, Miguet, Sweetlove, & Fell, 2009), enabling the simulation of growth and biomass formation of a heterotrophic cell suspension system. Since this first publication, a number of different GEMs has been published for a variety of species. We identified 30 different models reconstructed for 12 different plant species (Table 1). Given the fact that *A. thaliana* remains a model organism for plant metabolism, it is not surprising that GEMs reconstructed to study different metabolic features of this species represent more than one-third of identified plant GEMs. Quite interestingly, reconstructed plant GEMs cover a wide array of phylogenetic clades, with species from monocots to eudicots clades. At least one representative of Poaceae, Brassicaceae, Solanaceae, Fabaceae, Lamiaceae and Salicaceae families are present. This diversity of species could be very interesting, in particular to compare metabolic features for the identification of conserved or diverse patterns across species. It is interesting to note that most of the reconstructed models initially come from two data

**Table 1** Genome-scale metabolic models of plant species.

| Species | Classification | Publication | Reactions | Metabolites | Genes | Applications |
|---|---|---|---|---|---|---|
| *Arabidopsis thaliana* | Eudicot, Rosids, Brassicaceae | Poolman et al. (2009) | 1406 | 1253 | – | First plant GEM, modeling heterotrophic cell growth |
| | | de Oliveira Dal'Molin, Quek, Palfreyman, Brumbley, and Nielsen (2010a) | 1567 | 1748 | 1419 | AraGEM, modeling photosynthesis, photorespiration and respiration |
| | | Saha, Suthers, and Maranas (2011) | 1798 | 1820 | 1597 | GEM build alongside a rice model |
| | | Mintz-Oron et al. (2012) | 1363 | 1079 | 1791 | Emphasis on compartmentalization. Pipeline to reconstruct tissue specific models |
| | | Bekaert, Edger, Hudson, Pires, and Conant (2012) | 1785 | 1914 | – | Focus on glucosinolate biosynthesis, in particular on the metabolic cost of that biosynthesis |
| | | Cheung et al. (2013) | 2769 | 2739 | 2857 | Study maintenance cost, in particular in stress conditions, using Pareto optimality analysis |

*Continued*

**Table 1** Genome-scale metabolic models of plant species.—cont'd

| Species | Classification | Publication | Reactions | Metabolites | Genes | Applications |
|---|---|---|---|---|---|---|
| | | Cheung, Poolman, Fell, Ratcliffe, and Sweetlove (2014) | 5889 | 5515 | – | Comparison of C3/CAM photosynthesis. Modeling of day/night metabolism for leaves |
| | | de Oliveira Dal'Molin, Quek, Saa, and Nielsen (2015) | 1607 | 1782 | 1409 | Multi-tissue model (leaf, stem, root) across diurnal cycle. Study differences between source and sink tissues in particular on C and N allocation |
| | | Seaver et al. (2015) | 2801 | 2864 | 4262 | Creation of a so-called "Evidence-Based Model" from KEGG, AraCyc and existing *A. Thaliana* GEMs. A particular focus was put on compartmentalization |
| | | Shaw and Cheung (2018) | 11,320 | 10,664 | – | Multi-tissue model (leaf and root) across diurnal cycle. Study metabolism through different growth stages and resource availability with dFBA |
| | | Scheunemann, Brady, and Nikoloski (2018) | 2199 | 1813 | | Root-specific model constructed using transcriptomics and proteomics data, from the PlantSEED based Arabidopsis model |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Oryza sativa* | Monocot, Commelinids, Poaceae | Liu et al. (2013) | 3316 | 2986 | 4426 | Integration of GEM, protein-protein interaction, and gene regulatory networks to study phenotype-genotype interactions in rice |
| | | Poolman, Kundu, Shaw, and Fell (2013) | 1736 | 1484 | – | Leaf model to study redox, photorespiration and mitochondrial metabolism in response to varying light levels |
| | | Lakshmanan et al. (2015) | 2283 | 1999 | 2164 | Using transcriptomic data to study metabolism under different light conditions (blue, green, red, white and dark) |
| | | Chatterjee and Kundu (2015) | 1721 | 1544 | – | Focus on nitrogen assimilation and chlorophyll biosynthesis under different photorespiration conditions |
| | | Chatterjee, Huma, Shaw, and Kundu (2017) | 1136 | 1380 | 3602 | Simulating leaf metabolism across different light and enzymatic cost conditions, and for varying RuBisCO activity |

*Continued*

**Table 1** Genome-scale metabolic models of plant species.—cont'd

| Species | Classification | Publication | Reactions | Metabolites | Genes | Applications |
|---|---|---|---|---|---|---|
| *Zea mays* | Monocot, Commelinids, Poaceae | de Oliveira Dal'Molin, Quek, Palfreyman, Brumbley, and Nielsen (2010b) | 1588 | 1755 | 11,623 | C4GEM: comparison between maize, sorghum and sugarcane metabolism. Study flux distributions in mesophyll cells and the bundle sheath cells |
| | | Saha et al. (2011) | 1985 | 1825 | 1563 | FBA simulation under different conditions (photosynthesis, photorespiration, respiration) |
| | | Simons et al. (2014) | 8525 | 9153 | 5824 | Study carbon fixation and nitrogen assimilation under different nitrate conditions, and comparison with mutant phenotypes |
| | | Seaver et al. (2015) | 2629 | 2634 | 5540 | Three models for leaf, embryo and endosperm tissues, based on transcriptomic data |
| | | Bogart and Myers (2016) | 2720 (635) | 2725 (603) | 5204 (2140) | A comprehensive model and a high-confidence subset of this model (in parentheses). Introduce non-linear relationship between carbon dioxide and oxygen levels, and reaction rates |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Sorghum bicolor* | Monocot, Commelinids, Poaceae | de Oliveira Dal'Molin et al. (2010b) | 1588 | 1755 | 3557 | C4GEM: comparison between maize, sorghum and sugarcane metabolism. Study flux distributions in mesophyll cells and the bundle sheath cells |
| *Saccharum officinarum* | Monocot, Commelinids, Poaceae | de Oliveira Dal'Molin et al. (2010b) | 1588 | 1755 | 3881 | C4GEM: comparison between maize, sorghum and sugarcane metabolism. Study flux distributions in mesophyll cells and the bundle sheath cells |
| *Setaria italica* | Monocot, Commelinids, Poaceae | de Oliveira Dal'Molin et al. (2016) | 1515 | 1690 | – | Multi-omics data integration (transcriptome, proteome, metabolome) in mature and immature stem/leaf phytomers |
| *Solanum lycopersicum* | Eudicot, Asterids, Solanaceae | Yuan, Cheung, Poolman, Hilbers, and van Riel (2016) | 2143 | 1998 | 3410 | Simulate tomato leaf metabolism under heterotrophic and phototrophic conditions, and under drought stress |
| *Solanum tuberosum* | Eudicot, Asterids, Solanaceae | Botero, Restrepo, and Pinzón (2018) | 2072 | 1938 | 2751 | Simulate response to a pathogen by integrating transcriptomic data as constraints into fluxes of reactions |

*Continued*

**Table 1** Genome-scale metabolic models of plant species.—cont'd

| Species | Classification | Publication | Reactions | Metabolites | Genes | Applications |
|---|---|---|---|---|---|---|
| *Medicago truncatula* | Eudicot, Rosids, Fabaceae | Pfau et al. (2018) | 2909 | 2780 | 3403 | Study the effect of symbiosis on metabolic fluxes, in particular regarding oxygen availability |
| *Mentha x piperita* | Eudicot, Asterids, Lamiaceae | Johnson, Lange, Srividya, and Lange (2017) | 624 | 466 | 757 | Transcriptomics used to reconstruct the model and to perform modeling. Focus on monoterpenes biosynthesis |
| *Populus trichocarpa* | Eudicot, Rosids, Salicaceae | Sarkar and Maranas (2020) | 3282 | 2502 | 7188 | First GEM for a perennial woody tree. Reconstructed to help interpreting results of SNP analysis |
| *Glycine max* | Eudicot, Rosids, Fabaceae | Moreira et al. (2019) | 3001 | 2814 | 6127 | Studying the mobilization of soybean seed reserves during growth over 4 days. Highlight metabolic switches |

sources: KEGG (Kanehisa, 2000) and MetaCyc (Caspi et al., 2016), via the Plant Metabolic Network (Schläpfer et al., 2017). Having so little data-sources could become an issue, in particular regarding plant species, since compartmentalization is of great importance in plant metabolism. Those databases often lack precise information regarding the subcellular localization of reactions, and therefore the need for transport reactions across intracellular membranes. Such transport reactions are enabled by protein transporters that are often poorly characterized. The lack of information on subcellular localization of reactions and transporter protein can lead to modeling inconsistencies.

During the reconstruction of GEMs, metabolomics could play a determinant role. Indeed, the reactions present in a metabolic network describe how metabolites are produced and consumed. Theoretically, metabolomics could one day produce a comprehensive list of molecules present in a given tissue. If, in practice, the list of molecules is never complete, it can still give a large set of molecules either preponderant in the biomass, hence helping for the formulation of the biomass function, or present in very small quantities, as results of the secondary metabolism pathways, for instance. Such lists can be very useful to perform gap-filling, either from a topological (Prigent et al., 2017) or flux-based (Thiele, Vlassis, & Fleming, 2014) point of view.

## 2. Bases and specificities of plant metabolic modeling

Constraint-based modeling is a classical approach to study flux distributions in GEMs. Indeed, by assuming that the models are at quasi-steady states, flux-based methods such as Flux Balance Analysis (FBA) (Orth, Thiele, & Palsson, 2010) enable exploration of possible fluxes going throughout a model. Flux-based approaches have been widely used to study plant metabolism in conjunction with GEMs. Given the Quasi-Steady-State Assumption (QSSA), performing accurate modeling using GEMs is largely based on a good definition of influxes and outfluxes of the models. Those are usually defined as the entrance of metabolites from outside the metabolic system leading to the production of biomass. Plants possess a big specificity here, since they are gaining most of their energy and carbon from photosynthesis. This mechanism has thus been widely studied using GEMs. This aspect was tackled at first with the AraGEM model (de Oliveira Dal'Molin et al., 2010a) where photosynthesis and photorespiration were accurately modeled.

Plants are complex organisms possessing several organs and tissues, and some of them are not capable of photosynthesis. Hence, to model the biological reality, several multi-organ and multi-tissue models have been built and analyzed. We can, for example, cite (Seaver et al., 2015) where the authors used transcriptomic data to build specialized GEMs for maize leaves, endosperm and embryos, and specialized biomass functions for each model. More recently, Shaw and Cheung developed a multi-tissue model for *A. thaliana* by reconstructing a model including leaves and root metabolism during day and night phases (Shaw & Cheung, 2018). Those tissues could cross-talk via a common pool of metabolites representing the phloem. A similar approach was used previously to model interactions between leaves, stems and roots (de Oliveira Dal'Molin et al., 2015).

To perform accurate modeling, precise biomass and nutrient definitions are required but it is often not sufficient since the solution space remains very big. Hence, several types of data can be incorporated into GEMs to improve the quality of the models. Since the cost of generating omics data continues to decrease while the quality of such data increases, it is not surprising that transcriptomics, proteomics and metabolomics are more and more used to reduce the number of possible flux distributions in a GEM. While gene expression, protein concentration and fluxes do not always correlate, transcriptomics and proteomics are often used to build tissue/environment/time-specific GEMs that contain reactions that could be catalyzed only in some cells or conditions. On the other hand, metabolomics (including isotope labelling) gives information that can be directly used in simulations. For example, targeted metabolomics has been previously used to obtain semi-quantitative information on 100 metabolites (de Oliveira Dal'Molin et al., 2016), in conjunction with transcriptomics and proteomics data, to compare the abundance of metabolites in young and mature tissues of *Setaria italica*. Isotopic labeling such as $^{13}C$ can also be used to obtain very precise information on subsets of the metabolic models. Metabolic Flux Analysis (Wiechert, 2001) can be used to identify precise fluxes that would lead to a given distribution of labeled isotope metabolites. Such fluxes put big constraints on the other parts of the metabolism, as what was performed by Hay et al. (2014) on a quite small model where this technique improved the predictive power of the simulations.

We have seen previously how metabolomics can be used to improve the quality of GEMs, in particular during the gap-filling process and during simulations. GEMs could also be used to improve the quality of metabolomics data, in particular of Mass Spectrometry (MS) data. Indeed, metabolic

networks give access to a large number of metabolic reactions involving reactants and products. Hence, GEMs give access to thousands of metabolites that should be present at least under some conditions in the metabolome of a given species. Depending on the metabolites and the technique used, the signatures associated with those molecules should be recovered, in MS data for instance. This information could then be used to annotate some of the peaks, thus participating in the identification effort that remains highly challenging. This could lead to a virtuous loop where GEMs reactions coverage would be improved by metabolomics, and those GEMs would, in turn, improve metabolomics annotation.

All the GEMs presented in this part have been successfully used to model the metabolism of plant cells or tissues, sometimes interacting together. Yet, plants never live alone. Bacteria and fungi permanently accompany them during their normal or pathological development. The modeling of the metabolism of bacteria and fungi is far ahead of that of plants. It is now becoming necessary to integrate plant models with the symbionts, parasites and pathogens that accompany them.

## 3. Metabolic modeling of plant-microorganism interactions

### 3.1 Microorganisms involved in biotic interactions

GEMs were initially developed, since 1999, to study microbial strains for applications in medicine and biotechnology (Kim, Sohn, Kim, Kim, & Lee, 2012). Nowadays, a wide diversity of GEMs for microorganisms exist for more diverse applications, and the available GEMs include several microorganisms interacting with plants such as symbionts and pathogens, from bacteria and eukarya domains (Table 2).

Extension of GEMs to bacterial plant pathogens started in 2015 for the causal agent of soft-rot disease: *Pectobacterium carotovorum* (Wang et al., 2015). The metabolic model of *P. carotovorum* was reconstructed based on genomic annotation and *Escherichia coli* metabolic model. It was confronted with Biolog Phenotype Microarrays (PMs) experiments and applied to the discovery of agricultural bactericide targets. The next GEM, published in 2016 (Peyraud et al., 2016) was the model of *Ralstonia solanacearum*, the agent of bacterial wilt. It was reconstructed using a broader range of reference organisms (*E. coli*, *Pseudomonas aeruginosa*, *Ralstonia eutropha, Bacillus subtilis*), and included a macromolecule module representing the biosynthesis of macromolecular components (539 reactions), such as a wide range of

**Table 2** Genome-scale metabolic models of plant-interacting microorganisms.

| Species | Classification | Biotic interaction | Publication | Metabolites | Reactions | Genes | Applications |
|---|---|---|---|---|---|---|---|
| *Rhizobium etli* | α-Proteobacteria | Symbiotic | Resendis-Antonio, Reed, Encarnación, Collado-Vides, and Palsson (2007) | 371 | 387 | 363 | Metabolic pathways involved in nitrogen fixation |
| *Pectobacterium carotovorum* | γ-Proteobacteria | Pathogenic | Wang et al. (2015) | 1113 | 2235 | 1209 | Discovery of agricultural bactericide targets |
| *Ralstonia solanacearum* | β-Proteobacteria | Pathogenic | Peyraud, Cottret, Marmiesse, Gouzy, and Genin (2016) | 2644* | 2086* | 1476 | Resource allocation trade-off between growth and virulence |
| *Sinorhizobium meliloti* | α-Proteobacteria | Symbiotic | diCenzo et al. (2016) | 1579 | 1825 | 1575 | Metabolic reprogramming in different environments |
| *Phytophtora infestans* | Eukaryota, oomycete | Pathogenic | Rodenburg, Seidl, de Ridder, and Govers (2018) | 2394 | 2685 | 1301 | Mapping of transcriptomics data at different life-stages |
| *Pectobacterium parmentieri* | γ-Proteobacteria | Pathogenic | Zoledowska et al. (2019) | 2080 | 2182 | 1245 | Metabolic behaviors at different steps |
| *Sclerotinia sclerotiorum* | Eukaryota, fungus | Pathogenic | Peyraud, Mbengue, Barbacci, and Raffaele (2019) | 1525 | 1495 | 1039 | Metabolic heterogeneity/division of labor |
| *Xanthomonas oryzae* | γ-Proteobacteria | Pathogenic | Koduru et al. (2020) | 808 | 839 | 673 | Effects of nitrogen fertilizers, relation between quorum sensing and metabolism |
| *Sinorhizobium fredii* | α-Proteobacteria | Symbiotic | Contador, Lo, Chan, and Lam (2020) | 508 | 535 | 541 | Symbiosis yields in different symbiotic organisms/upon different environmental conditions |
| *Xylella fastidiosa* | γ-Proteobacteria | Pathogenic | Gerlin et al. (2020) | 1107 | 1158 | 473 | Metabolic determinants of the slow-growing (fastidious) phenotype |

[a]Including 539 reactions and 367 metabolites corresponding to the biosynthesis of macromolecules (e.g. type-three effectors, plant cell-wall degrading enzymes).

virulence factors. The model was also intensively calibrated: Biolog PMs, bacterial cultures, chemical analysis of supernatants, virulence factors quantification, both on the wild-type strain and mutants reduced in virulence. The authors demonstrated that a resource allocation trade-off controlled by the population level occurs between the production of virulence factors and fast proliferation. Metabolic modeling thus helps understanding the metabolic adaptation of pathogens to plant environments. Three other bacterial plant pathogen GEMs were published:

**(i)** *Pectobacterium parmentieri* (Zoledowska et al., 2019). The reconstruction used the online tool KBase (Arkin et al., 2018), and Biolog PMs data. It predicted the metabolic behaviors of the bacterium in two conditions: persistence in soil and plant colonization in the rhizosphere.

**(ii)** *Xanthomonas oryzae* (Koduru et al., 2020). The reconstruction was based on KEGG (Kanehisa, 2000) and MetaCyc (Caspi et al., 2016) databases and refined by adding specific pathways of the organism. The authors studied the impact of quorum sensing deletion on metabolic pathways. Experiments were performed to validate the model: bacterial cultures, excretion quantifications and amino acid consumption.

**(iii)** *Xylella fastidiosa* (Gerlin et al., 2020). The GEM was reconstructed using five reference organisms using the same method as *R. solanacearum*, and was refined with Biolog PMs. This work showed how metabolism was involved in the self-limiting behavior of the bacterium inside the host. In particular, *X. fastidiosa* lacks several key metabolic reactions, which contribute to the slow-growing phenotype of the organism and is suspected to be an advantage to remain under the radar of the plant immune system.

So far, fewer GEMs are available for bacterial symbionts. In 2007, a metabolic model of a legume symbiont (*Rhizobium etli*) was reconstructed using the KEGG database (Resendis-Antonio et al., 2007). This bacterium, located in legume roots, transforms atmospheric nitrogen into ammonium and thus provides nitrogen to the plant in exchange of organic matter. The qualification of the model as a "genome-scale" metabolic model is questionable as only 387 reactions are included in the model, versus between 1000 and 2500 reactions for most metabolic models. This model compared in silico metabolic pathways involved in nitrogen fixation to experiments. More recently, diCenzo et al. (diCenzo et al., 2016) applied metabolic modeling to another legume symbiont *Sinorhizobium meliloti*, a major model organism of nitrogen fixation by bacteria, using KBase (Arkin et al., 2018)

for reconstruction. This model was used to simulate metabolic reprogramming in different niches, and the importance of the symbiotic plasmid for growth in the rhizosphere environment. *S. meliloti* metabolic network was also used as a reference organism to generate the metabolic model of another legume symbiont, *Sinorhizobium fredii* (Contador et al., 2020). On this study, the authors compared the symbiotic yields between existing legume symbionts models.

Metabolic modeling is more easily applicable and extended to prokaryotes than eukaryotes, because of the added complexity of eukaryotic cell compartmentalization, which increases the curation and validation duration. Thus, to our knowledge, only two eukaryotic microbe models related to plants are available. A model of the pathogenic oomycete *Phytophtora infestans*, the causal agent of the late blight on potato and tomato (Rodenburg et al., 2018) was generated using hidden Markov models and KEGG database and was used to map transcriptomics data. It was followed by the metabolic model of *Sclerotinia sclerotiorum* (Peyraud et al., 2019), based on *Saccharomyces cerevisiae* and *Aspergillus niger* as reference organisms. It investigated the division of labor and metabolic heterogeneity *in planta*.

## 3.2 Plant-microorganism interactions

Several metabolic models of plant-interacting microorganisms are now available to the scientific community (Table 2). On the other hand, plant GEMs were also largely developed for both model organisms and crops (Table 1). It is thus tempting to reuse these constructions to build plant-microorganism interaction metabolic models, in order to better understand trophic interactions, the effect of one organism on the other and the ways to favor or restrict an interaction. However, this raises the question of how to model multi-organism systems. Metabolic modeling has already been applied to other multi-organism systems. The first published model studied the cooperation of two microorganisms for degrading organic substrates in the context of anaerobic digestion (Stolyar et al., 2007). In a biotechnological context, it helped to understand and quantify how cooperation between strains could improve the yield of a desired compound (Salimi, Zhuang, & Mahadevan, 2010). An immune cell-pathogen system was also built to provide medical solutions against *Mycobacterium tuberculosis* (Bordbar, Lewis, Schellenberger, Palsson, & Jamshidi, 2010).

For plant-microorganism interactions, three models were published so far. The first studied the process of nitrogen fixation through legume-rhizobia

symbiosis (Pfau et al., 2018), with the model organisms *Medicago truncatula* (legume) and *S. meliloti* (symbiont). The authors combined a two-compartment model of *M. truncatula* (root and shoot) and a model of *S. meliloti*. In the model, *S. meliloti* is able to acquire atmospheric nitrogen ($N_2$) and convert it to ammonium and amino acids. These nitrogen sources are subsequently provided to the plant root. In exchange, the plant root provides oxygen and organic carbon to the symbiont. Simulations showed that symbiosis was not beneficial for the global growth rate (plant and symbiont) if enough nitrogen was available in the soil as nitrate ($NO_3^-$) and ammonium ($NH_4^+$). On the contrary, the growth increased when low nitrogen was available in the soil. In the same study, the authors investigated the impact of varying oxygen supply on the symbiont central metabolism, as oxygen is suspected to be a limiting factor for this symbiosis. In this interaction model, a curated metabolic model of the plant *M. truncatula* was used. However, the symbiont was not modeled using a curated and published metabolic model (as the one presented in the precedent section), but an automatic extraction from an online metabolic database (MetaCyc). This approximation limits the precision and the reliability of the study to predict fluxes inside the bacterium, and to associate bacterial phenotypes to metabolic genes. A new model of *M. truncatula*-*S. meliloti* interaction was published by diCenzo and collaborators (diCenzo, Tesi, Pfau, Mengoni, & Fondi, 2020), named ViNE (*Virtual Nodule Environment*). The bacterium part of the model was based on a curated and refined GEM (diCenzo et al., 2018, 2016). The multi-tissue plant part of the model was the previously published GEM of *M. truncatula* (Pfau et al., 2018) and its multi-tissue modeling methodology. Some quantitative aspects of the interaction in accordance with experimental data were included in the simulations (shoot and root representing 98% of the biomass versus 2% for the nodule). This allowed to validate the model prediction (growth rates, nitrogen fixation rates, mutant phenotypes). It also provided novel insights on the symbiotic interaction such as costs of nitrogen fixation and nodule formation, impacts of nodule size/fixation efficiency on the system, or bacteroid carbon sources (diCenzo et al., 2020).

Rodenburg et al. (Rodenburg et al., 2019) developed a metabolic model for a plant-pathogen interaction. They combined a metabolic model of tomato (*S. lycopersicum*) leaf and *P. infestans*. The two-organism metabolic model was manually curated to represent the interaction in agreement with the biological knowledge, such as the dependence of *P. infestans* to thiamine produced by the plant. To study flux distribution, different potential objective functions were tested: minimization of import fluxes from the leaf to the

pathogen, minimization/maximization of the number of metabolic reactions used by the pathogen, and minimization of both the pathogen's import fluxes and the pathogen's metabolic reactions. Flux coupling analysis methodology suggested plant or pathogen pathways that could be used to control the pathogen spread. Finally, transcriptomics and metabolomics data at different stages of the infection were mapped onto the metabolic network to highlight the dynamics of metabolism during the infection process. This mapping illustrated the shutting down of alternative pathways as the infection progressed, and identified potential pathogen's targets and plant pathways to engineer to block the infection. However, these study results are limited by the lack of physiological/metabolomic experiments on *P. infestans*-tomato leaf interaction. Quantitative data (e.g. growth rates, organisms' biomass ratio) and evidence on the exchanged molecules would provide additional insights on tomato late blight.

## 3.3 Toward a quantitative and dynamic representation of plant-microorganism interaction

Metabolic modeling of plant-microorganisms interaction is already emerging (diCenzo et al., 2020; Pfau et al., 2018; Rodenburg et al., 2019), and many will be in the future, since the number of plant and microorganism models increases each year. The model developed by diCenzo et al. (2020) included a quantitative representation of the interaction and addressed the question of the availability of the resources at the interaction site. The numerous insights provided by the study confirms that these considerations are central to investigate metabolic interactions.

Thus, we propose three main methodological points that should be addressed in the future metabolic models of interactions: the quantitative aspect of the interaction, the application of QSSA and the availability of the resources between organisms (Fig. 1).

### *3.3.1 Quantitative representation of the interaction*

So far, most of the models available lack the integration of plant and microorganism quantifications; the analysis of the interaction is only qualitative. Most of the approaches unraveled which metabolic fluxes were activated in the presence or absence of the organism interacting with plants, but they did not estimate whether these flux variations were significant or not at the scale of whole plant metabolism. Being able to quantify the metabolic effect of one organism on the other would help to better understand the interaction. For example, Rodenburg et al. were able to highlight by FCA the plant

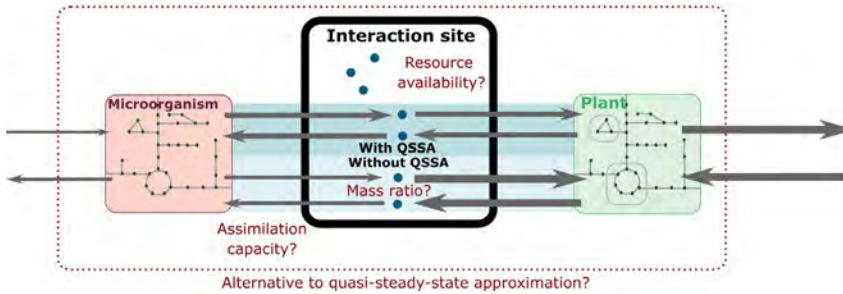



**Fig. 1** Toward a quantitative and dynamic metabolic model of biotic interactions. The principal methodological and experimental issues that must be addressed are represented in red: quantification of the microorganism/plant ratio, use of a methodology that relaxes the QSSA growth hypothesis, and accurate representation of the resources available in the interaction site. Exchange fluxes are represented in two scenarios: assuming QSSA, metabolic flux from plant to interaction site and interaction site to microorganism must be equal (in mol·time unit$^{-1}$) while without QSSA, this constraint is relaxed. Relaxing equality in exchange fluxes could be relevant in several cases, as the microorganism and the plant can have contrasted assimilation fluxes.

metabolites potentially assimilated by the pathogen as nutritive sources (Rodenburg et al., 2019). Implementation of quantitative data (ratios, growth rates) would allow verifying if the proposed scenarios were realistic. Being able to quantify the metabolic effect of one organism on the other would help to understand better the interaction. It would be interesting to estimate at which population density the pathogen or symbiont impacts negatively or positively plant growth, as in the *Virtual Nodule Environment* metabolic model. It would also be interesting to study how much the microorganism's growth relies on virulence mechanisms, and if it can coerce the plant to provide nutrients in the interaction site. Ultimately, this study would allow unraveling how symbiont organisms could be optimized and selected to favor plant growth and agricultural yields or conversely, how pathogens could be controlled in their growth and virulence to avoid being deleterious to the plant. Yet, building a quantitative metabolic model would need experimental work to calibrate the model. Still, some of the experiments could be simple to perform, such as mass estimations of the ratio between the plant and the interacting organism, ideally at different stages of plant growth and development.

### 3.3.2 Quasi-steady-state approximation in a multiorganism system

To estimate fluxes, metabolic modeling relies on simplifications, as the numerous kinetic parameters required at the global cellular scale are majorly

unknown (Orth et al., 2010). One of the hypotheses behind the simplification of the model, the QSSA, stipulates that concentrations of internal metabolites are assumed constant. It is justified by the fact that internal metabolites do not accumulate, they are usually consumed immediately as they are produced by the cell to synthesize macromolecules such as membrane lipids or proteins (Orth et al., 2010). Metabolic modeling of a multi-organism system at QSSA would imply that metabolites in the whole system (metabolites of the plant, the interacting organism and in the interaction site) are consumed immediately as they are produced, and consequently that both plant and interacting organism grow at the same velocity. While this could potentially be a reasonable hypothesis for symbiosis, balanced growth seems inadequate to simulate the dynamics of pathogenesis. Indeed, in several cases, the pathogen's proliferation will accelerate as it deploys its virulence mechanisms against the host, and oppositely, the plant growth will decelerate. To tackle this problem and simulate the dynamics of plant-pathogen interaction, different options could be considered. First, it is possible to perform the simulation at a precise instant and consider that on a relatively short time interval, the system could be considered at quasi-steady-state. This instantaneous simulation would then be repeated at different time steps to access the dynamics of the interaction, by updating the biomass of each organism at each time step. This process would, however, require the acquisition of at least growth of the two organisms throughout the experiment. Another solution is to consider that separately, the organisms are at QSSA, and allow non-QSSA behaviors for the exchanged metabolites. This would require implementing kinetics in the model, and thus estimate their parameters if it is possible.

Alternatively, metabolic modeling frameworks were developed to overpass the limitations of the QSSA hypothesis. To represent the dynamics of microbial communities, Zomorrodi et al. (Zomorrodi, Islam, & Maranas, 2014; Zomorrodi & Maranas, 2012) proposed a multi-level and multi-objective framework, which is in theory adaptable for parasitic systems. To overpass the QSSA hypothesis, Baroukh, Muñoz-Tamayo, Steyer, and Bernard (2014), proposed to reduce metabolic networks into sub-networks at QSSA, allowing dynamics for the exchange metabolites between subnetworks, generating a smaller metabolic network that could be modeled without hypothesizing QSSA. These two methodologies were not applied to host-symbiont or host-pathogen interactions yet, but are encouraging directions to take.

### 3.3.3 Resource availability at the interaction site

To provide a reliable representation of biotic interactions, a critical point is the representation of the resources available at the interaction site between the two organisms. Metabolic modeling of an organism strongly depends on the calibration of the uptake rate of the substrates and the carbon/nitrogen sources modeled. Comparing the resource usage in the simulation with experimental data also provides a better understanding of the interaction (diCenzo et al., 2020). Upon biotic interaction, these aspects are dependent on the representation of the interaction site. Thus, it is necessary to: (i) define which resources are made available by the plant/the interacting organism, (ii) quantify at which amount they are present, (iii) indicate if these resources could be assimilated by the plant/the interacting organism, and (iv) quantify the assimilation capacity of the resources by the plant/the interacting organism. These details are not always available, in particular the quantity available of the resource and the organisms' assimilation capacities, and would require substantial data acquisition.

Currently, the generation of pathogens and symbionts metabolic models is growing fast. It provides large resources to understand how microorganisms interact with plants. The next step, as the number of plant metabolic models are also growing rapidly, would be to integrate pathogen/symbiont and plant in a global metabolic model. Even though three studies were published and provided some insights (diCenzo et al., 2020; Pfau et al., 2018; Rodenburg et al., 2019), in particular when quantitative aspects were addressed, this approach is still poorly used. Indeed, quantitative and dynamic metabolic models of biotic interactions are arduous, because they require both the application of non-standard metabolic modeling approaches and the integration of several experimental data, some of them difficult to acquire.

## 4. Visualizing metabolism

Data related to metabolism often makes little sense if not put into context, i.e. confronted with the knowledge that has been accumulated about the organism metabolism. The sheer number of known metabolites and biochemical reactions linking them require computational solutions to be exploited. Yet, it is sometimes hard to formulate the interpretation requirements as a problem that can be solved programmatically.

Visual inspection of data and knowledge side by side, highlighting discriminant metabolites on a pathway map, for example, became a handy tool

to formulate hypotheses and design subsequent analysis. Knowledge related to metabolism can be intuitively represented as a network of metabolites connected by biochemical transformations, and the practical aspects of the visual exploration of such networks will be briefly introduced hereafter.

While the reconstruction of a metabolic network is a difficult task, as mentioned in the first section of this chapter, it is not to be confused with the creation of its visual representation, the network drawing, yet another far from trivial task on its own. A single network may be represented by several drawings of various usability and purposes, and this section aims at providing some insight about choosing the right one and avoiding common pitfalls.

The science of information visualization has provided numerous methods and tools that help to deal with complex and abstract data, including large networks (Herman, Melancon, & Marshall, 2000; von Landesberger et al., 2011). Many algorithms have been produced to create network layouts that maximize some aesthetic criterion, such as reducing cluttering, edge crossing or node overlaps (Ware, Purchase, Colpoys, & McGill, 2002; Yoghourdjian et al., 2018).

Such techniques have found an abundance of applications in network-related sciences, such as sociology or physics. Yet, their popularity has not pervaded much in the metabolism research community. One of the main reasons may be the discrepancy between the drawing produced automatically and the visual semantics expected of metabolic representations. The most striking example is the drawing of whole networks obtained using force-directed algorithms, which earned in the field the tongue-in-cheek title of "hairballs" or "ridiculograms," referring to their little use beyond advertising the complexity of a research question (Fig. 2).

The "quality" of visualization depends on a given task to be performed or a particular property to be shown. Many layouts are prone to reveal the global or meso-scale organization of a system represented as a network. Force-directed layouts, for example, can give a glimpse of a community organization in a network, with visible groups of nodes that seem to be more connected with each other than with those of the other groups.

While understanding the global properties of metabolism as a system is a relevant and still an ongoing research question to this day, metabolic network visualizations are mainly used to decipher the role of specific elements (metabolite or reaction), or group of elements, in the system, rather than properties of the system itself. We will hence focus on these applications for the rest of this section.

The complex chains of interactions that can relate the concentrations of two metabolites, involving numerous branching, usually makes visual interpretation challenging. One approach consists in dividing the metabolic

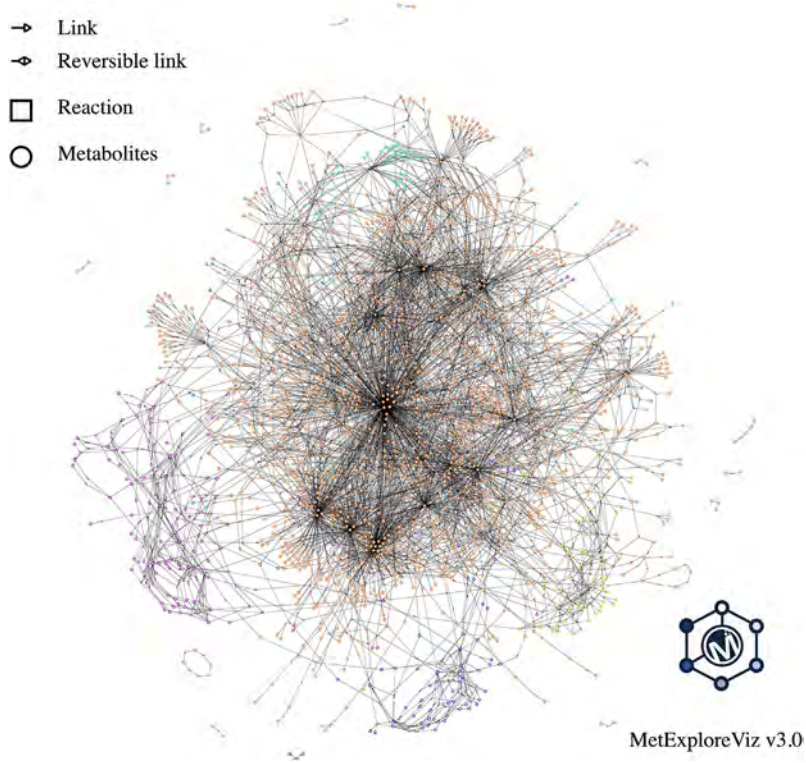



**Fig. 2** ARAGEM metabolic network drawn using MetExploreViz with a force-directed layout.

network into smaller sub-parts that can be drawn manually and inspected more easily (Reimand et al., 2019; Saraiya, North, & Duca, 2005; Schuster, Pfeiffer, Moldenhauer, Koch, & Dandekar, 2002).

The definition of these sub-parts had historically been driven by their functional purposes (e.g. glycolysis, amino acid biosynthesis). Reasoning over individual metabolic pathways is still the most used strategy, following the aforementioned "divide and conquer" approach. However, the rise of global and untargeted experimental approaches (metabolomics, transcriptomics …) tends to dampen such practice, since observing entities of interest spanning multiple pathways is not rare (not to mention the subjective definition of pathways and consequently the lack of consensus between pathway databases). Thus, the call for global visualization of metabolism has grown strong (Dinkla & Westenberg, 2012; Gehlenborg et al., 2010).

The first global visualizations were proposed in the late 1960s, with the representation of G. Michal (Michal, 1999) and around the same time D. Nicholson (Nicholson, 2005). Not oriented toward the identification

of large-scale structures, they opted for a planar representation dedicated to "knowledge navigation," and are often referred to as maps or atlases.[a]

Many proposed views today still look like those original drawings, with the addition of a layer of interactivity made possible by technological advances since then. Arguably the most peculiar and yet popular descendant of those maps has been the KEGG atlas (Okuda et al., 2008). Using colors for the related sub-parts, straight lines and smooth bends, and evenly spaced nodes, the KEGG atlas (Fig. 3) share a striking resemblance with Harry Beck's design of subway maps (Degani, 2013), which itself share similarities with electric diagrams.

The key of the worldwide application of such design to subway maps lies in the fact that it is task-oriented, it is focused solely on what those maps are used for: finding a path from station A to station B. Prior to Beck's design, the London underground map was drawn as an overlay on the classical London city map. Despite being familiar, it had several drawbacks regarding path search: stations from the city center were cluttered due to a higher density than country-side stations, tunnel twists and turns were hard to follow, and a lot of the displayed information were related to the surface and thus of little relevance for underground navigation (Tufte, 1997).

Path search is undoubtedly a relevant task in a metabolic network context, allowing finding biochemical transformations that could relate two compounds, which might explain the appeal of the KEGG Atlas. However, the analogy falls short when considering the number of "lines" and "stations" on a metabolic network, which makes performing this task far from an easy ride.

Another subtle difference with impairing consequences is related to topology. While not strictly planar, an underground railroad network can be projected on a two-dimensional map with little distortion, due to a small depth compared to the surface spanning. A well-known and more difficult case is one of the world maps. Projecting the surface of a near-spherical planet on a 2D map is at the cost of some distortions (Mulcahy & Clarke, 2001), which can have consequences on the way we perceive the represented object. The relative difference of areas between the African continent and Antarctica or Greenland is commonly biased by the classical Mercator representation (representation that has the benefits of representing parallels and meridians as straight lines, which have made it a fair visualization for nautical related tasks).

[a] http://biochemical-pathways.com/#/map/1.

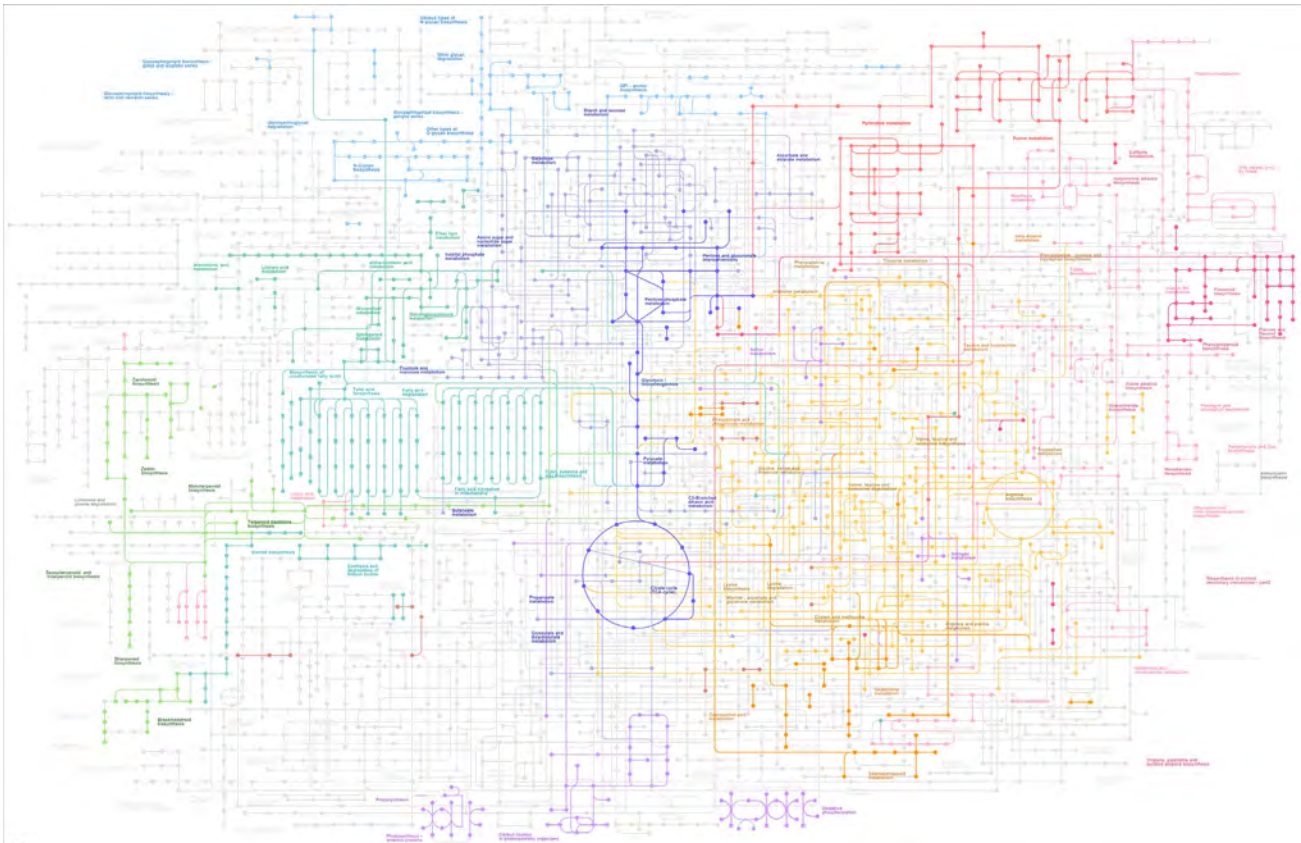


**Fig. 3** KEGG overall representation of *A. thaliana* metabolic pathways.

Metabolic networks have a far more complex "shape," which makes the representation in a low dimensional space tedious. One should be aware that metabolic maps usually rely on some tricks to flatten the view, duplicating nodes in several places, or removing some edges, which is far from facilitating path search tasks or neighborhood exploration. Scratching the surface often reveals that most "hand-drawn" global visualizations of metabolism tend indeed to be "patchworks" of local pathways, where the necessary distortions are oriented to conserve their integrity, thus bearing the same bias as individual pathway views.

The old saying that "the map is not the territory" thus applies very well to metabolic networks, and their representations are still an open challenge.

While existing solutions can still be useful for navigating the network from an initial focal point to extract knowledge, one must be cautious when reasoning from sets of nodes. As extrapolating walk distance between stations from a subway map would be ill-advised, extending navigational "metabolic atlas" to other tasks can be biased. Therefore, from previously mentioned points, the perceived distances in metabolic drawings do not necessarily represent the "true" metabolic distances (usually estimated from path-length distances in the network). While it is tempting to draw biological hypotheses from a group of metabolites that would appear grouped or spread on the map, their relative positions bear little information and can be misleading. A drawing dedicated to this specific task must be considered, using, for example, multidimensional scaling, at the cost of less readable paths.

As the size of metabolic models has been subject to tremendous growth (Table 1), the manual drawing of their underlying network has become a fastidious task. Beyond scaling issues, hand-drawn maps' main weakness is their rigidity toward a constantly expanding understanding of metabolism, as small changes might lead to re-drawn large portions of the map. The future of metabolic network visualization may thus lie in the adaptation of drawing algorithms toward better-defined tasks.

The most commonly used graph-drawing algorithm to embed a network in a two- or three-dimensional space is the force-directed layout. There exist many variations of this method but, even if there was in past decades a clear improvement of the algorithms in terms of complexity and reproducibility, the underlying concept remains the same. These algorithms consider the network as a physical system made of physical bodies connected by a system of forces (like springs). The assumption is that such a system will reach an equilibrium status which will position nodes and edges in the representation space in such a way that aesthetic criteria will be satisfied (e.g. even spread of

nodes, balanced length of edges). These algorithms are broadly used since they generate a representation for any kind of network, as they do not require any topological prerequisite. The main drawback, in the context of metabolic network studies, is that metabolic networks have topological features that usually lead to cluttered drawings. More precisely, dense parts of the network (nodes with high degrees and strong clustering index) tend to be overloaded in the center of the representation while, in the periphery, long isolated chains radiate in every direction, giving far more room to less central parts of metabolism. This representation captures well the overall organization of the metabolic network, but when it comes to navigating the network to glean some information about the role of a particular element, it is far from ideal.

At the other hand of the spectrum of graph drawing methods, lies planar representations (Didimo, Liotta, & Patrignani, 2019). This class of methods aims to provide a drawing with a good spread of nodes in the representation space while minimizing edge crossings by using edge bending. These drawings are quite efficient for instance for physical network representations such as electrical diagrams or software development schemas such as UML representations of program libraries. To achieve these goals the strategy is first to extract a subgraph which is planar (which can be drawn in two dimensions with no edge crossing) and then define an algorithmic solution to add the remaining edges to the initial drawing while minimizing the number of edge crossings. The advantage of this class of methods in the context of metabolic networks is the overall resemblance with atlas and maps, sharing the comfort of evenly spaced nodes for navigation. However, the main drawback is that, since metabolic networks are far from being planar, the algorithm will tend to create overly long edges that make visual path search quite challenging.

Based on the pros and cons of the various methods, some algorithms were proposed to combine the benefit of the existing approaches and try to fit biochemists' drawing conventions. The shared principle is to divide the network into sub-parts and apply the best-suited drawing algorithms to each part. For example, a circular layout could be applied to a sub-part containing a cycle, and a hierarchical layout could be applied to an acyclic sub-part. It was first designed to draw metabolic pathways (Fig. 4A) (Becker & Rojas, 2001). However, as acknowledged by authors, the approach is hardly scaling up to the network level. In fact, the challenge is to combine these separated drawings since subnetworks are usually highly overlapping. It thus requires additional operations (rotations, translation) to assemble sub-drawings in a harmonious arrangement.

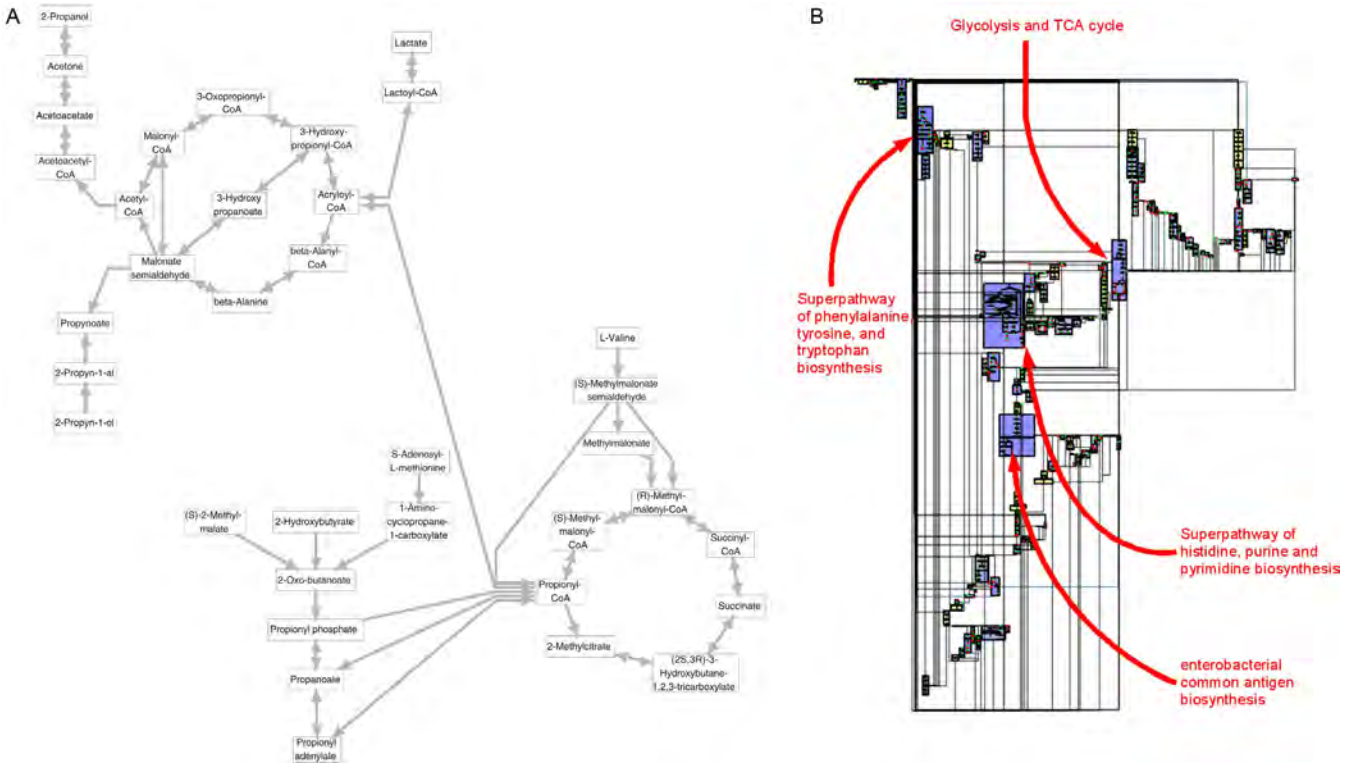



**Fig. 4** Representation of metabolic pathways. (A) Automatic drawing of propanoate metabolism from (Becker & Rojas, 2001). (B) Automatic drawing of *E. coli* metabolic network using a combination of cyclic and hierarchical drawing algorithms for pathways and a planar graph drawing for the edges connecting these pathways. *From Bourqui et al. (2007).*

In their article, Bourqui et al. (2007) propose to take advantage of the decomposition of the metabolic network into pathways. To limit the challenge of overlapping pathways, they use a coloration algorithm to detect the maximum number of pathways, which can be drawn without influencing others. Then, pathways are arranged using a planar drawing with cycle and hierarchical substructure search applied internally to each pathway (Fig. 4B).

Even if automatic drawings provide a good first sketch of the drawing of metabolic networks, it still requires some manual curation and modifications. For example, a handcrafted list of "side compounds" is usually removed from the view, since those compounds, including water or ATP, appear overly connected (they are hubs, to speak in network terms) and cause drastic cluttering. Nevertheless, one should note that some highly connected nodes, such as Pyruvate, make the network harder to draw, yet their removal would cause crucial loss of information. The automation of the hub removal is thus far from trivial as careful expert curation is required. Hence, more and more tools provide interactive edition features on top of initial automatic drawings. This is, for instance, the case of Cytoscape (Shannon et al., 2003), or MetExploreViz (Chazalviel et al., 2018) which is designed especially for metabolic network representations (bipartite default style, edges bundling for reactions, etc).

Overall, moving from a segmented view using pathways to a more holistic, system-level representation is at cost of time-consuming manual curation, despite some progress regarding automatic drawing to generate "seeds". Beyond the creation of these multi-purpose global views aiming at knowledge exploration, it is possible to imagine task-oriented representations (see example in Fig. 5, adapted from (Frainay et al., 2018)).

In that case, it is of interest to apply more data-driven drawings of metabolic networks such as attribute layouts or hive plots (Krzywinski, Birol, Jones, & Marra, 2012) as support to convey a particular finding and identify intriguing patterns. Finally, it is important to notice that the quality of representation has not only to be assessed in terms of aesthetic criteria but also based on their ability to improve the user capacity to perform a search task (e.g. looking for a metabolic path between two metabolites of interest).

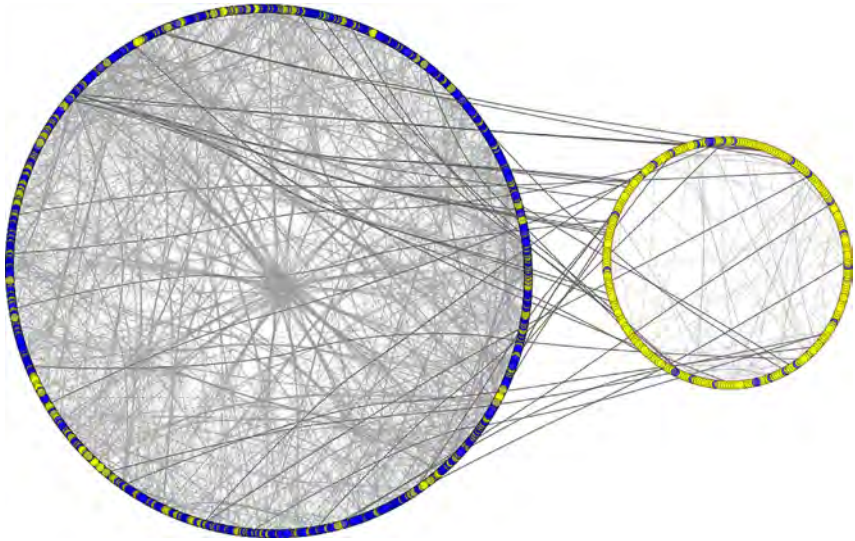


**Fig. 5** An unconventional and task-oriented view of a metabolic network. The task here is the comparison of two groups of nodes partitioning the network. A circular layout has been applied to the two groups. This view emphasizes the difference between the two groups in terms of size and inner density, shows the relative scarcity of links between the two groups, and the overrepresentation of a given property in one of them, using color mapping. Drawn with Cytoscape using "group attributes layout" method.

## 7. Conclusion

As shown in this chapter, the development of omics-related techniques in plants have allowed metabolic modeling to become widespread. Since the acquisition of plant complete genomes remains complicated, compared to obtaining such data for bacteria, for example, plants GEMs reconstruction and analysis have benefited from all the methods that have been developed for metabolic modeling of bacteria and other eukaryotes.

Even if reconstructing and analyzing plant GEMs remain difficult (the number of existing plant GEMs speaks for itself), the number of GEMs for "non-Arabidopsis" species is growing, especially in the past few years. The next two big steps to improve modeling are to better take into account the compartmentalization of the models, and metabolomics will play a huge role here, and to better take into account the interaction among models, either among different tissues inside a plant or with the microbiome interacting with the plant. Once again, metabolomics will have a major role here, in particular by providing lists of metabolites shared among the different

organs and/or species. Complementary to metabolomics data acquisition, developing biologically relevant models will also require careful considerations and adaptation of existing methodologies. Indeed, constraint-based modeling relies on mathematical hypotheses, such as QSSA, which may not match the complexity of compartmentalized, multi-tissue or multi-organism systems.

Finally, modeling plant metabolism will face the same challenges as modeling all other kinds of species: the amount of omics data to integrate into the models will keep growing, and modeling techniques will have to cope with such data. Improvements in the visualization of metabolic networks will undoubtedly help to provide context to the various data that may be associated with GEMs, and hence improve the understanding of such big networks.

## References


Arkin, A. P., Cottingham, R. W., Henry, C. S., Harris, N. L., Stevens, R. L., Maslov, S., et al. (2018). KBase: The United States Department of energy systems biology knowledgebase. *Nature Biotechnology*, *36*(7), 566–569. https://doi.org/10.1038/nbt.4163.

Baroukh, C., Muñoz-Tamayo, R., Steyer, J.-P., & Bernard, O. (2014). DRUM: A new framework for metabolic modeling under non-balanced growth. Application to the carbon metabolism of unicellular microalgae. *PLoS One*, *9*(8), e104499. https://doi.org/10.1371/journal.pone.0104499.

Becker, M. Y., & Rojas, I. (2001). A graph layout algorithm for drawing metabolic pathways. *Bioinformatics*, *17*(5), 461–467. https://doi.org/10.1093/bioinformatics/17.5.461.

Bekaert, M., Edger, P. P., Hudson, C. M., Pires, J. C., & Conant, G. C. (2012). Metabolic and evolutionary costs of herbivory defense: Systems biology of glucosinolate synthesis. *New Phytologist*, *196*(2), 596–605. https://doi.org/10.1111/j.1469-8137.2012.04302.x.

Bogart, E., & Myers, C. R. (2016). Multiscale metabolic modeling of C4 plants: Connecting nonlinear genome-scale models to leaf-scale metabolism in developing maize leaves. *PLoS One*, *11*(3), e0151722. https://doi.org/10.1371/journal.pone.0151722.

Bordbar, A., Lewis, N. E., Schellenberger, J., Palsson, B.Ø., & Jamshidi, N. (2010). Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Molecular Systems Biology*, *6*(1), 422. https://doi.org/10.1038/msb.2010.68.

Botero, K., Restrepo, S., & Pinzón, A. (2018). A genome-scale metabolic model of potato late blight suggests a photosynthesis suppression mechanism. *BMC Genomics*, *19*(S8), 863. https://doi.org/10.1186/s12864-018-5192-x.

Bourqui, R., Cottret, L., Lacroix, V., Auber, D., Mary, P., Sagot, M.-F., et al. (2007). Metabolic network visualization eliminating node redundance and preserving metabolic pathways. *BMC Systems Biology*, *1*(1), 29. https://doi.org/10.1186/1752-0509-1-29.

Caspi, R., Billington, R., Ferrer, L., Foerster, H., Fulcher, C. A., Keseler, I. M., et al. (2016). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Research*, *44*(D1), D471–D480. https://doi.org/10.1093/nar/gkv1164.

Chatterjee, A., Huma, B., Shaw, R., & Kundu, S. (2017). Reconstruction of Oryza sativa indica genome scale metabolic model and its responses to varying RuBisCO activity, light intensity, and enzymatic cost conditions. *Frontiers in Plant Science*, *8*, 2060. https://doi.org/10.3389/fpls.2017.02060.

Chatterjee, A., & Kundu, S. (2015). Revisiting the chlorophyll biosynthesis pathway using genome scale metabolic model of Oryza sativa japonica. *Scientific Reports*, *5*(1), 14975. https://doi.org/10.1038/srep14975.

Chazalviel, M., Frainay, C., Poupin, N., Vinson, F., Merlet, B., Gloaguen, Y., et al. (2018). MetExploreViz: Web component for interactive metabolic network visualization. *Bioinformatics*, *34*(2), 312–313. https://doi.org/10.1093/bioinformatics/btx588.

Cheung, C. Y. M., Poolman, M. G., Fell, D. A., Ratcliffe, R. G., & Sweetlove, L. J. (2014). A diel flux balance model captures interactions between light and dark metabolism during day-night cycles in $C_3$ and Crassulacean acid metabolism leaves. *Plant Physiology*, *165*(2), 917–929. https://doi.org/10.1104/pp.113.234468.

Cheung, C. Y. M., Williams, T. C. R., Poolman, M. G., Fell, D. A., Ratcliffe, R. G., & Sweetlove, L. J. (2013). A method for accounting for maintenance costs in flux balance analysis improves the prediction of plant cell metabolic phenotypes under stress conditions. *The Plant Journal*, *75*(6), 1050–1061. https://doi.org/10.1111/tpj.12252.

Contador, C. A., Lo, S.-K., Chan, S. H. J., & Lam, H.-M. (2020). Metabolic analyses of nitrogen fixation in the soybean microsymbiont *Sinorhizobium fredii* using constraint-based modeling. *MSystems*, *5*(1), e00516-19. /msystems/5/1/msys.00516-19.atom https://doi.org/10.1128/mSystems.00516-19.

de Oliveira Dal'Molin, C. G., Orellana, C., Gebbie, L., Steen, J., Hodson, M. P., Chrysanthopoulos, P., et al. (2016). Metabolic reconstruction of Setaria italica: A systems biology approach for integrating tissue-specific omics and pathway analysis of bioenergy grasses. *Frontiers in Plant Science*, 7. https://doi.org/10.3389/fpls.2016.01138.

de Oliveira Dal'Molin, C. G., Quek, L.-E., Palfreyman, R. W., Brumbley, S. M., & Nielsen, L. K. (2010a). AraGEM, a genome-scale reconstruction of the primary metabolic network in Arabidopsis. *Plant Physiology*, *152*(2), 579–589. https://doi.org/10.1104/pp.109.148817.

de Oliveira Dal'Molin, C. G., Quek, L.-E., Palfreyman, R. W., Brumbley, S. M., & Nielsen, L. K. (2010b). C4GEM, a genome-scale metabolic model to study $C_4$ plant metabolism. *Plant Physiology*, *154*(4), 1871–1885. https://doi.org/10.1104/pp.110.166488.

de Oliveira Dal'Molin, C. G., Quek, L.-E., Saa, P. A., & Nielsen, L. K. (2015). A multi-tissue genome-scale metabolic modeling framework for the analysis of whole plant systems. *Frontiers in Plant Science*, *6*. https://doi.org/10.3389/fpls.2015.00004.

Degani, A. (2013). A tale of two maps: Analysis of the London underground "Diagram". *Ergonomics in Design: The Quarterly of Human Factors Applications*, *21*(3), 7–16. https://doi.org/10.1177/1064804613489125.

diCenzo, G. C., Benedict, A. B., Fondi, M., Walker, G. C., Finan, T. M., Mengoni, A., et al. (2018). Robustness encoded across essential and accessory replicons of the ecologically versatile bacterium Sinorhizobium meliloti. *PLoS Genetics*, *14*(4), e1007357. https://doi.org/10.1371/journal.pgen.1007357.

diCenzo, G. C., Checcucci, A., Bazzicalupo, M., Mengoni, A., Viti, C., Dziewit, L., et al. (2016). Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in Sinorhizobium meliloti. *Nature Communications*, *7*(1), 12219. https://doi.org/10.1038/ncomms12219.

diCenzo, G. C., Tesi, M., Pfau, T., Mengoni, A., & Fondi, M. (2020). Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. *Nature Communications*, *11*(1), 2574. https://doi.org/10.1038/s41467-020-16484-2.

Didimo, W., Liotta, G., & Patrignani, M. (2019). HV-planarity: Algorithms and complexity. *Journal of Computer and System Sciences*, *99*, 72–90. https://doi.org/10.1016/j.jcss.2018.08.003.

Dinkla, K., & Westenberg, M. A. (2012). Network visualization in cell biology. *Tsinghua Science and Technology*, *17*(4), 365–382. https://doi.org/10.1109/TST.2012.6297584.

Frainay, C., Schymanski, E., Neumann, S., Merlet, B., Salek, R., Jourdan, F., et al. (2018). Mind the gap: Mapping mass spectral databases in genome-scale metabolic networks reveals poorly covered areas. *Metabolites*, *8*(3), 51. https://doi.org/10.3390/metabo8030051.

Gehlenborg, N., O'Donoghue, S. I., Baliga, N. S., Goesmann, A., Hibbs, M. A., Kitano, H., et al. (2010). Visualization of omics data for systems biology. *Nature Methods*, *7*(S3), S56–S68. https://doi.org/10.1038/nmeth.1436.

Gerlin, L., Cottret, L., Cesbron, S., Taghouti, G., Jacques, M.-A., Genin, S., et al. (2020). Genome-scale investigation of the metabolic determinants generating bacterial fastidious growth. *MSystems*, *5*(2), e00698-19. /msystems/5/2/msys.00698-19.atom https://doi.org/10.1128/mSystems.00698-19.

Hay, J. O., Shi, H., Heinzel, N., Hebbelmann, I., Rolletschek, H., & Schwender, J. (2014). Integration of a constraint-based metabolic model of Brassica napus developing seeds with 13C-metabolic flux analysis. *Frontiers in Plant Science*, *5*. https://doi.org/10.3389/fpls.2014.00724.

Herman, I., Melancon, G., & Marshall, M. S. (2000). Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics*, *6*(1), 24–43. https://doi.org/10.1109/2945.841119.

Johnson, S. R., Lange, I., Srividya, N., & Lange, B. M. (2017). Bioenergetics of monoterpenoid essential oil biosynthesis in nonphotosynthetic glandular trichomes. *Plant Physiology*, *175*(2), 681–695. https://doi.org/10.1104/pp.17.00551.

Kanehisa, M. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27–30. https://doi.org/10.1093/nar/28.1.27.

Kim, T. Y., Sohn, S. B., Kim, Y. B., Kim, W. J., & Lee, S. Y. (2012). Recent advances in reconstruction and applications of genome-scale metabolic models. *Current Opinion in Biotechnology*, *23*(4), 617–623. https://doi.org/10.1016/j.copbio.2011.10.007.

Koduru, L., Kim, H. Y., Lakshmanan, M., Mohanty, B., Lee, Y. Q., Lee, C. H., et al. (2020). Genome-scale metabolic reconstruction and in silico analysis of the rice leaf blight pathogen, *Xanthomonas oryzae*. *Molecular Plant Pathology*, *21*(4), 527–540. https://doi.org/10.1111/mpp.12914.

Krzywinski, M., Birol, I., Jones, S. J., & Marra, M. A. (2012). Hive plots—Rational approach to visualizing networks. *Briefings in Bioinformatics*, *13*(5), 627–644. https://doi.org/10.1093/bib/bbr069.

Lakshmanan, M., Lim, S.-H., Mohanty, B., Kim, J. K., Ha, S.-H., & Lee, D.-Y. (2015). Unraveling the light-specific metabolic and regulatory signatures of rice through combined in silico modeling and multi-omics analysis. *Plant Physiology*, 01379.2015. https://doi.org/10.1104/pp.15.01379.

Liu, L., Mei, Q., Yu, Z., Sun, T., Zhang, Z., & Chen, M. (2013). *An integrative bioinformatics framework for genome-scale multiple level network reconstruction of rice*. https://doi.org/10.2390/BIECOLL-JIB-2013-223.

Michal, G. (1999). *Biochemical pathways*. John Wiley & Sons, Inc.

Mintz-Oron, S., Meir, S., Malitsky, S., Ruppin, E., Aharoni, A., & Shlomi, T. (2012). Reconstruction of Arabidopsis metabolic network models accounting for subcellular compartmentalization and tissue-specificity. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(1), 339–344. https://doi.org/10.1073/pnas.1100358109.

Moreira, T. B., Shaw, R., Luo, X., Ganguly, O., Kim, H.-S., Coelho, L. G. F., et al. (2019). A genome-scale metabolic model of soybean (*Glycine max*) highlights metabolic fluxes in seedlings. *Plant Physiology*, *180*(4), 1912–1929. https://doi.org/10.1104/pp.19.00122.

Mulcahy, K. A., & Clarke, K. C. (2001). Symbolization of map projection distortion: A review. *Cartography and Geographic Information Science*, *28*(3), 167–182. https://doi.org/10.1559/152304001782153044.

Nicholson, D. (2005). From metabolic pathways charts to animaps in 50 years. *Biochemistry and Molecular Biology Education*, *33*(3), 156–158. https://doi.org/10.1002/bmb.2005.494033032480.

Okuda, S., Yamada, T., Hamajima, M., Itoh, M., Katayama, T., Bork, P., et al. (2008). KEGG atlas mapping for global analysis of metabolic pathways. *Nucleic Acids Research*, *36*, W423–W426 (web server) https://doi.org/10.1093/nar/gkn282.

Orth, J. D., Thiele, I., & Palsson, B.Ø. (2010). What is flux balance analysis? *Nature Biotechnology*, *28*(3), 245–248. https://doi.org/10.1038/nbt.1614.

Peyraud, R., Cottret, L., Marmiesse, L., Gouzy, J., & Genin, S. (2016). A resource allocation trade-off between virulence and proliferation drives metabolic versatility in the plant pathogen Ralstonia solanacearum. *PLoS Pathogens*, *12*(10), e1005939. https://doi.org/10.1371/journal.ppat.1005939.

Peyraud, R., Mbengue, M., Barbacci, A., & Raffaele, S. (2019). Intercellular cooperation in a fungal plant pathogen facilitates host colonization. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(8), 3193–3201. https://doi.org/10.1073/pnas.1811267116.

Pfau, T., Christian, N., Masakapalli, S. K., Sweetlove, L. J., Poolman, M. G., & Ebenhöh, O. (2018). The intertwined metabolism during symbiotic nitrogen fixation elucidated by metabolic modelling. *Scientific Reports*, *8*(1), 12504. https://doi.org/10.1038/s41598-018-30884-x.

Poolman, M. G., Kundu, S., Shaw, R., & Fell, D. A. (2013). Responses to light intensity in a genome-scale model of rice metabolism. *Plant Physiology*, *162*(2), 1060–1072. https://doi.org/10.1104/pp.113.216762.

Poolman, M. G., Miguet, L., Sweetlove, L. J., & Fell, D. A. (2009). A genome-scale metabolic model of Arabidopsis and some of its properties. *Plant Physiology*, *151*(3), 1570–1581. https://doi.org/10.1104/pp.109.141267.

Prigent, S., Frioux, C., Dittami, S. M., Thiele, S., Larhlimi, A., Collet, G., et al. (2017). Meneco, a topology-based gap-filling tool applicable to degraded genome-wide metabolic networks. *PLoS Computational Biology*, *13*(1), e1005276. https://doi.org/10.1371/journal.pcbi.1005276.

Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A., et al. (2019). Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nature Protocols*, *14*(2), 482–517. https://doi.org/10.1038/s41596-018-0103-9.

Resendis-Antonio, O., Reed, J. L., Encarnación, S., Collado-Vides, J., & Palsson, B.Ø. (2007). Metabolic reconstruction and modeling of nitrogen fixation in Rhizobium etli. *PLoS Computational Biology*, *3*(10), e192. https://doi.org/10.1371/journal.pcbi.0030192.

Rodenburg, S. Y. A., Seidl, M. F., de Ridder, D., & Govers, F. (2018). Genome-wide characterization of *Phytophthora infestans* metabolism: A systems biology approach: A metabolic model for *Phytophthora infestans*. *Molecular Plant Pathology*, *19*(6), 1403–1413. https://doi.org/10.1111/mpp.12623.

Rodenburg, S. Y. A., Seidl, M. F., Judelson, H. S., Vu, A. L., Govers, F., & de Ridder, D. (2019). Metabolic model of the *Phytophthora infestans*-tomato interaction reveals metabolic switches during host colonization. *MBio*, *10*(4). e00454-19, /mbio/10/4/mBio.00454-19.atom https://doi.org/10.1128/mBio.00454-19.

Saha, R., Suthers, P. F., & Maranas, C. D. (2011). Zea mays iRS1563: A comprehensive genome-scale metabolic reconstruction of maize metabolism. *PLoS One*, *6*(7), e21784. https://doi.org/10.1371/journal.pone.0021784.

Salimi, F., Zhuang, K., & Mahadevan, R. (2010). Genome-scale metabolic modeling of a clostridial co-culture for consolidated bioprocessing. *Biotechnology Journal*, *5*(7), 726–738. https://doi.org/10.1002/biot.201000159.

Saraiya, P., North, C., & Duca, K. (2005). Visualizing biological pathways: Requirements analysis, systems evaluation and research agenda. *Information Visualization*, *4*(3), 191–205. https://doi.org/10.1057/palgrave.ivs.9500102.

Sarkar, D., & Maranas, C. D. (2020). SNPeffect: Identifying functional roles of SNPs using metabolic networks. *The Plant Journal*. tpj.14746 https://doi.org/10.1111/tpj.14746.

Scheunemann, M., Brady, S. M., & Nikoloski, Z. (2018). Integration of large-scale data for extraction of integrated Arabidopsis root cell-type specific models. *Scientific Reports*, *8*(1), 7919. https://doi.org/10.1038/s41598-018-26232-8.

Schläpfer, P., Zhang, P., Wang, C., Kim, T., Banf, M., Chae, L., et al. (2017). Genome-wide prediction of metabolic enzymes, pathways, and gene clusters in plants. *Plant Physiology*, *173*(4), 2041–2059. https://doi.org/10.1104/pp.16.01942.

Schuster, S., Pfeiffer, T., Moldenhauer, F., Koch, I., & Dandekar, T. (2002). Exploring the pathway structure of metabolism: Decomposition into subnetworks and application to mycoplasma pneumoniae. *Bioinformatics*, *18*(2), 351–361. https://doi.org/10.1093/bioinformatics/18.2.351.

Seaver, S. M. D., Bradbury, L. M. T., Frelin, O., Zarecki, R., Ruppin, E., Hanson, A. D., et al. (2015). Improved evidence-based genome-scale metabolic models for maize leaf, embryo, and endosperm. *Frontiers in Plant Science*, *6*. https://doi.org/10.3389/fpls.2015.00142.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*(11), 2498–2504. https://doi.org/10.1101/gr.1239303.

Shaw, R., & Cheung, C. Y. M. (2018). A dynamic multi-tissue flux balance model captures carbon and nitrogen metabolism and optimal resource partitioning during Arabidopsis growth. *Frontiers in Plant Science*, *9*, 884. https://doi.org/10.3389/fpls.2018.00884.

Simons, M., Saha, R., Amiour, N., Kumar, A., Guillard, L., Clement, G., et al. (2014). Assessing the metabolic impact of nitrogen availability using a compartmentalized maize leaf genome-scale model. *Plant Physiology*, *166*(3), 1659–1674. https://doi.org/10.1104/pp.114.245787.

Stolyar, S., Van Dien, S., Hillesland, K. L., Pinel, N., Lie, T. J., Leigh, J. A., et al. (2007). Metabolic modeling of a mutualistic microbial community. *Molecular Systems Biology*, *3*(1), 92. https://doi.org/10.1038/msb4100131.

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, *408*(6814), 796–815. https://doi.org/10.1038/35048692.

Thiele, I., & Palsson, B.Ø. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols*, *5*(1), 93–121. https://doi.org/10.1038/nprot.2009.203.

Thiele, I., Vlassis, N., & Fleming, R. M. T. (2014). fastGapFill: Efficient gap filling in metabolic networks. *Bioinformatics*, *30*(17), 2529–2531. https://doi.org/10.1093/bioinformatics/btu321.

Tufte, E. R. (1997). *Visual explanations: Images and quantities, evidence and narrative*. Graphics Press.

von Landesberger, T., Kuijper, A., Schreck, T., Kohlhammer, J. J., van Wijk, J., Fekete, J.-D., et al. (2011). Visual analysis of large graphs: State-of-the-art and future research challenges. *Computer Graphics Forum*, *30*(6), 1719–1749. https://doi.org/10.1111/j.1467-8659.2011.01898.x.

Wang, C., Deng, Z.-L., Xie, Z.-M., Chu, X.-Y., Chang, J.-W., Kong, D.-X., et al. (2015). Construction of a genome-scale metabolic network of the plant pathogen *Pectobacterium carotovorum* provides new strategies for bactericide discovery. *FEBS Letters*, *589*(3), 285–294. https://doi.org/10.1016/j.febslet.2014.12.010.

Ware, C., Purchase, H., Colpoys, L., & McGill, M. (2002). Cognitive measurements of graph aesthetics. *Information Visualization*, *1*(2), 103–110. https://doi.org/10.1057/palgrave.ivs.9500013.

Wiechert, W. (2001). 13C metabolic flux analysis. *Metabolic Engineering*, *3*(3), 195–206. https://doi.org/10.1006/mben.2001.0187.

Yoghourdjian, V., Archambault, D., Diehl, S., Dwyer, T., Klein, K., Purchase, H. C., et al. (2018). Exploring the limits of complexity: A survey of empirical studies on graph visualisation. *Visual Informatics*, *2*(4), 264–282. https://doi.org/10.1016/j.visinf.2018.12.006.

Yuan, H., Cheung, C. Y. M., Poolman, M. G., Hilbers, P. A. J., & van Riel, N. A. W. (2016). A genome-scale metabolic network reconstruction of tomato (*Solanum lycopersicum* L.) and its application to photorespiratory metabolism. *The Plant Journal*, *85*(2), 289–304. https://doi.org/10.1111/tpj.13075.

Zoledowska, S., Presta, L., Fondi, M., Decorosi, F., Giovannetti, L., Mengoni, A., et al. (2019). Metabolic modeling of Pectobacterium parmentieri SCC3193 provides insights into metabolic pathways of plant pathogenic Bacteria. *Microorganisms*, 7(4), 101. https://doi.org/10.3390/microorganisms7040101.

Zomorrodi, A. R., Islam, M. M., & Maranas, C. D. (2014). D-OptCom: Dynamic multi-level and multi-objective metabolic modeling of microbial communities. *ACS Synthetic Biology*, *3*(4), 247–257. https://doi.org/10.1021/sb4001307.

Zomorrodi, A. R., & Maranas, C. D. (2012). OptCom: A multi-level optimization framework for the metabolic modeling and analysis of microbial communities. *PLoS Computational Biology*, *8*(2), e1002363. https://doi.org/10.1371/journal.pcbi.1002363.

# Appendix II – Supplementary Materials of "A multi-organ metabolic model of tomato predicts plant responses to nutritional and genetic perturbations"

**Supplementary Figure 1.** Additional physiological data. Fresh:dry weight ratio for each organ, growth curves with ln(dry weight), linear regression of fresh/dry weights and organ weights, daily variation of transpiration.

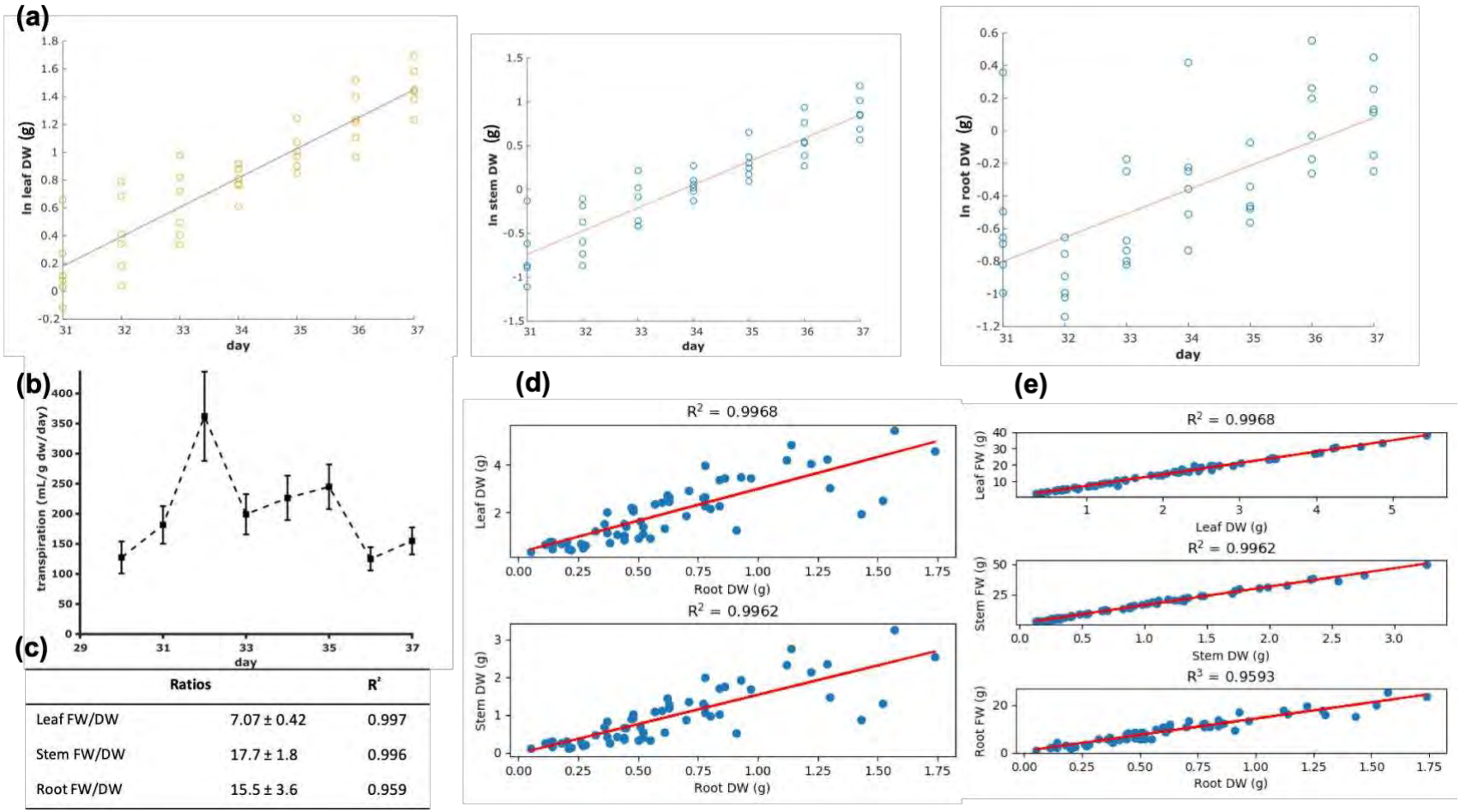
**Supplementary Figure 2.** Additional xylem metabolomics data. Xylem constitution at each sampling day.

**Supplementary Figure 3.** Additional organ metabolomics data. Biomass constitution at each sampling day.

**Supplementary Figure 4.** Additional representation of model responses to plant physical variations. Metabolic fluxes in photosystem II as a function of stem contribution to photosynthesis.

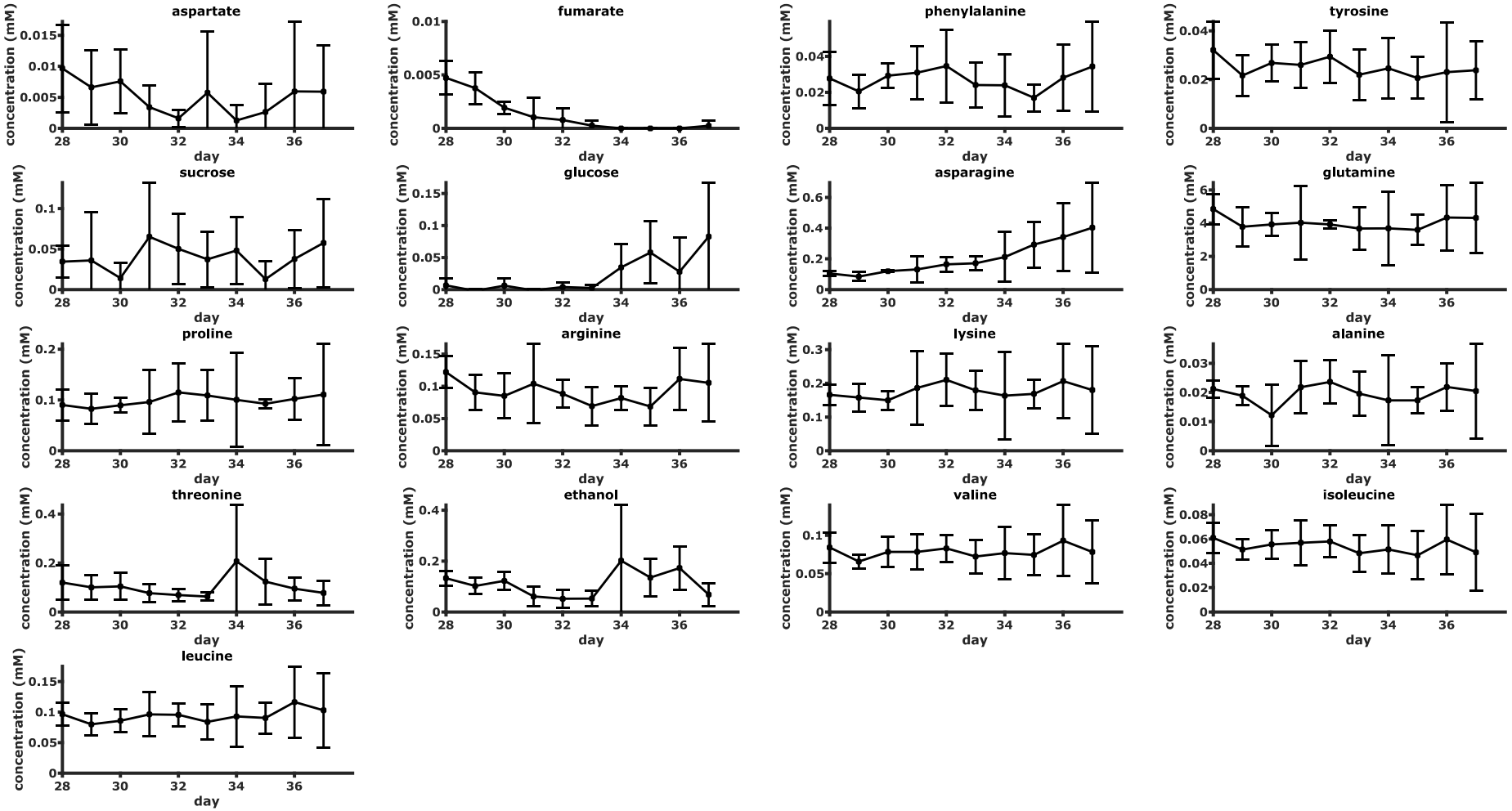
**Supplementary Figure 5.** Flow chart of the multi-organ modeling pipeline.

**Supplementary Text 1.** Additional analyses. Effects of ATP cost variation, Impact of physiological changes on the plant at high percentages, Impact of the uptake ratio between nitrates and ammonium, Stoichiometric Balance Cycles.

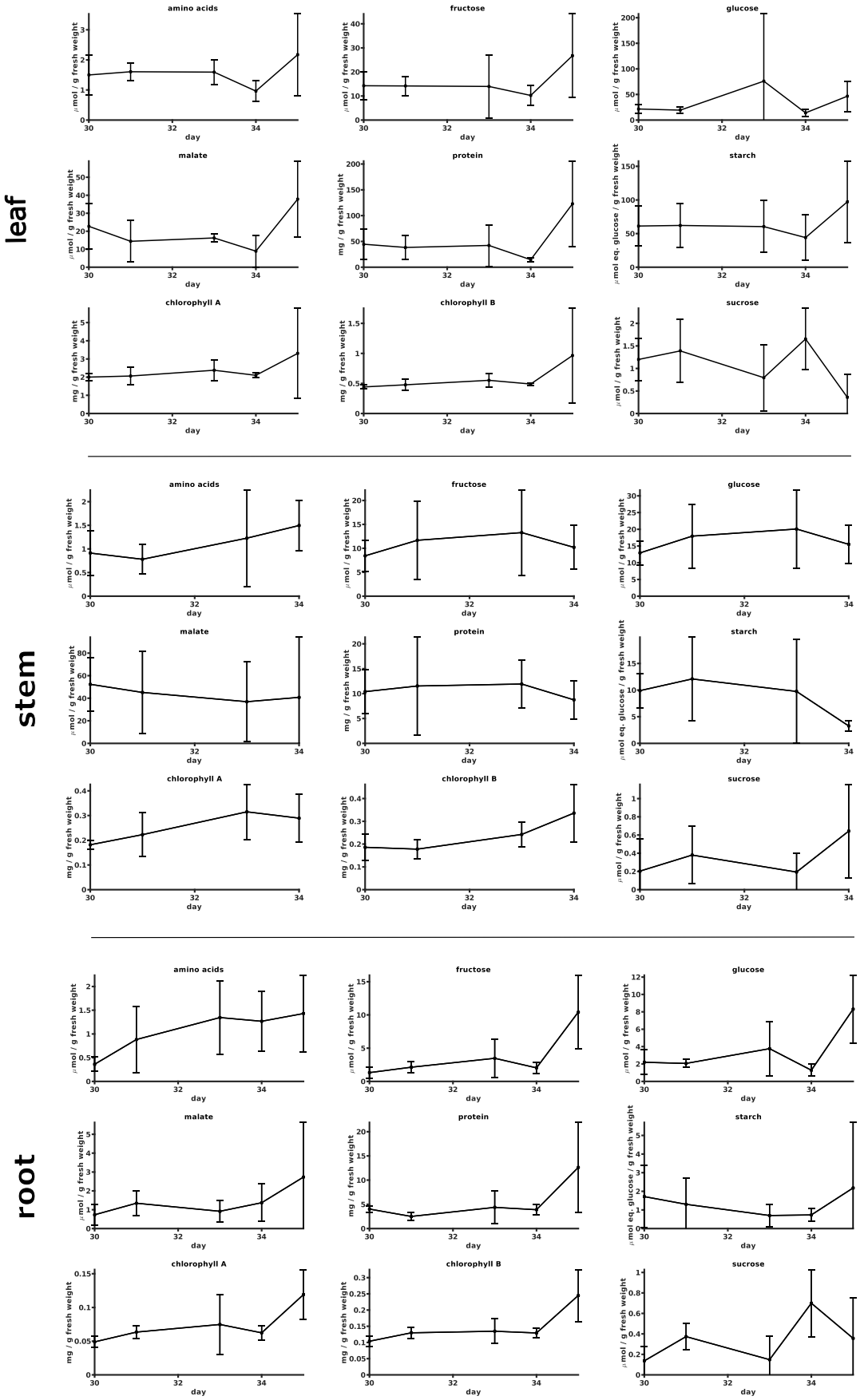



| Ratios | | $R^2$ |
|---|---|---|
| Leaf FW/DW | 7.07 ± 0.42 | 0.997 |
| Stem FW/DW | 17.7 ± 1.8 | 0.996 |
| Root FW/DW | 15.5 ± 3.6 | 0.959 |

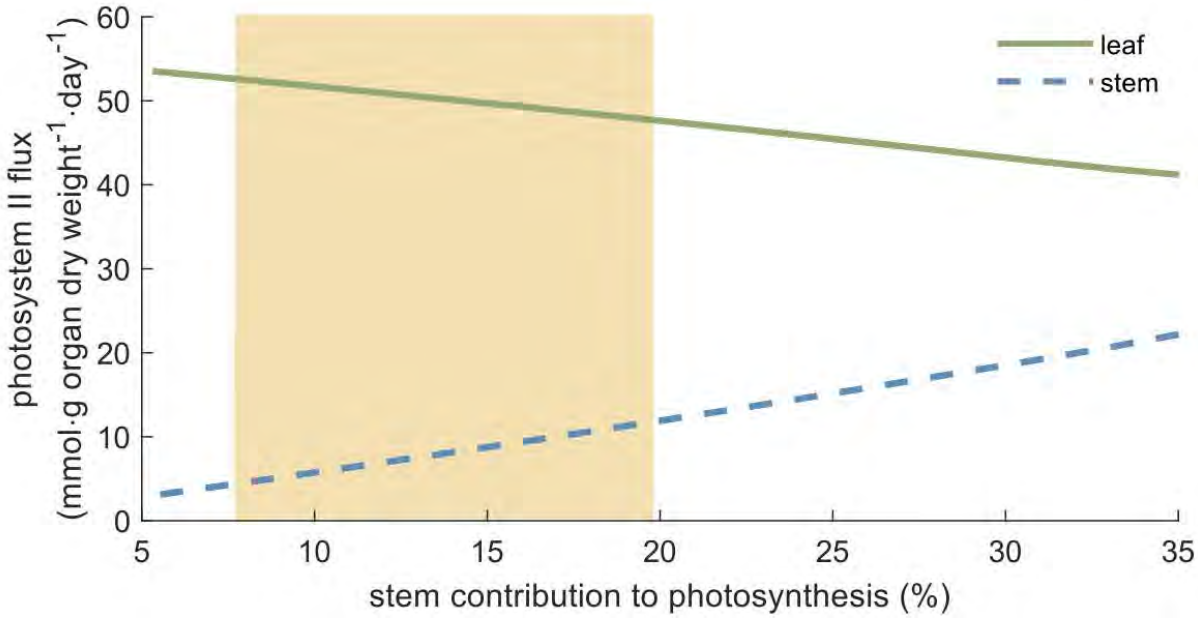
**Supplemental Figure S1. Additional physiological data.** (a): Logarithm regression of fresh weight over time for respectively leaf, stem and root. Theses curves allowed to infer growth rate of each organ. (b) Transpiration of plants over time normalized by biomass weight. Values represented are mean ± standard deviation. Sample size: 6 plants per day (c) Fresh:dry weight ratio for each organ (leaf, stem, root). Values represented are mean ± standard deviation. Sample size: 54 plants (d) Linear regression between dry weight of leaf (resp. stem) and root (e) Linear regression between fresh and dry weight for each organ (leaf, stem, root).



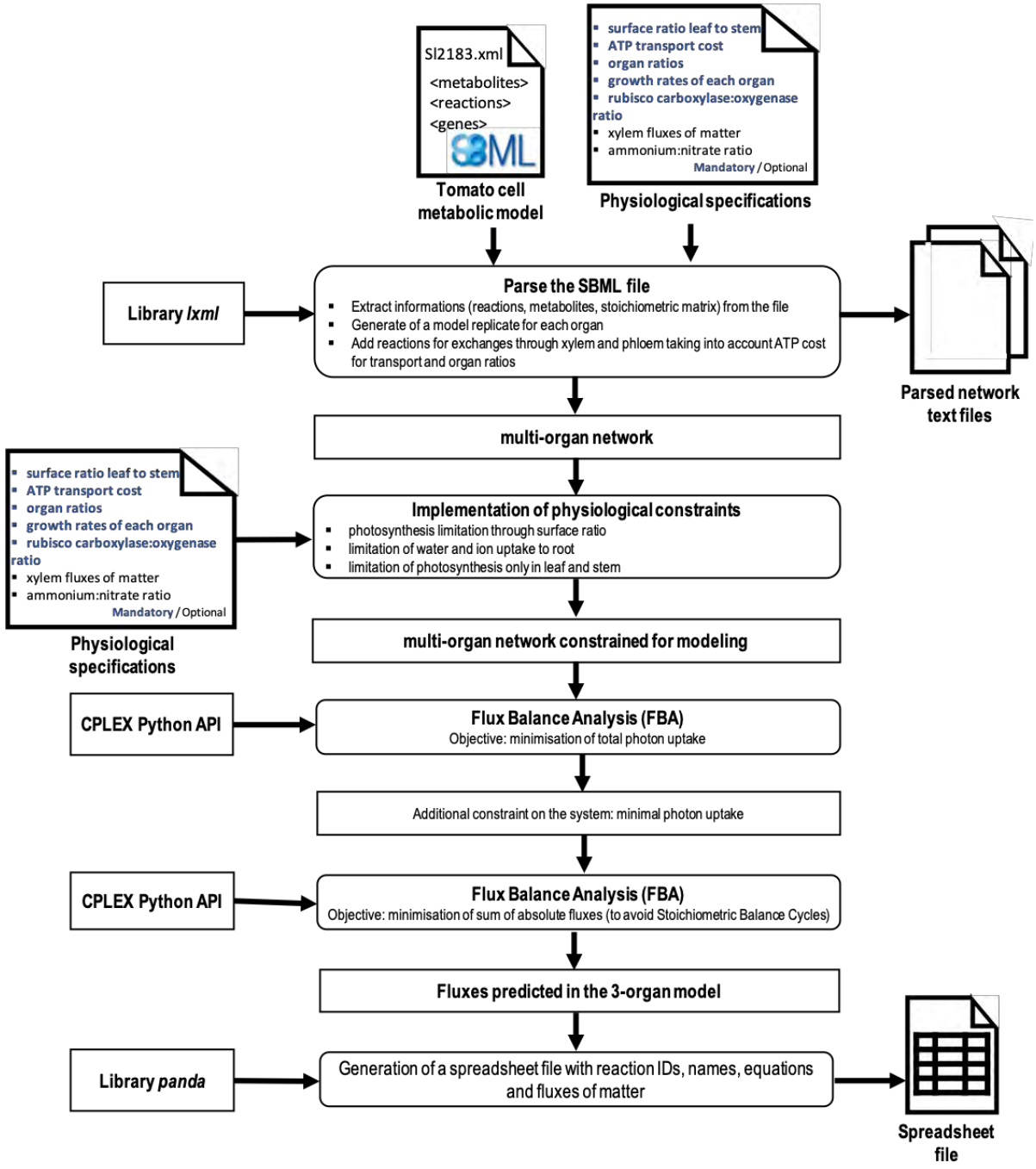

**Supplemental Figure S2. Additional xylem metabolomics data.** Organic xylem constitution at each sampling day. Values are mean ± standard deviation. Sample size: 33 plants.




**Supplemental Figure S3. Additional organ metabolomics data.** Biomass constitution at each sampling day. Values are mean ± standard deviation. Sample size: at least 19 plants.




**Supplemental Figure S4. Additional representation of model responses to plant physical variations.** Metabolic fluxes in photosystem II as a function of stem contribution to photosynthesis.




**Supplemental Figure S5. Flow chart of the multi-organ modeling pipeline.**

# Supplemental Text: additional analyses

## 1. Effect of ATP cost variation

In the multi-organ metabolic model, a transport cost (ATP cost) between organ was added. Transporting k µmol/g organ dry weight/day would consume k·(transport cost) µmol/g organ dry weigh/day of ATP. In the modeling results presented in the paper, the ATP cost is fixed to 0.5. This cost could impact our conclusions, so we provide here the effects of a reasonable change of this variable on our prediction for two use cases.

### *Xylem composition* (Fig. A)

Xylem composition is directly impacted by the transport cost, so it is the use case the most sensitive to this value. We showed in the result section that we were able to predict, without any constraint other than the ATP cost, that glutamine was the major organic element of xylem sap. We were also able to predict the presence of other amino acid in a more diluted proportion. This conclusion is still true for all the ATP costs tested between 0.02 and 1.5. At the peculiar case of no ATP cost (0) and at very high ATP cost (>1.5), the predictions are no longer consistent. Our conclusions on xylem sap composition are thus robust to a reasonable variation of this parameter.

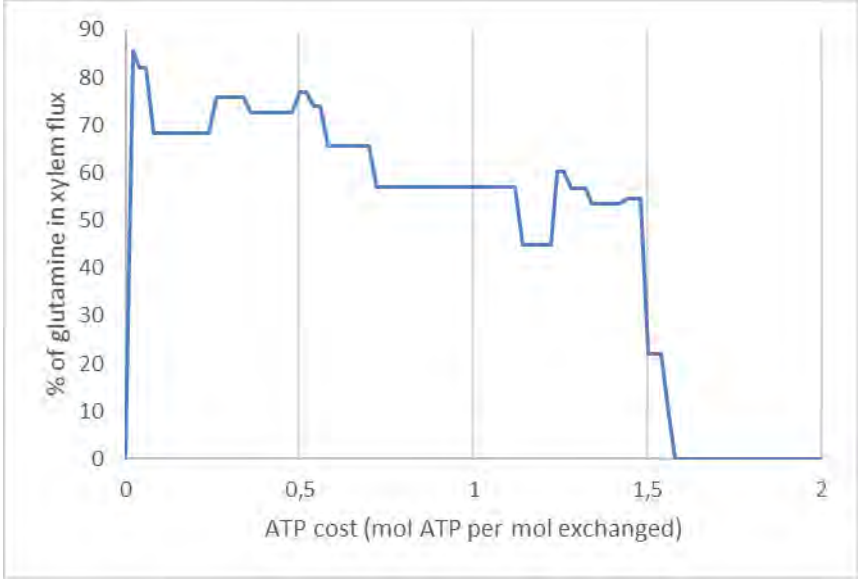



**Figure A – Glutamine content in the xylem for different ATP costs.**

### *Impact of nitrogen limitation on growth*

Use case 4 in the main article presents the impact of nitrogen limitation on the plant growth rate. A linear decrease of the growth rate related to the decrease of the assimilation of nitrogen is observed, consistently with experimental data. This result holds for any ATP cost value between 0 and 4 tested (Figure B).

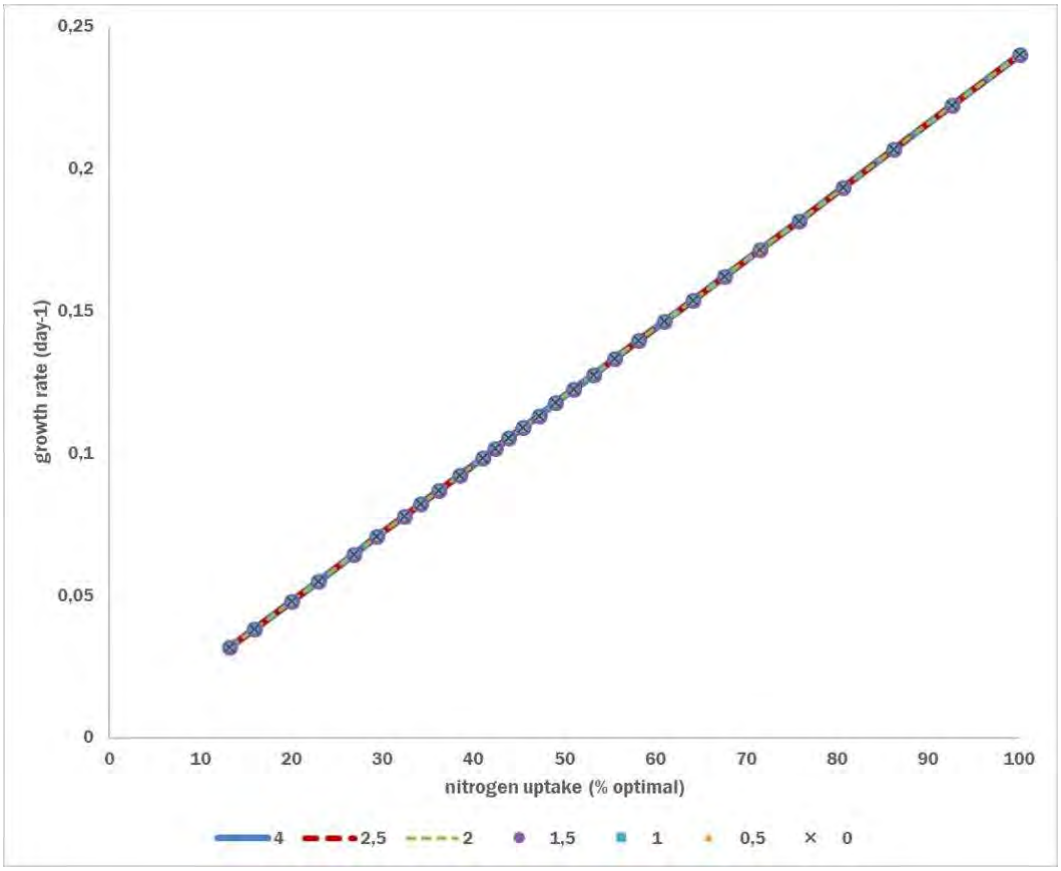



**Figure B – Effect of nitrogen uptake on plant growth rate at different ATP costs**. Details are provided in the main article, section Use case 4.

## 2. Impact of physiological changes on the plant at high percentages

Stem contribution to photosynthesis and stem weight impact on plant metabolism was tested for a limited range of values that we considered as biologically relevant. Below (Figure C and D) are shown simulations results for a wider range of values.

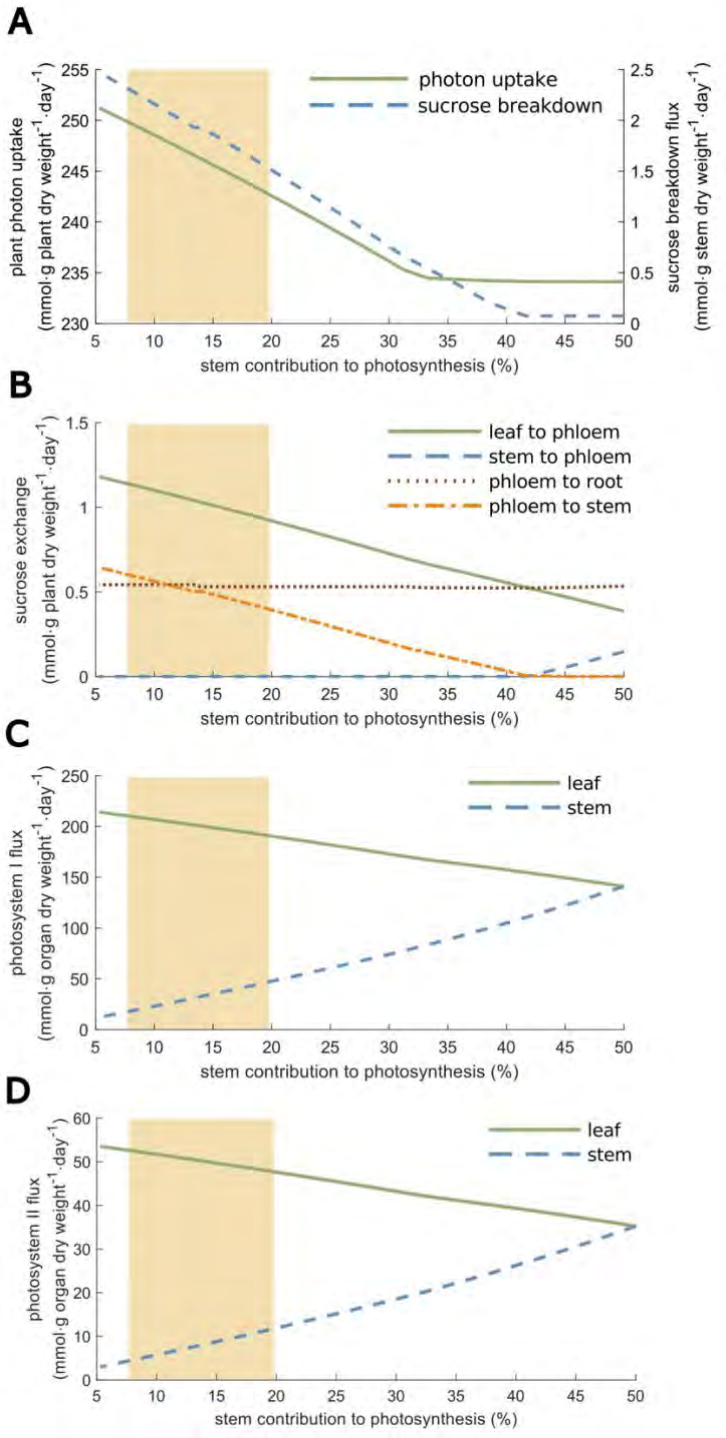



**Figure C – Fluxes prediction of photosynthesis-related reactions (photon uptake, sucrose exchange, photosystem I, photosystem II) in response to different leaf:stem photosynthetic capabilities.** The yellow area represents the range of experimentally observed values: [mean – standard deviation ; mean + standard deviation]. A: Effect of stem contribution to photosynthesis on plant photon uptake and sucrose breakdown fluxes. B: Effect of stem contribution to photosynthesis on sucrose exchange reactions between organs. C and D. Effect of stem contribution to photosynthesis on the photosystem I (C) and II (D) flux.

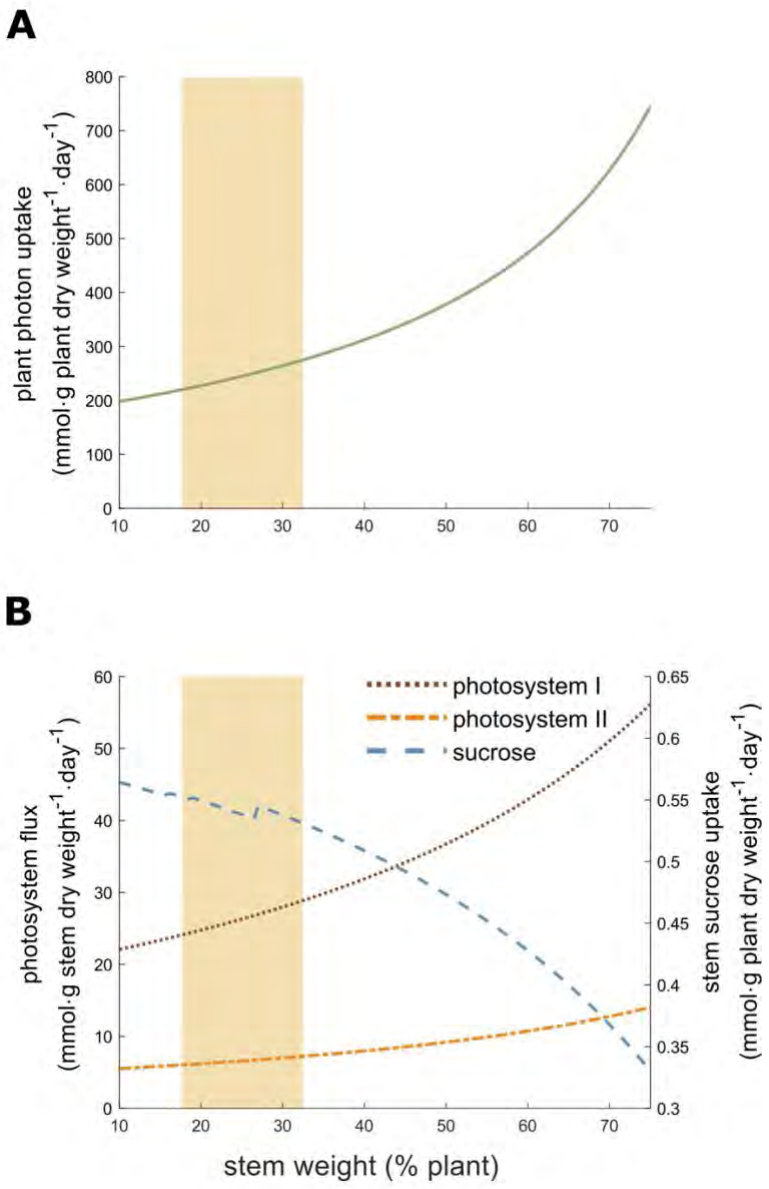



**Figure D - Effect of stem weight on photon uptake (a), sucrose uptake and stem photosystem fluxes (b).** The yellow area represents the range of experimentally observed values: mean ± standard deviation.

## 3. Impact of the uptake ratio between nitrates and ammonium

Uptake ratio of ammonium:nitrate ($NH_4^+$:$NO_3^-$) in root was not constrained in our simulations (except for use case 4 for which there was only nitrates, in agreement with the experimental data). However, the model allows to study the impact of the nitrogen source on plant metabolism (Figure E). The simulation results show that xylem content differs depending on the $NH_4^+$:$NO_3^-$ uptake ratio. The experimentally observed xylem organic content, such as glutamine and asparagine, is achieved with simulations where ammonium input is high but decreases with a $NH_4^+$:$NO_3^-$ ratio below 0.25. We also see that photon demand increases when $NH_4^+$ uptake is reduced. Results are consistent on a qualitative basis with a study on the impact of the nitrogen source on tomato xylem sap composition performed by [3]. Indeed, the authors found a significant decrease in the overall amino acid content and particularly glutamine when $NH_4^+$ supply is reduced and $NO_3^-$ increased.

If there is no constraint on the nitrogen uptake source, the model predicts a $NH_4^+$:$NO_3^-$ uptake ratio of 121.

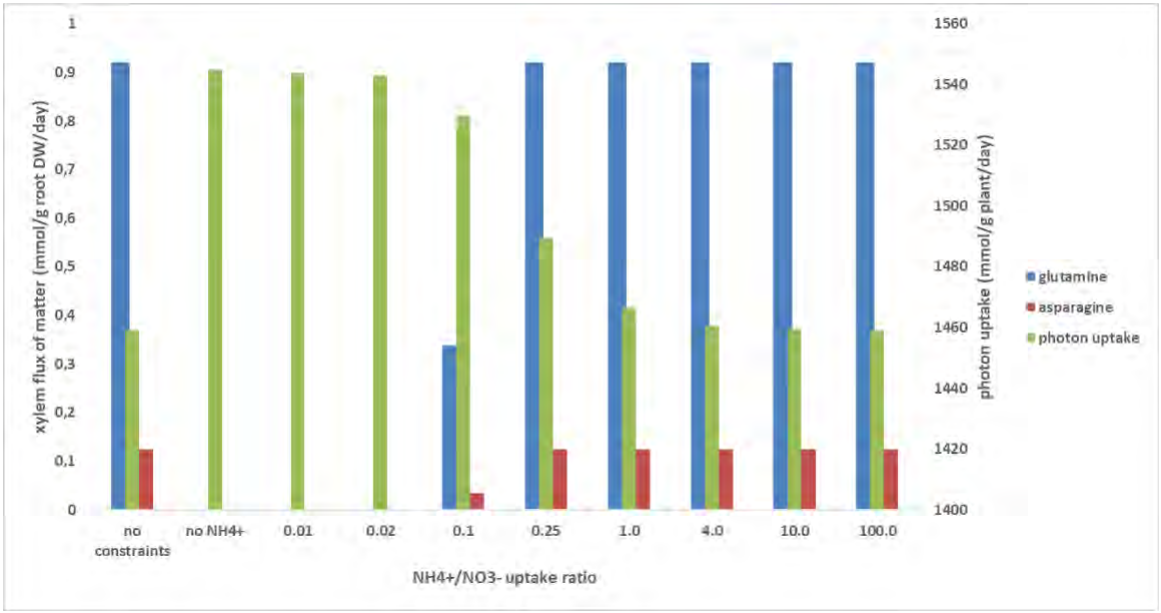



**Figure E - Effect of stem weight on photon uptake (a), sucrose uptake and stem photosystem fluxes (b).** The yellow area represents the range of experimentally observed values: mean ± standard deviation.

**Reference:** Bialczyk, J., et al. "Composition of the Xylem Sap of Tomato Seedlings Cultivated on Media with HCO 3 − and Nitrogen Source as NO 3 − or NH 4 +." *Plant and Soil 2004 263:1*, vol. 263, no. 1, Springer, 2004, pp. 265–72, doi:10.1023/B:PLSO.0000047739.11698.CA.

## 4. Stoichiometric Balance Cycles

We analyzed the potential Stoichiometric Balance Cycles (SBCs) potentially generated by our model, and found that it was present in 194 reactions of the cell model (in more than 2000 reactions). We manually curated some of the cycles and reached 175 reactions with SBCs. The remaining cycles cannot be removed as they are present because of shuttles between organelles or the use of either nad/nadh or nadp/nadph as electron donor for the enzyme complex of a reaction. We checked that these reactions do not impact our simulations since they are not artificially generating free matter or energy or interfering with our analyses of the model. We also verified that they were all removed by the flux minimization second objective function that is performed for each simulation.

# Appendix III – Supplementary Materials of "Unravelling physiological signatures of tomato bacterial wilt and xylem metabolites exploited by *Ralstonia solanacearum*"

**Supplementary Text**

## Generation of *R. solanacearum* triple mutant strain defective for glucose, sucrose and fructose assimilation

Strain GRS941 was engineered by cumulating deletions in the glucose transport operon (RSp1632-RSp1635), the sucrose assimilation operon (RSp1280-RSp1286) and the fructose transport operon (RSc2861-RSc2863).

An unmarked internal deletion in of the glucose transport operon (ΔRSp1632-RSp1635) gene was created using the *sacB*-mediated counter selection system (Schäfer et al., 1994). Briefly, upstream (RSp1632) and downstream (RSp1635) regions were PCR amplified using the primer pairs 1632A/1632B and 1635C/1635D, respectively, and cloned into the *Eco*RI-*Hin*dIII-digested pk18mobsacb to generate pK18-ΔRSp1632/1635 (listed below). The circularized
plasmid was recombined in strain GMI1000 through natural transformation (Perrier et al., 2018).

Kanamycin-resistant and sucrose-sensitive recombinant clones were first selected and in a second step, after overnight culture in BG medium (Plener et al., 2010), kanamycin-sensitive and sucrose-resistant clones were screened by PCR using the primer pair 1632A/1635D to identify a ΔRSp1632-RSp1635 recombinant. The resulting mutant was named GRS903.

A deletion of the RSp1280-RSp1286 region was created by double recombination using a selectable antibiotic interposon cassette from pHP45 (Prentki & Krisch, 1984): upstream (RSp1286) and downstream (RSp1280) regions were PCR amplified using the primer pairs 1280R1 /1280Sc and ScrR-Xba/ScrR-R1, respectively, cloned in the pGEMT vector (Promega), and then the Ω interposon was inserted in the unique *Eco*RI site. The resulting plasmid pSG804 was recombined in strain GRS903 through natural transformation as described above, and spectinomycin-resistant clones were selected and PCR-validated, yielding strain GRS936.

A deletion of the *fruBKA* operon (RSc2861-RSc2863) was created by double recombination using the selectable integrative plasmid pCM184 (Marx & Lindstrom, 2002): upstream (RSc2861) and downstream (RSc2863) regions were PCR amplified using the primer pairs fru-bgl-2/fru-nde-2 and fru-Sac2/fru-Sc1, respectively, and cloned in pCM184 as *Bgl*II-*Nde*I and *Sac*II-*Sac*I inserts, respectively. This recombinant plasmid, pSG753, was recombined in strain GRS936 through natural transformation as described above, then gentamycin-resistant clones were selected and PCR-validated, yielding strain GRS941.

List of oligonucleotides used in this study

| | |
|---|---|
| 1632A | GAATTCAGGAGACAGTCCATGGAGCTC |
| 1632B | TCTAGATCAGCGCGCCGGTGGTG |
| 1635C | TCTAGATCTCCGTCACGGGCATG |
| 1635D | AAGCTTGACGATCATCTGCCACG |
| 1280R1 | GCCTGCTCACGCGCCCGGAGG |
| 1280Sc | GAGCTCGCCCAACAAGTTCGAG |
| ScrR-Xba | TCTAGATCAATTCGTACACCAC |
| ScrR-R1 | GAATTCAACCTCGGCGCTATTGC |
| fru-bgl-2 | AGATCTGCTGGTGCGGCCGAACAGC |
| fru-nde-2 | CATATGTGCTCGTCCGACTGCG |
| fru-Sac2 | CTGGCGCGCATGCTGGACAG |
| fru-Sc1 | GAGCTCGGCGTGTAGACCTGGC |

## References


Marx CJ, Lidstrom ME (2002) Broad-host-range cre-lox system for antibiotic marker recycling in gram-negative bacteria. Biotechniques 33:1062-1067. doi: 10.2144/02335rr01.

Perrier A, Barberis P, Genin S (2018) Introduction of genetic material in *Ralstonia solanacearum* through natural transformation and conjugation. Methods Mol Biol. 1734:201-207. doi: 10.1007/978-1-4939-7604-1_16.

Plener L, Manfredi P, Valls M, Genin S (2010) PrhG, a transcriptional regulator responding to growth conditions, is involved in the control of the type III secretion system regulon in *Ralstonia solanacearum*. J Bacteriol. 192:1011-1019. doi: 10.1128/JB.01189-09.

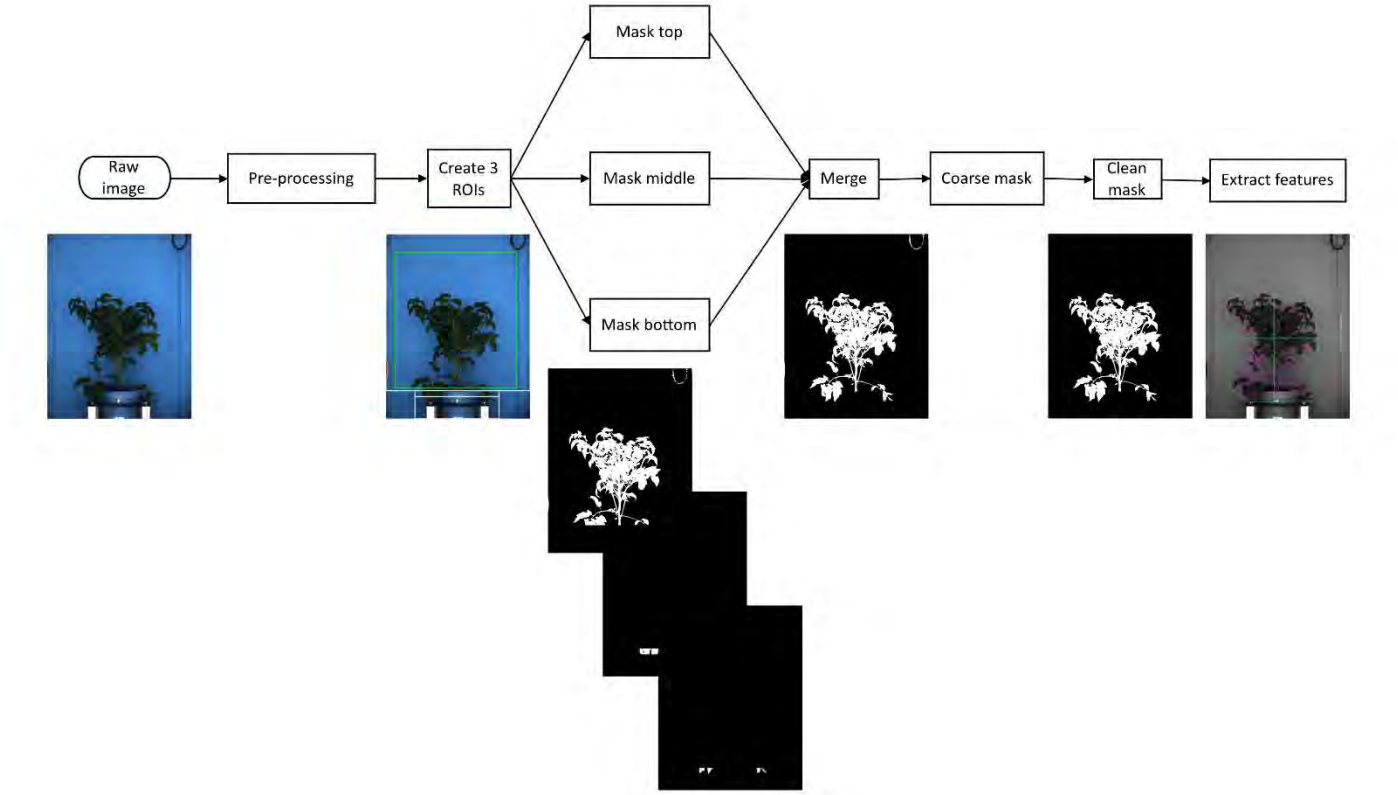
Prentki P, Krisch HM (1984) *In vitro* insertional mutagenesis with a selectable DNA fragment. Gene 29:303-313. doi: 10.1016/0378-1119(84)90059-3.

Schäfer A, Tauch A, Jäger W, Kalinowski J, Thierbach G, Pühler A (1994) Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene 145: 69–73.

## Supplementary Figures and Tables

### Figure S1. Step-by-step segmentation pipeline.

Are represented the major steps of the segmentation process. The JSON script of this specific pipeline is available as supplementary file 1 and can be uploaded, run and modified using the IPSO-Phen software package available at Github *https://github.com/tpmp-inra/ipso_phen*.

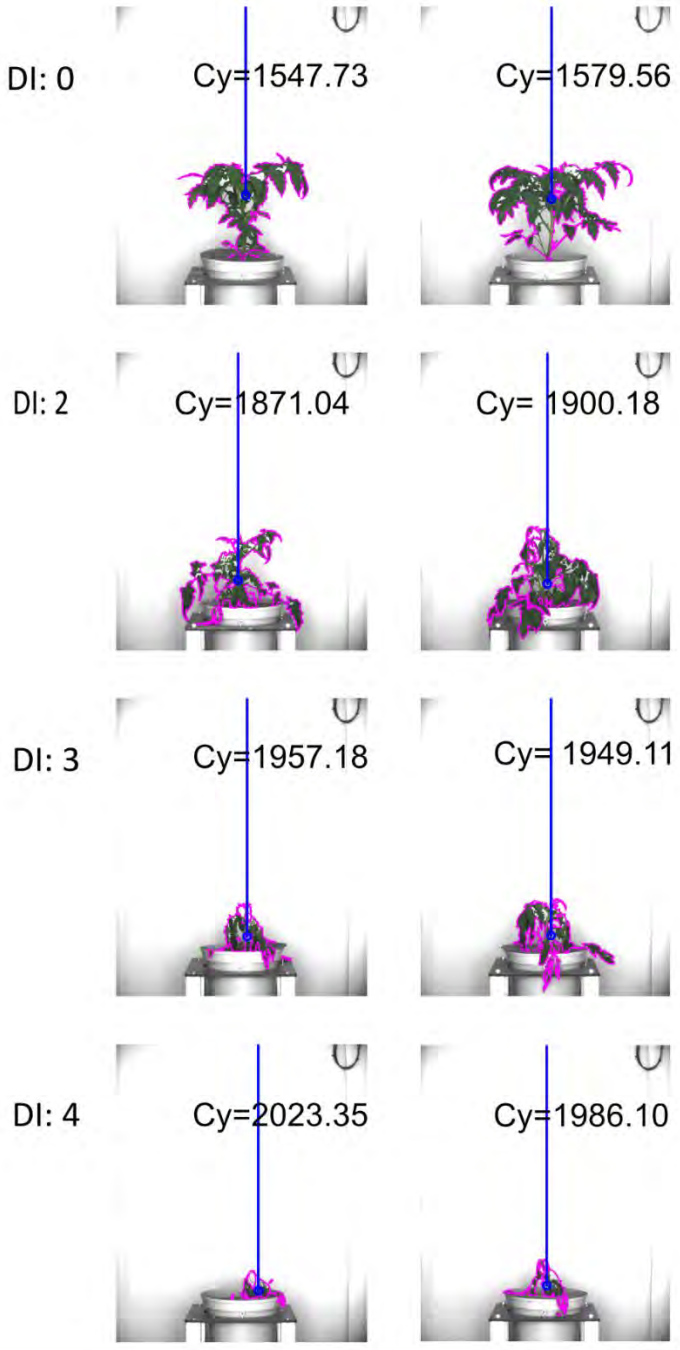



### Figure S2. The centroid-y trait (Cy).

The Centroïd-y traitor “Cy” is the y coordinate of the center of gravity of the plant area (Cy is one os the trait extracted as a final step of the image analysis process (see Supplementary figure X0). The origin of the Y axis is set at the top of the image. Plant growth, as an upwards expansion, lowers the Cy value. Bacterial wilt, with leaves dropping and later decaying, increases the Cy value.

**Figure S3. Examples of Cy values on healthy and diseased plants.**

Two distinctive plant images corresponding to disease index O, or healthy, and diseased (DI=2, 3 or 4) are showed with their respective Cy values (in pixel).



## Figure S4. Biological replicate of the physiological measurements.

Impact of *R. solanacearum* colonization on plant physiology (dry weight/transpiration) and monitoring of disease kinetics (disease index), imaging parameters (chlorophyll/center of gravity position) and water content. Chlorophyll content was estimated through RGB values (automatic phenotyping) of plant images as described by Liang et al (Liang et al., 2017). Center of gravity position (vertical) was estimated through plant imaging (automatic phenotyping) with 0 pixel representing the top of the plant. The data presented are from an experiment on 90 plants with at least 3 plants per conditions at each sampling point. Fresh weight and dry weight were used to determine the proportion of water in healthy and infected tissues. Bars indicate standard deviation. Healthy and infected plant data were compared by Wilcoxon-Mann-Whitney test (*: p-value<0.05, **: p-value<0.01, ***: p-value<0.001, ****: p-value <0.0001).

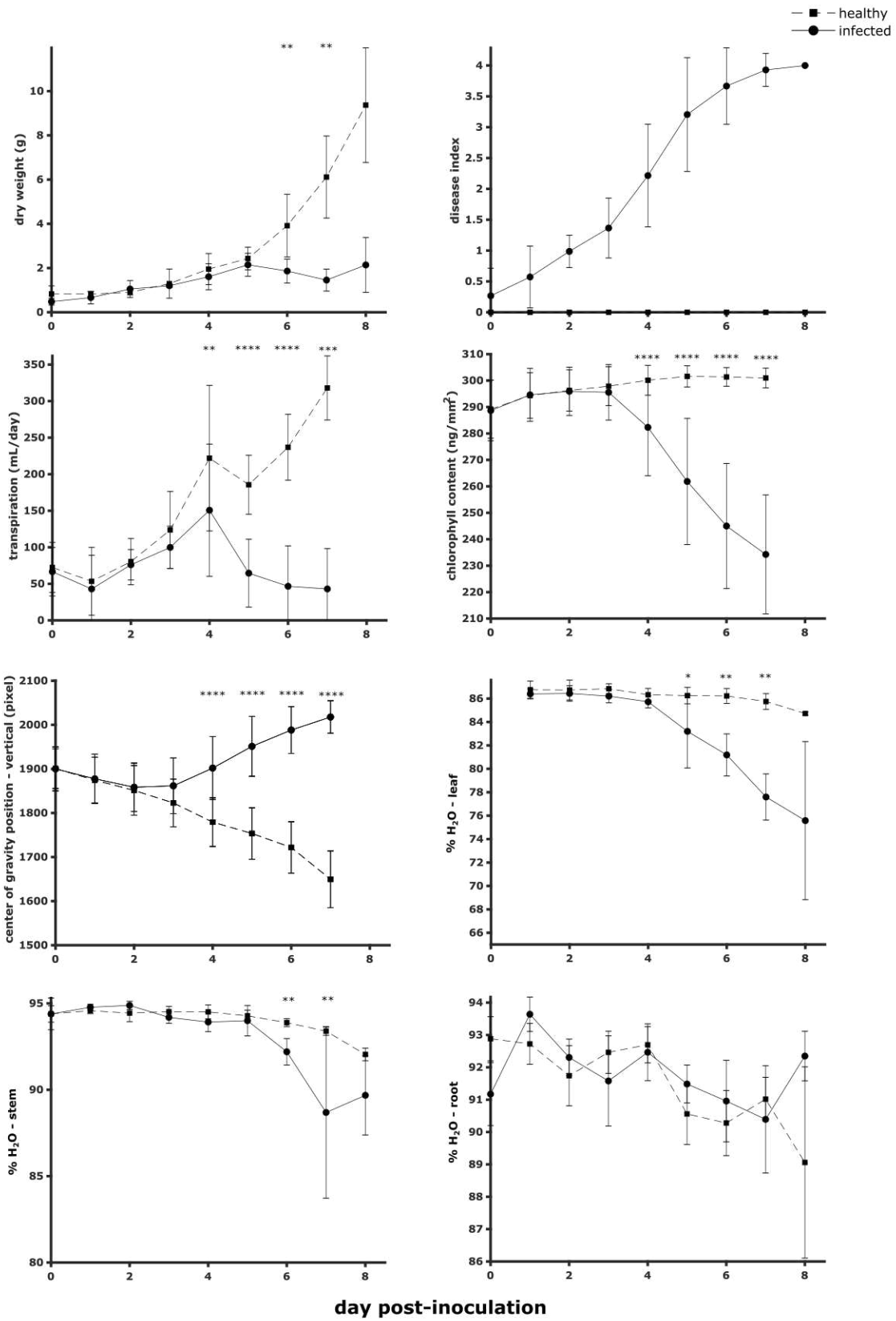


**Figure S5. Impact of *R. solanacearum* infection on plant tissues metabolic content.**

The mean and standard deviation displayed were obtained after combining the acquired data of the two biological replicates (total of 12 samples for leaf, 10 or 12 for stem, 12 for root at each time point). Healthy and infected plant data were compared by Wilcoxon-Mann-Whitney test (no significant change on the metabolites presented on this figure).

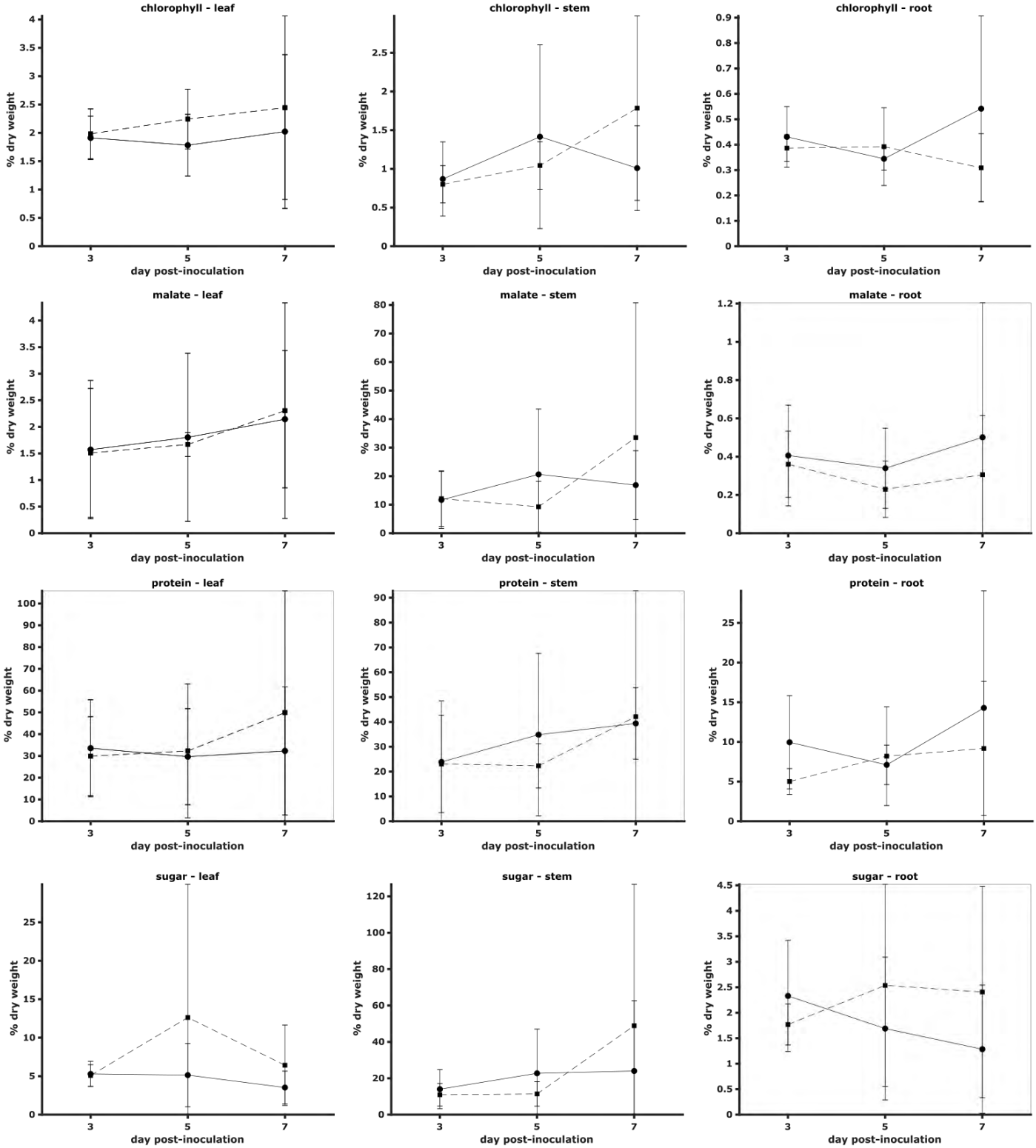
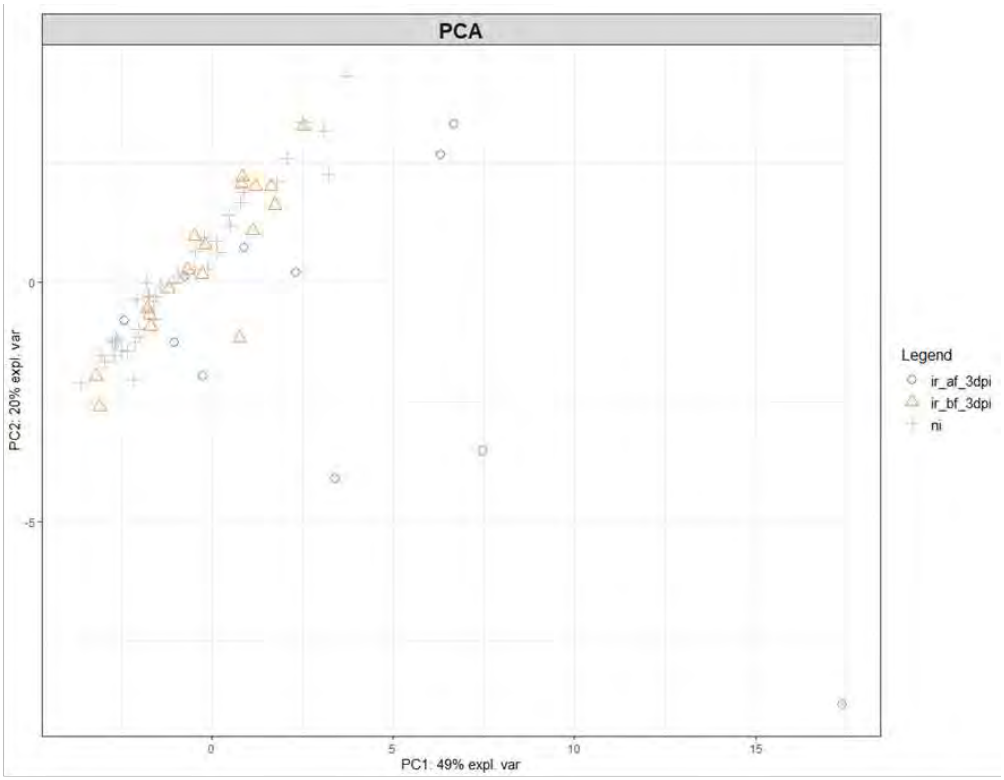


**Figure S6. PCA on the biological replicate data: experimental points (A) and variables (B).**

As CFU per gram of fresh weight data were not available, infection stage (ni: non-infected, ir_bf_3dpi: infected, sampling before or at 3dpi, ir_af_3dpi: infected, sampling after 3dpi) was mapped on the analyzed data. Similarly to the PCA presented in the Results section, glutamine are the compounds the most correlated with no infection or limited infection stage as it is oriented at the top left (B) which is associated with no or early infection. Oppositely, putrescine and acetate are at bottom right (B) which is associated with late infection. Other compounds associated with late infection include acetate, also noticed in the Results figure as associated with the infection. Other compounds appear to be associated with the infection so as fumarate but this association was not confirmed by the replicate presented in the Results figure.



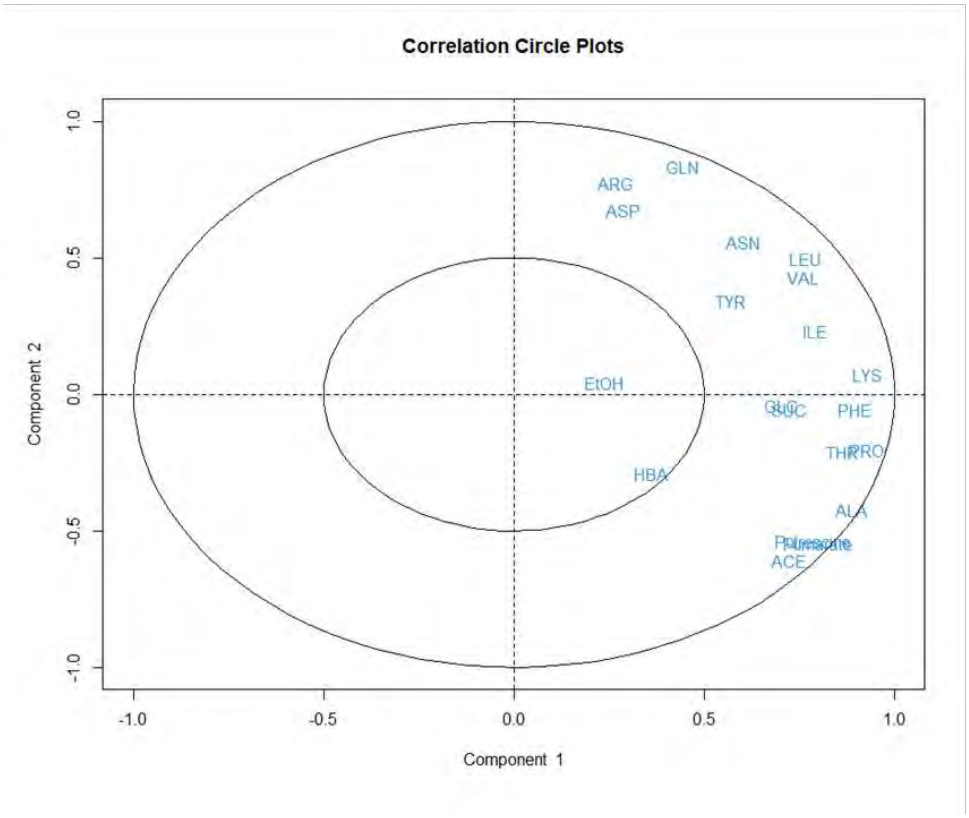



**Figure S7. Explained variance of the principal components.**

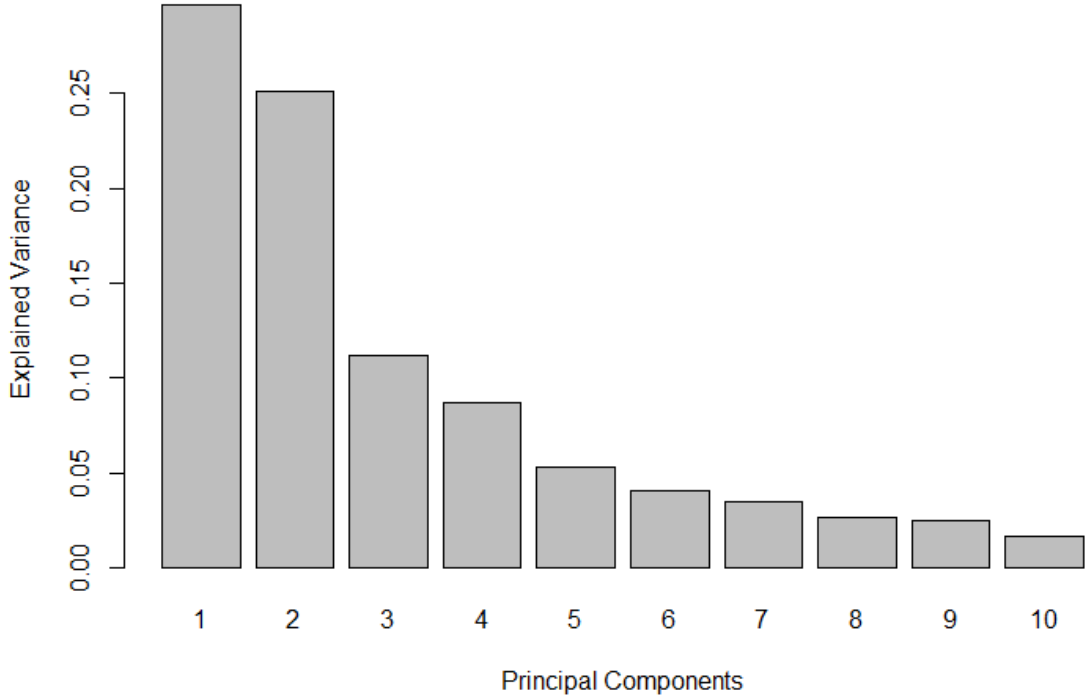


**Figure S8. Concentration profiles for alanine and other metabolites with non-statistically significant infection effect.**

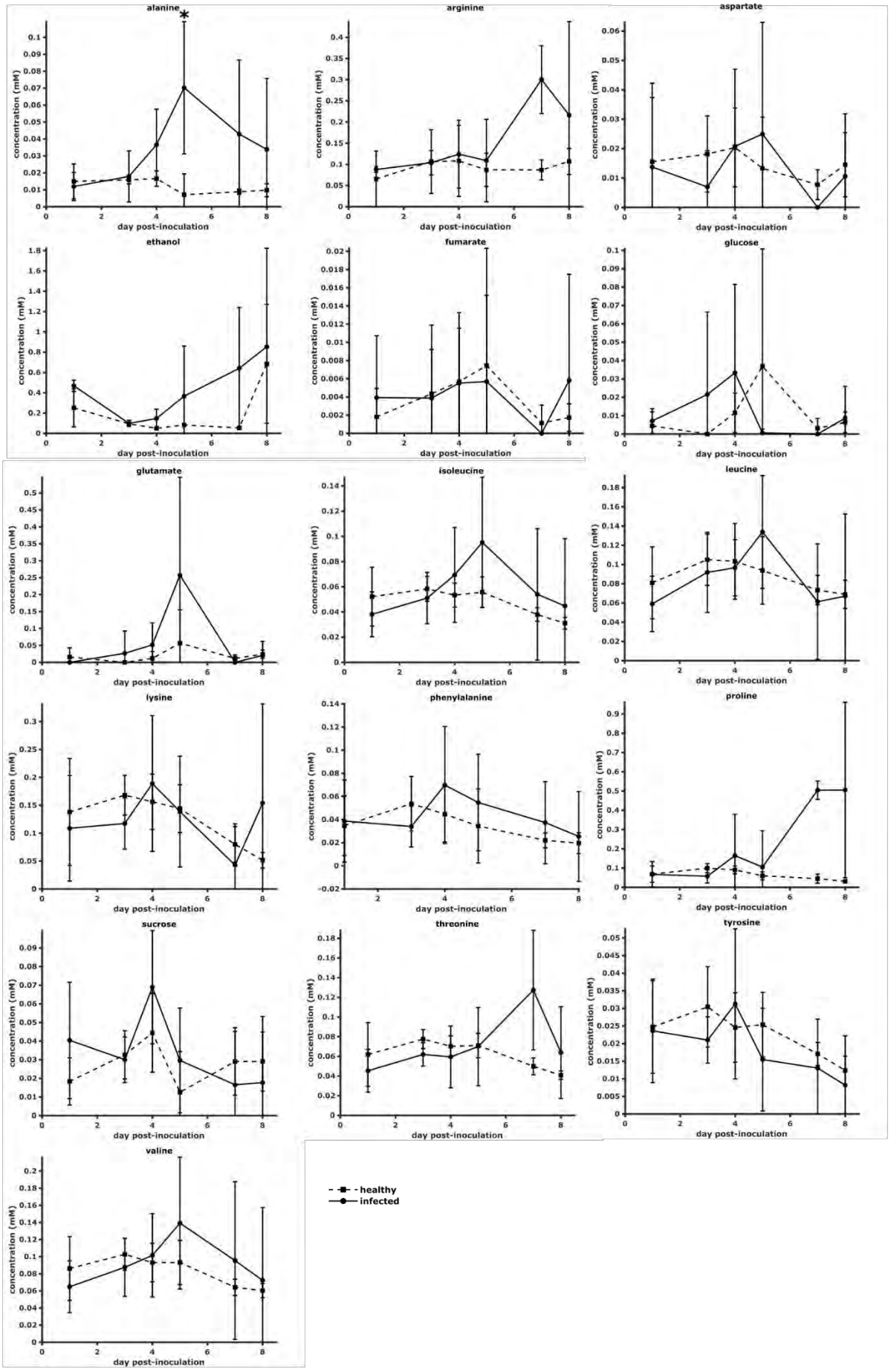


**Fig S9. In vitro and in planta phenotyping of R. solanacearum triple mutant strain defective for glucose, sucrose and fructose assimilation.**

Growth on minimal medium supplemented with glutamine (A) and sucrose (B) as sole carbon source monitored in microplates with three independent biological replicates. Means and standard deviations are represented by the lines and colored areas, respectively. Plant dry weight (C, F) and bacterial population (H) were determined on at least 9 plants per condition and per day day for infected WT/GRS941 (except 3 at 8 dpi for F) and at least 3 plants per condition and per day for healthy plants. Disease index, transpiration and chlorophyll content (D, E, G) were determined on at least 7 plants per condition and per day for infected WT/GRS941 (except 3 at 8 dpi for E, G) and at least 3 plants per day for healthy plants. E, F, G, H represent an independent experimental repeat, and C, D are complementary figures of the experiment presented in the results section of the article. Means are presented and bars indicate standard-deviation. Infection by wild-type (WT) and GRS941 strains data were compared by Wilcoxon-Mann-Whitney test (*: p-value<0.05, **: p-value<0.01).

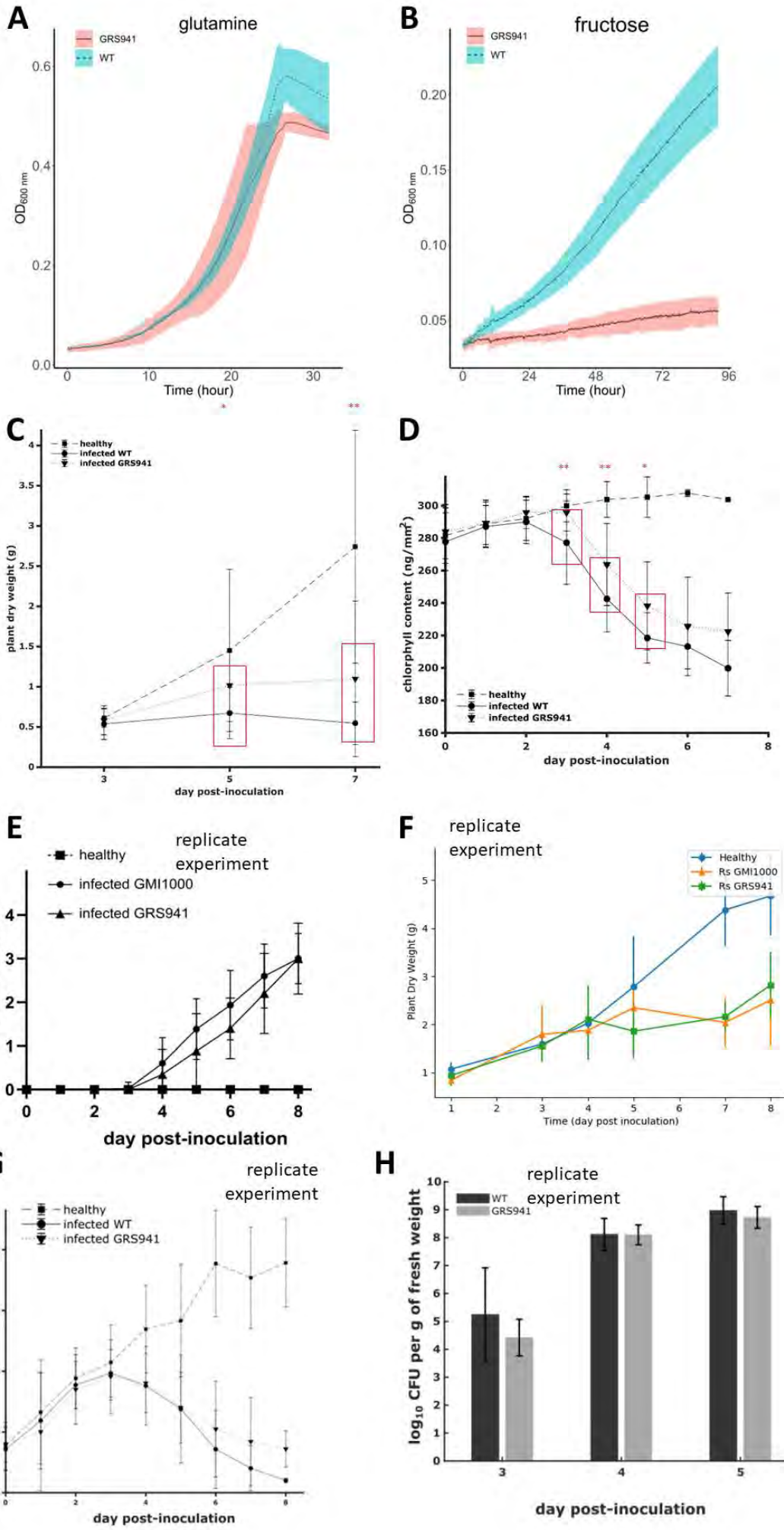

A
glutamine
GRS941
WT
OD600 nm
Time (hour)
B
fructose
GRS941
WT
OD600 nm
Time (hour)
C
healthy
infected WT
infected GRS941
plant dry weight (g)
day post-inoculation
D
healthy
infected WT
infected GRS941
chlorphyll content (ng/mm2)
day post-inoculation
E
replicate
experiment
healthy
infected GMI1000
infected GRS941
disease index
day post-inoculation
F
replicate
experiment
Healthy
Rs GMI1000
Rs GRS941
Plant Dry Weight (g)
Time (day post inoculation)
G
replicate
experiment
healthy
infected WT
infected GRS941
transpiration (mL)
day post-inoculation
H
replicate
experiment
WT
GRS941
log10 CFU per g of fresh weight
day post-inoculation

**Fig S10. NMR spectrum example.**

A. PHE: phenylalanine, TYR: tyrosine, SUC: sucrose, GLC: glucose, ASP: aspartate, ASN: asparagine, GLN: glutamine, GLU: Glutamate, PRO: proline, ARG: arginine, ACE: Acetate, Putr: Putrescine, LYS: lysine, ALA: Alanine, THR: threonine, EtOH: ethanol, 3HB: 3-hydroxybutyric acid, ILE: isoleucine, LEU: leucine, VAL: valine. B and C highlight the difference of spectrum aspects between glutamine and glutamate (or its conjugated acid glutamic acid), and a xylem sample with very few glutamate but high glutamine content: on B, the spectrum profile between 3.7 and 3.8 ppm is different (aspect and number of peaks) and on C, glutamine and glutamate profiles are very different: location and number of group of peaks.

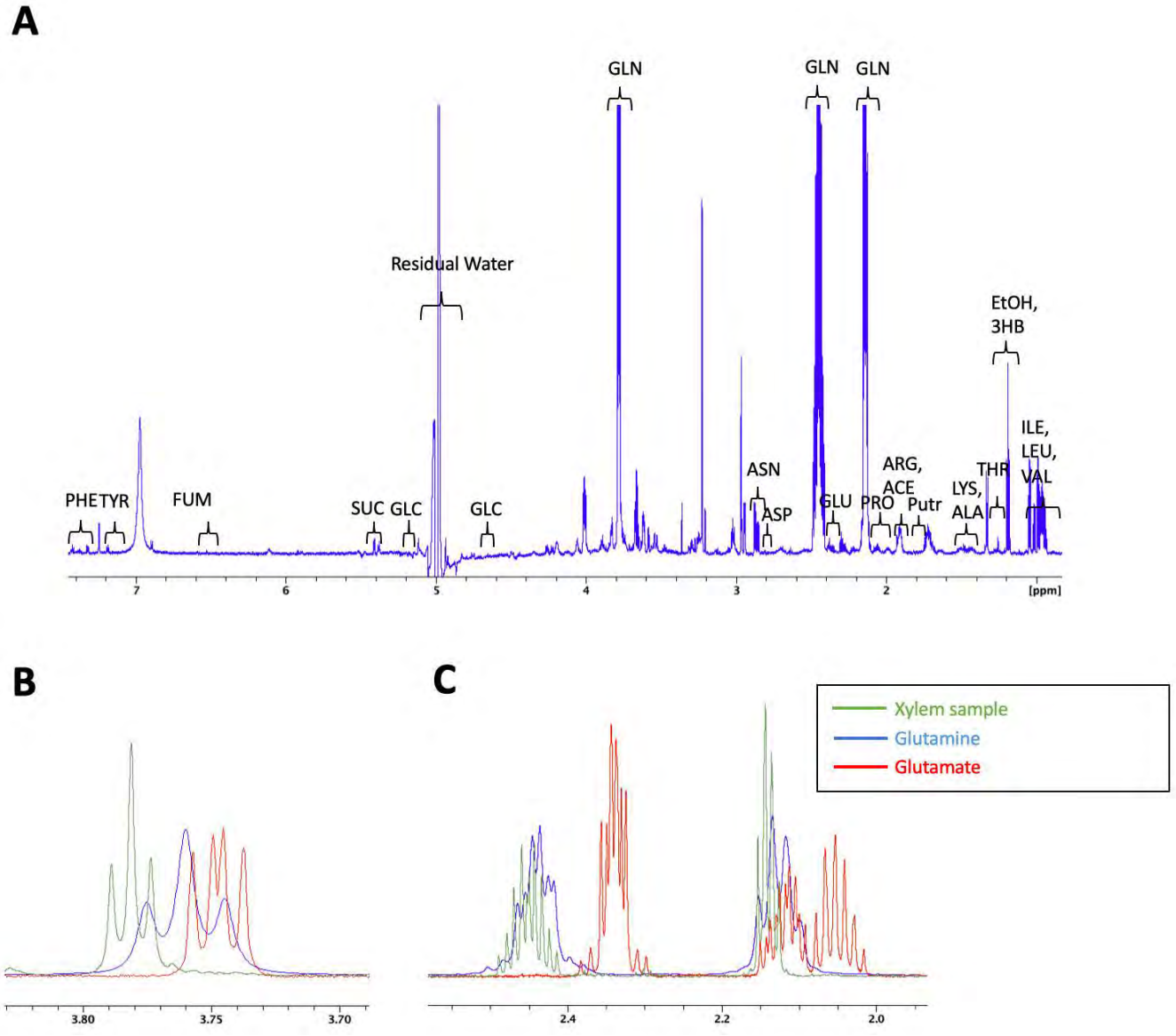


**Table S1. Biological replicate of the composition of non-infected xylem.**

| Metabolite | Concentration in healthy plant (mM) |
|---|---|
| glutamine | 3.259 ± 1.633 |
| asparagine | 0.143 ± 0.078 |
| tyrosine | 0.025 ± 0.007 |
| glucose | 0.031 ± 0.058 |
| sucrose | 0.070 ± 0.052 |
| fumarate | 0.002 ± 0.003 |
| aspartate | 0.017 ± 0.014 |
| phenylalanine | 0.029 ± 0.011 |
| leucine | 0.081 ± 0.029 |
| valine | 0.076 ± 0.026 |
| proline | 0.118 ± 0.058 |
| lysine | 0.163 ± 0.052 |
| isoleucine | 0.052 ± 0.017 |
| arginine | 0.132 ± 0.051 |
| ethanol | 0.771 ± 0.910 |
| putrescine | not detected |
| threonine | 0.058 ± 0.019 |
| glutamate | not detected |
| acetate | not detected |
| 3-hydroxybutyric acid | not detected |
| alanine | 0.021 ± 0.007 |

**Table S2. Loading weights on the second component of the PCA.**

| Metabolite | Loading weight |
|---|---|
| PUTR | 0.401 |
| GLN | -0.350 |
| ACE | 0.339 |
| ASN | -0.325 |
| PRO | 0.300 |
| ARG | 0.242 |
| ALA | 0.242 |
| HBA | 0.2309 |
| TYR | -0.217 |
| THR | 0.212 |
| SUC | -0.177 |
| GLC | -0.169 |
| ILE | 0.146 |
| ETOH | 0.116 |
| VAL | 0.116 |
| FUM | -0.113 |
| ASP | -0.091 |
| GLU | 0.086 |
| PHE | -0.075 |
| LEU | 0.020 |

# Appendix IV - Polyamines: double agents in disease and plant immunity

**Article**

Polyamines: double agents in disease and plant immunity

**Journal**

*Trends in Plant Science* (published online 11 June 2021)

**Contributions**

- Writing the first draft of the Box 1 and the sections “Polyamines: pleiotropic molecules” and “Plant polyamines as immunity players”
- Designing Figure I (in Box 1)
- Reviewing and editing the complete article with the co-authors

Trends in Plant Science

CellPress

Review

# Polyamines: double agents in disease and plant immunity

Léo Gerlin,[1,2,3] Caroline Baroukh,[1,2,3,@] and Stéphane Genin[1,2,*]

**Polyamines (PAs) are ubiquitous amine molecules found in all living organisms. In plants, beside their role in signaling and protection against abiotic stresses, there is increasing evidence that PAs have a major role in the interaction between plants and pathogens. Plant PAs are involved in immunity against pathogens, notably by amplifying pattern-triggered immunity (PTI) responses through the production of reactive oxygen species (ROS). In response, pathogens use phytotoxins and effectors to manipulate the levels of PAs in the plant, most likely to their own benefit. It also appears that pathogenic microorganisms produce PAs during infection, sometimes in large quantities. This may reflect different infectious strategies based on the selective exploitation of these molecules and the functions they perform in the cell.**

## Highlights

Polyamines (PAs) are a relatively overlooked component of the bacterial and plant metabolomes. The multiple roles of PAs have blurred our understanding of the function of these molecules in the context of plant–pathogen interactions.

Plant PAs are involved in plant immunity, because they serve as essential modulators of the redox status of the cell, contributing to potentiation of the oxidative burst as well as the synthesis of other antimicrobial compounds.

Perturbation of plant PA homeostasis appears as a common strategy of pathogens during infection. Several mechanisms involving pathogen effectors and phytotoxins lead to manipulation of plant PA biosynthetic pathways, although the outcome of these actions may have different purposes.

## Polyamines: pleiotropic molecules

PAs are organic molecules with several amine functions. Five major biogenic PAs, synthesized from amino acids, are found in the three domains of life: putrescine, spermidine, spermine, thermospermine, and cadaverine (Box 1). Putrescine and spermidine are the core PAs, present in all PA-synthesizing organisms, while some species lack spermine, thermospermine, or cadaverine synthesis. [1] These molecules are produced by most organisms except some auxotrophic protozoans [2] and rare families of bacteria, which grow without any PA [1,3].

The amine functions of PAs confer them a polycationic nature at physiological pH. Thus, they are able to stabilize or destabilize anionic macromolecules or negatively charged ions [4,5]. This chemical property makes PAs pleiotropic molecules, contributing to diverse molecular and biochemical processes. In eukaryotes and Archaea, PAs are essential for growth [1], because they (most often spermidine) are required for hypusine modification of the eukaryotic translation initiation factor 5A (eIF5A) or the archaeal homolog [6]. PAs are not strictly essential in bacteria, but numerous reports show their importance for the growth of several species, and their role in adaptive functions, such as virulence factor production or biofilm formation in human pathogens [7].

Due to their ubiquitous roles, it is complicated to dissect the multiple contributions of PAs, which are produced both by plants and pathogens. In this review, we summarize knowledge of their role during plant–pathogen interactions, because PAs increasingly appear as major players. In addition, recent evidence indicates an active role of pathogen virulence factors in modifying PA homeostasis.

## Metabolism of polyamines in plant pathogens

PAs have many roles in plant pathogens. Yet, since these molecules can be essential for growth, it is difficult to assess their precise role in pathogenesis. In particular, all pathogenic fungi, as eukaryotes, need spermidine for growth. As a result, organisms with mutations affecting the synthesis of putrescine and spermidine become auxotrophic for the missing PA, and this has been reported for several fungi, including *Stagonospora nodorum* or *Ustilago maydis* (Figure 1).

[1]LIPME, Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France
[2]Laboratory website: www.lipme.fr
[3]These authors contributed equally to this work

*Correspondence: stephane.genin@inrae.fr (S. Genin).
@Twitter: @LIPME_Toulouse (C. Baroukh).


© 2021 Elsevier Ltd. All rights reserved.

Several fungal species impaired in PA synthesis are also reduced in terms of their pathogenicity or are even avirulent on plants [8–10]. Since plants have an alternative pathway to synthesize spermidine, fungal ornithine decarboxylase (ODC) appears a target of choice for developing control strategies, without any secondary effects on the PA metabolism of the host [11]. However, this strategy only works if the auxotrophic mutant is avirulent, which is not always the case [12].

The differential incidence of the loss of the ability to synthesize PAs among pathogenic fungi might be due to the varying ability of the corresponding mutants to obtain PAs from the plant. For example, *Tapesia yallundae*, even if requiring PAs to differentiate into infective structures *in vitro*, is still infectious [12]. The authors of this study hypothesized that *T. yallundea* obtains sufficient PAs from the host plant for normal growth and differentiation. Infection of maize kernels by an *Aspergillus flavus* spermidine synthase (SPS) mutant resulted in a significant reduction in fungal growth, sporulation, and aflatoxin production [13]. Yet, the pathogen appeared to be able to import PAs from the host, since expression analyses showed significant increase in the

## Box 1. Chemical structures and metabolic pathways of polyamines

### Putrescine

Putrescine is a four-carbon diamine, at the crossroads of the amino acid and triamine/tetraamine biosynthetic pathways (Figure I). A first pathway produces putrescine through direct conversion of ornithine by ornithine decarboxylase (ODC). This pathway is identified in eukaryotes and bacteria, although it does not occur in some plants, such as arabidopsis, and some bacteria [72]. Alternatively, plants and bacteria share a putrescine biosynthetic pathway using arginine as a substrate. Arginine is converted into agmatine by arginine decarboxylase (ADC). In some organisms, agmatine is directly converted into putrescine by agmatinase, while, in others, agmatine is converted into *N*-carbamoylputrescine by agmatine deaminase and afterwards into putrescine by *N*-carbamoyl putrescine amidohydrolase [1].

### Spermidine and higher polyamines

Spermidine is a seven-carbon triamine, whereas spermine and thermospermine are ten-carbon tetraamine isomers. Their common precursor is putrescine. Spermidine is generated from putrescine and decarboxylated-*S*-adenosylmethionine (dcSAM) using SPS, and similarly, spermine or thermospermine synthase generate spermine or thermospermine from spermidine and dcSAM [1]. Some bacteria have an alternative pathway to generate spermidine from aspartate semialdehyde (AspSA) and putrescine [1]. This alternative pathway relies on two steps: first, AspSA and putrescine are converted into carboxyspermidine through carboxyspermidine dehydrogenase and then, carboxyspermidine is converted into spermidine through carboxyspermidine decarboxylase.

### Cadaverine

Cadaverine is a five-carbon diamine generated from lysine through lysine decarboxylase [73]. This pathway and the resulting PA are not branched to the four other main PAs.

### PA catabolism

In plants, spermidine, spermine, thermospermine, cadaverine, and putrescine can be catabolized by amine oxidases (AOs). All AO reactions consume $H_2O$ and $O_2$, and produce $H_2O_2$ and additional substrates. Diamine oxidase (DAO) oxidizes putrescine and cadaverine, and converts them into ammonia and 4-aminobutanal (4AB) and 5-aminopentanal (5AP), respectively [73,74]. Polyamine oxidase (PAO) oxidizes spermidine, spermine, and thermospermine. PAO back-converts spermine or thermospermine into spermidine, and spermidine into putrescine. Additionally, these reactions release 3-aminopropanal (3AP) [75–77]. Alternatively, in some plant species, PAO does not catalyze back-conversion reactions: instead, it converts spermine (spermidine) into diaminopropane and (3-aminopropyl)-4-aminobutanal (4-aminobutanal) (not shown) [74]. In fungi and animals, spermidine, spermine, and putrescine are acetylated before being oxidized into spermine, putrescine, and *N*-acetyl-4-aminobutanal, respectively [11,74,78]. These catabolic pathways can lead to the production of GABA [74,78]. Amine oxidase reactions are poorly documented in bacteria. *N*-acetylated forms of spermidine, spermine, and putrescine are observed in plants [79–81] and bacteria [82–84], but the conversion mechanisms are not entirely characterized.

### Other biogenic PAs

Several other PAs have been identified in plants or bacteria, such as 1,3-diaminopropane, norspermidine, homospermidine, aminopropylcadaverine, aminobutylcadaverine, norspermine, and aminopropylhomospermidine. However, their occurence and biological role are generally less ubiquitous than the five PAs presented in Figure I [1].

## Glossary

**Pattern-triggered immunity (PTI) and effector-triggered immunity (ETI):** two intertwined layers of the plant innate immune system. PTI is activated by microbial patterns via cell surface-localized pattern recognition receptors, whereas ETI is triggered by pathogen effectors mainly via intracellularly localized receptors.

**Phytotoxins:** small molecules produced by plant pathogens and acting as virulence factors, often inducing chlorosis or necrosis of plant tissues; their synthesis is generally of nonribosomal origin.

**Reactive oxygen species (ROS) burst:** rapid, transient production of huge amounts of ROS, and one of the earliest observable aspects of a plant immune response.

**Type 3 secretion system (T3SS):** a complex bacterial structure enabling many Gram-negative pathogens to inject virulence proteins (called effector proteins) directly into the host cell cytoplasm.

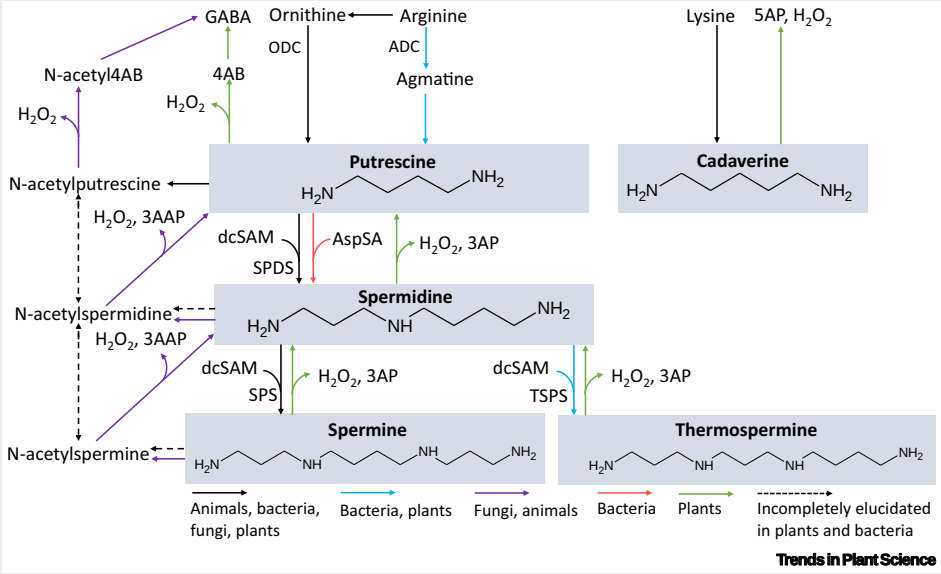



Figure I. Chemical structures and metabolic pathways of five main polyamines. A single arrow represents a single reaction or a sequence of reactions. For clarity, not all co-substrates and co-products are depicted.

expression of arginine decarboxylase (ADC) genes in maize and PA uptake transporters in the fungus [13]. In *Magnaporthe oryzae*, PAs are critical for the cell-wall integrity of the appressorium infection structure and the turgor build-up needed for rice cuticle penetration [14].

In bacteria, the essentiality of PAs for growth is not observed in all species. For instance, PAs are not essential for *Dickeya zeae* viability [15], whereas *Agrobacterium tumefaciens* requires spermidine or its precursor, 1,3-diaminopropane, for growth [16], and a *Ralstonia solanacearum* ODC mutant is auxotroph for putrescine [17]. However, for *D. zeae*, which does not require PA for growth, a mutant deficient in putrescine synthesis showed moderately decreased virulence in rice seeds, and a triple mutant deficient in both putrescine synthesis and transport was avirulent [15]. Exogenous addition of putrescine was shown to induce bacterial swimming motility and biofilm production, suggesting that this PA molecule acts as an inducing signal to promote bacterial motility and systemic infection [15]. In several bacteria, PAs are involved into biofilm formation and mobility [18], but there is no common behavior in this regard in plant pathogenic bacteria. Contrary to what is observed in *D. zeae*, no link between PA metabolism and biofilm formation or motility was observed in *R. solanacearum* or *Pseudomonas syringae* [17,19], whereas an *A. tumefaciens* ODC mutant produced more biofilm compared with the wild-type [20].

Other indirect evidence supports a specific role of PAs produced by pathogens during the parasitic phase. *R. solanaceraum* and *P. syringae* excrete abundant amounts of putrescine both *in vitro* and *in planta* [17,19,21,22]. For *R. solanacearum*, putrescine is considered a virulence factor since putrescine-pretreated tomato plants show accelerated disease symptoms and systemic spread of the pathogen [17]. However, in *P. syringae*, even when putrescine excretion was stimulated in virulence-inducing conditions *in vitro*, there was no obvious incidence on virulence, because no enhancement of disease symptoms and bacterial colonization or any difference in the expression of plant hormone signaling pathways were observed [19]. The reasons for the high production of putrescine by these bacteria remain unclear, and it is difficult

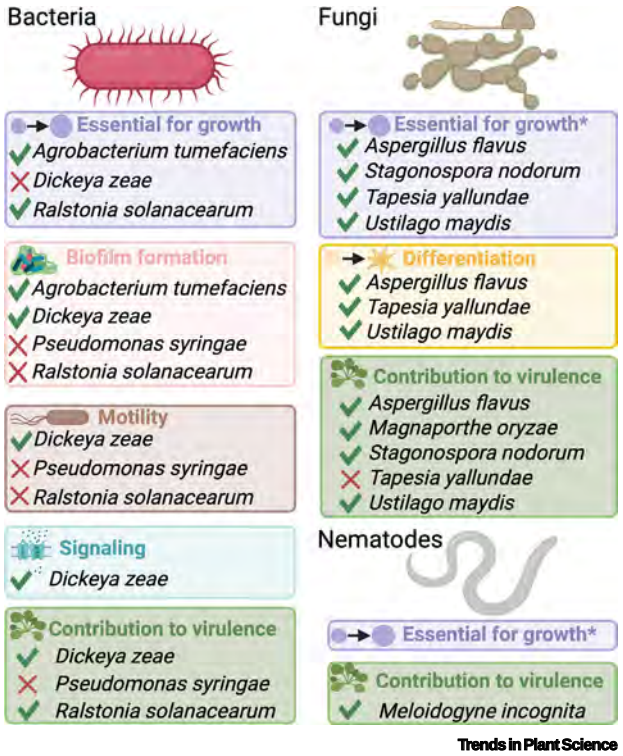



Figure 1. Documented roles of polyamines (PAs) in plant pathogens. Green tick: reported role of PAs in phenotype for the corresponding organism; Red cross: no role reported for PAs in the phenotype. *Assumed to be true for all species in this category. References: Bacteria: *Agrobacterium tumefaciens* [16,20], *Dickeya zeae* [15], *Pseudomonas syringae* [19], and *Ralstonia solanacearum* [17]; fungi: *Aspergillus flavus* [13], *Magnaporthe oryzae* [14], *Stagonospora nodorum* [70], *Tapesia yallundae* [12], and *Ustilago maydis* [8–10,68,71]; and Nematodes: *Meloidogyne incognita* [23]. Figure created using BioRender (https://biorender.com/).

to attribute a specific function to these excreted PAs, unless they are simply the result of metabolic activity associated with bacterial growth.

Finally, plant pathogens can also sense and respond to plant PAs. This is the case with some root-knot nematodes, where specific PAs (1,3-diaminopropane, putrescine, and cadaverine) are detected to locate the appropriate host plants [23]. Putrescine was also reported as an inducer of mycotoxin production in *Fusarium graminearum* [24], indicating that host PAs can act as chemical attractants to trigger expression of virulence factors. It is also hypothesized that putrescine in rice seeds triggers motility and biofilm production in *D. zeae*, both traits required for entrance into the host [15].

In conclusion, PAs have multiple roles in microorganisms, with significant variations between organisms, in particular with regard to their requirement for growth; therefore, it is difficult to establish general rules. PAs have been reported to have either little or major contributions to virulence, depending on the pathogen, and probably the host. It remains to be investigated whether the contribution of PAs to virulence implies the involvement of other mechanisms, such as free radical ion scavenging or participating in the synthesis of complex molecules (e.g., siderophores), as has been described for other pathogenic organisms [18].

## Plant polyamines as immunity players

In addition to their role in cell proliferation described for all eukaryotes, PAs contribute to diverse pathways involved in plant development and abiotic stress responses [25]. In arabidopsis (*Arabidopsis thaliana)*, cold, drought, and salt induce the expression of genes required for the PA biosynthesis pathway, and transgenic plants overproducing PAs have increased abiotic stress tolerance [25]. Many studies have also pointed out a role of PAs in plant immune responses

to pathogens [26,27]. However, due to their pleiotropic nature, it is difficult to disentangle the general functions of PAs from those specifically dedicated to immunity. In addition, homeostasis of the PA pool is subject to complex crosstalk with hormonal and signaling pathways, including those of abscisic acid (ABA), salicylic acid (SA), nitric oxide (NO), and calcium [25].

Modulations of PA levels and PA biosynthetic/catabolic enzyme activities in pathogenic fungi interactions have been described since the 1980s. Variations in PA levels in response to pathogens were observed with different pathosystems and different types of interaction, either compatible or incompatible [28]. The increase/decrease pattern of PAs or associated enzyme activities may differ depending on the PA being studied and the type of interaction, although PA increases are more frequently observed than decreases.

Recent transcriptomics and metabolomics analyses of the interaction between *A. flavus* and both resistant/susceptible maize lines revealed an increase in PA levels in both lines, but with different profiles between resistant and susceptible plants, as well as the upregulation of PA biosynthesis and catabolism genes [13]. In bacterial pathogens, studies have all highlighted a pattern of putrescine accumulation, which may be a conserved trait, because it is found in different plants (*A. thaliana*, *Solanum lycopersicum*) interacting with different pathogens (*P. syringae*, *Pseudomonas viridiflava*, and *R. solanacearum*) [19,29–32]. Putrescine accumulation was triggered in the *A. thaliana–P. syringae* pathosystem by both **PTI and effector-triggered immunity (**ETI; see Glossary) [30,31] and probably relies on induced expression of one or both of the ADC genes present in arabidopsis [29].

Different mechanisms are proposed to explain the recurrent variations in PA levels in plant cells during biotic interactions. First, PAs have a major role in the control of hydrogen peroxide ($H_2O_2$) production, $H_2O_2$ being the prevalent ROS in plants [30,31]. PAs participate in enhancing $H_2O_2$ levels through their catabolism by amine oxidases, which generates $H_2O_2$ (Box 1). By contrast, higher PAs, such as spermine, can also act as hydroxyl radical scavengers [33–35]. Therefore, through the control of ROS levels, PAs have a key role in redox homeostasis. In addition, $H_2O_2$ acts directly in the signaling cascade inducing responses to stress, promoting phytoalexin synthesis [36–38] or cell wall rigidification [39]. Hydrogen peroxide radicals are also tightly linked with other major signaling molecules in plants, such as NO, ABA, SA, and calcium [25,40–42]. Finally, $H_2O_2$ has direct roles in responses to pathogens, being a precursor for radicals contributing to PTI and ETI responses [33]. Thus, PAs actively contribute to potentiation of the **ROS burst** induced in response to the bacterial flagellar epitope flg22 [43]. Loss-of-function mutations in arabidopsis ADC2, which is required for putrescine synthesis, compromise basal defenses mediated by ROS-dependent SA pathways [30,31].

A second way in which plant PAs contribute to defensive responses to pathogens is as precursors of hydroxycinnamic acid amides (HCAAs). HCAAs result from the association of a PA and a phenolic molecule, and are potent antimicrobial molecules [44]. HCAAs have been implicated in incompatible interactions between plant and fungi, as well as for tomato infected by *R. solanacearum* [32,45]. Here again, there is possible crosstalk between the regulation of PAs and hormonal pathways, since it has been reported that ethylene synthesis is attenuated by an increase in spermidine in the plant, which then promotes greater susceptibility of the tomato to *Botrytis cinerea* [46]. Of note, *N*-acylated polyamines (phenolamides), which are present ubiquitously in plants, were shown to be essential for local and systemic defenses against insect herbivores [47,48]. Putrescine is also one of the precursors of nicotine, a defensive neurotoxin of *Nicotiana* spp. against herbivores [49].

Overall, the modulation of PA levels appears to be a conserved trait of the plant immune response to eukaryotic and prokaryotic pathogens, but it is difficult to interpret the purposes of these variations, given that PAs are also involved in plant development and responses to abiotic stresses. Increasing evidence connects PAs to ROS, which are proven key contributors to the plant immune response [30,50]. However, PAs are involved in other defense mechanisms, such as the production of HCAA, or through crosstalk with hormonal pathways, such as those of ethylene or SA [30,31,44,46].

## Toxins inhibiting plant polyamine pathways

It is well known that multiple bacterial **phytotoxins** act on many plant metabolic pathways [51,52]. In recent years, the view has emerged that several of these phytotoxins, especially produced by various *P. syringae* strains, target more specifically key steps in PA biosynthesis or their ability to mount the immune oxidative burst (Figure 2). Phaseolotoxin inhibits ornithine transcarbamylase (OTC), a critical enzyme in the urea cycle, converting ornithine and carbamoyl phosphate to citrulline [53,54]. Hence, phaseolotoxin promotes virulence by antagonizing the synthesis of arginine and PAs. Another molecule produced by *P. syringae* pathovars, mangotoxin, was also shown to inhibit the enzyme ornithine *N*-acetyl transferase and, thus, blocks ornithine synthesis [55]. More recently, phevamine A, another small molecule produced by some *P. syringae* and *Erwinia amylovora* strains, was also shown to target the arginine/PA pathway [43]. Phevamine A is a virulence factor the synthesis of which is co-regulated with that of the **type 3 secretion system (T3SS)** [43,56]. This molecule suppresses the potentiation effect of PAs and arginine on the ROS burst generated upon recognition of bacterial flagellin [43]. Therefore, the emerging picture is that *P. syringae* uses either phaseolotoxin or phevamine A to attenuate the host immune response by altering PA biosynthesis or the signaling capacity of PAs, respectively (Figure 2).

Interestingly, there is evidence that another *P. syringae* phytotoxin, coronatine, also interferes with PA biosynthesis. Coronatine acts as a structural mimic of the plant hormone jasmonic acid-isoleucine and promotes entry of bacteria into the plant by stimulating the opening of stomata and suppression of SA-dependent host defenses [57,58]. Another action of coronatine in dampening plant immunity is to induce acetylation of putrescine by *A. thaliana N*-ACETYL TRANSFERASE ACTIVITY 1 (NATA1) [59]. NATA1 catalyzes the acetylation of putrescine to *N*-acetylputrescine [60] and thereby competes with SPS for a common substrate, reducing the accumulation and role of spermidine in the potentiation of microbe-associated molecular pattern (MAMP)-induced ROS bursts.

## Effector proteins also manipulate plant polyamine pools

Besides phytotoxins, pathogens also produce effector proteins the functions of which are to subvert several cellular functions of the host, in particular the immune response [61]. However, a recent example illustrates that the PA biosynthetic pathway in plants is also an operational area for effectors, primarily not with the aim of bypassing the mechanisms of immunity. Brg11, a *R. solanacearum* transcription activation-like (TAL) effector specifically targets ADC promoters to increase host ADC transcript levels [62] (Figure 2). Brg11 also induces distinct ADC transcript variants that mediate high levels of translation; metabolic analyses showed that Brg11-mediated activation leads to significant increases in putrescine levels. Wu *et al.* [62] proposed that Brg11 modifies the host environment specifically to inhibit competing pathogens while having little effect on the population of *R. solanacearum* that produces the effector. Brg11-mediated accumulation of putrescine attenuates the *in planta* growth of *P. syringae*, but not that of *R. solanacearum*, supporting the view that the action of Brg11 benefits the latter via suppression of niche competitors [62]. This model implies a differential susceptibility of resident bacterial populations to host

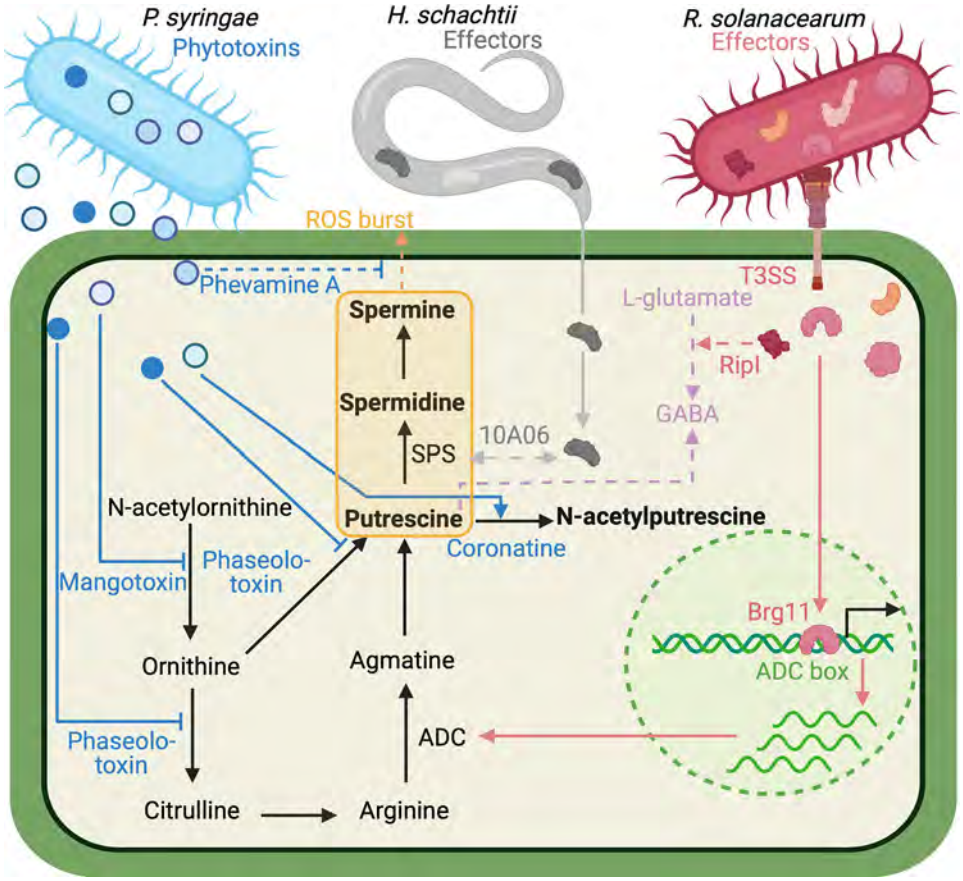



Figure 2. Action of pathogenic effector molecules on plant polyamine (PA) biosynthetic pathways. Examples include phytotoxins (in blue) from *Pseudomonas syringae* and effector proteins from *Heterodera schachtii* (gray) and *Ralstonia solanacearum* (red). The PA pool is symbolized by the orange box; PA precursor reaction pathways are shown in black. Phytotoxins either inhibit (flat-headed arrows) or promote (pointed arrow) specific reactions, while phevamine A attenuates (broken arrow) the PA-dependent reactive oxygen species (ROS) burst [43]. *H. schachtii* effector 10A06 interacts with the spermidine synthase (SPS) enzyme [60]. The transcription activation-like (TAL) effector Brg11 translocated inside the plant cell by the type 3 secretion system (T3SS) migrates to the plant nucleus (green circle) to elevate transcription of the arginine decarboxylase (ADC) gene after binding to the upstream DNA control sequence, the ADC box. Brg11-dependent transcription activation leads to enhanced synthesis of ADC and, in turn, to augmented putrescine levels. The resulting PA boost elicits biotic stress defense responses to ward off *R. solanacearum*-rivaling bacteria [62]. The figure also shows the link between polyamine catabolism and gamma-aminobutyric acid (GABA) synthesis, as well as the role of the effector RipI, which promotes the synthesis of GABA by *R. solanacearum* as a preferred metabolic source [66]. Interaction of the *Xanthomonas* effector AvrBsT with ADC [64] is not shown. Figure created using BioRender (https://biorender.com/).

putrescine-dependent immunity; however, the mechanism by which *R. solanacearum* specifically escapes the effects of this response remains unknown. Interestingly, a TAL effector, TAL17d, with the same ability to bind to the ADC promoters as Brg11, was also identified in the pathogen *Xanthomonas campestris* [63], suggesting a similar way to subvert expression of the host putrescine biosynthetic pathway. Another effector, AvrBsT, also translocated by the *Xanthomonas* T3SS, has been shown to physically interact with the ADC enzyme [64], but the mechanistic consequences of this interaction have not yet been elucidated.

The catabolism of PAs is also linked to γ-aminobutyric acid (GABA) metabolism (Figure 2). GABA is a versatile signaling molecule, which modulates the immune response in plants and accumulates in the apoplast during pathogen infection [65]. The recent finding that another *R. solanacearum* effector, RipI, induces GABA synthesis to support its nutrition *in planta* [66] further illustrates the complexity of the picture and shows how the effects of bacterial effectors simultaneously impact interconnected metabolic pathways for radically different physiological outcomes.

Finally, a translocated effector of the cyst nematode *Heterodera schachtii* was shown to interact with an arabidopsis SPS protein (SPDS2), thereby increasing spermidine content and, subsequently, PA oxidase activity [67]. The action of this effector, named 10A06, was presumed to stimulate the ROS-scavenging properties of plant spermidine as a safeguard for the developing syncytium, required for successful cyst nematode parasitism [67]. Beside this example, there are still few indications that nonbacterial pathogens directly manipulate PA levels in plants, although the fungus *U. maydis* has been reported to induce increased putrescine levels in maize tumors [68]. However, it is telling that a mycorrhizal fungus effector was recently shown to increase PA biosynthesis to promote successful symbiont colonization [69], thus suggesting that fine tuning of host PA levels controls interactions with a broad spectrum of microorganisms, not only pathogens.

## Concluding remarks and future perspectives

The metabolism of PAs is essential in both hosts and pathogens, with a wide array of described functions, which makes it difficult to determine their real causal effects in observed physiological and phenotypical responses (Figure 3, key figure). Pathogens induce perturbations in the homeostasis of PAs in plants by acting positively or negatively through various sophisticated mechanisms on plant PA biosynthetic and regulatory pathways (see Outstanding questions). Distinct pathogens appear to have opposite strategies in this regard, illustrated by *P. syringae* and *R. solanacearum*. *P. syringae* represses the PA biosynthetic pathways, while *R. solanacearum* induces the increased synthesis of putrescine. However, the recurrence of these actions in terms of altering the metabolism of PAs (through various phytotoxins or effectors) highlights the importance of the host PA pool in the immunity process.

PAs appear to have an essential role in both ROS production and detoxification in plants, which links them directly to the immune response. Their role is dual, given that they are involved in both resistance and sensitivity depending on the pathogen, although this latter feature has not yet been explained fully. On the pathogen side, there appears to be a differential vulnerability to PA-dependent immunity between infecting microorganisms, the mechanisms of which also remain unknown.

The dual role of PAs can also be emphasized in pathogens, some of which subtly manipulate host PA metabolism and simultaneously secrete abundant quantities of PAs for purposes that remain elusive. Several hypotheses, probably not exclusive, can be formulated for the role of PAs produced by pathogens, such as a basic cellular function during infection, a metabolic by-product, or a bona fide virulence factor (see also Outstanding questions).

### Outstanding questions

Why do plant pathogens excrete PAs? Do these PAs perform different functions depending on the pathogen? Does this excretion of PAs contribute to pathogenicity? If so, by what mechanisms?

What are the mechanisms underlying the control of PA pools in plant cells? Do putrescine and higher PAs (spermine and spermidine) perform distinct functions in plant immunity? Is the importance of PA pools in plants directly related to induction of the oxidative burst or other defense or signaling pathways?

Is perturbation of plant PA homeostasis a common strategy of pathogens during infection or does this perturbation have several purposes and, therefore, is specific to each pathogen? Is there a larger set of pathogenic effectors and toxins influencing PA levels in plants?

To what extent is modulation of the levels of plant PAs key to controlling a broader set of interactions, such as mutualist organisms or the microbiota?

**Acknowledgments**

Work in our laboratory is supported by the French Laboratory of Excellence project 'TULIP' (grant number ANR-10-LABX-41; ANR-11-IDEX-0002-02). L.G. was funded by a grant from the French Ministry of National Education and Research.

## Key figure

### Multiple roles of polyamines (PAs) during plant–pathogen interactions

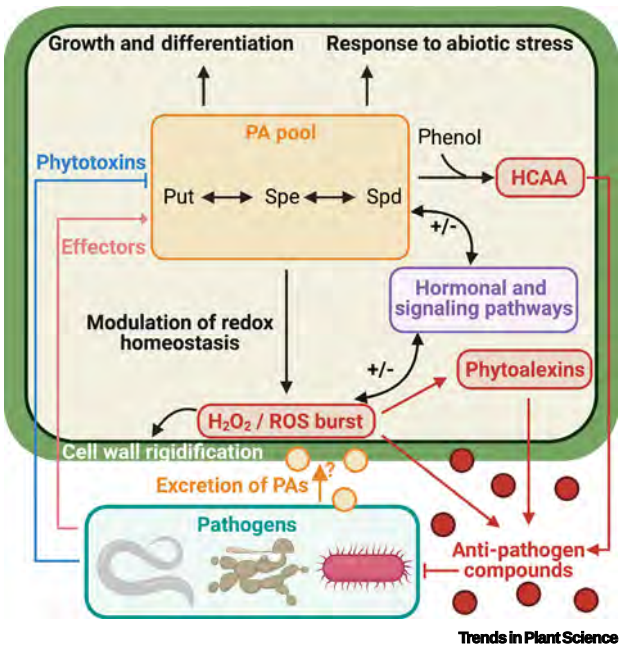



Figure 3. Beyond its major roles in cell growth, differentiation, and resistance to abiotic stress, the plant PA pool (Put, putrescine; Spe, spermine; Spd, spermidine) is also mobilized during interactions with pathogens. PAs are used to modulate the balance of reactive oxygen species (ROS) generation and elimination (redox homeostasis) and, therefore, the early oxidative burst observed in response to pathogen infection. Plant PAs are also precursors of hydroxycinnamic acid amides (HCAAs) that act, together with phytoalexins and ROS, as defense molecules. ROS are also involved in strengthening the plant wall to counteract infection. By contrast, pathogens have the capacity to manipulate plant PA homeostasis through the action of effector proteins or phytotoxins. For example, phytotoxins inhibit PA biosynthesis and potentiation of the ROS burst to reduce the plant defense response, while the Brg11 effector promotes putrescine synthesis to enhance plant immunity against competing bacteria. Pathogens also excrete PAs, but it is still unclear whether this is directly involved in the infection process (question mark). Perception of plant PAs by pathogens as signal molecules is not shown. Figure created using BioRender (https://biorender.com/).

## Declaration of interests

None declared by authors.

## References


1. Michael, A.J. (2016) Polyamines in eukaryotes, bacteria, and archaea. *J. Biol. Chem.* 291, 14896–14903
2. Li, B. *et al.* (2015) Different polyamine pathways from bacteria have replaced eukaryotic spermidine biosynthesis in ciliates *Tetrahymena thermophila* and *Paramecium tetaurelia*. *Mol. Microbiol.* 97, 791–807
3. Li, B. *et al.* (2019) Polyamine-independent growth and biofilm formation, and functional spermidine/spermine N-acetyltransferases in *Staphylococcus aureus* and *Enterococcus faecalis*. *Mol. Microbiol.* 111, 159–175
4. Dobrovinskaya, O.R. *et al.* (1999) Asymmetric block of the plant vacuolar $Ca^{2+}$-permeable channel by organic cations. *Eur. Biophys. J.* 28, 552–563
5. Dobrovinskaya, O.R. *et al.* (1999) Inhibition of vacuolar ion channels by polyamines. *J. Membr. Biol.* 167, 127–140
6. Chattopadhyay, M.K. *et al.* (2008) Hypusine modification for growth is the major function of spermidine in *Saccharomyces cerevisiae* polyamine auxotrophs grown in limiting spermidine. *Proc. Natl. Acad. Sci. U. S. A.* 105, 6554–6559
7. Di Martino, M.L. *et al.* (2013) Polyamines: emerging players in bacteria-host interactions. *Int. J. Med. Microbiol.* 303, 484–491
8. Valdés-Santiago, L. *et al.* (2010) Life without putrescine: disruption of the gene-encoding polyamine oxidase in *Ustilago maydis* odc mutants. *FEMS Yeast Res.* 10, 928–940
9. Valdés-Santiago, L. *et al.* (2009) *Ustilago maydis* spermidine synthase is encoded by a chimeric gene, required for morphogenesis, and indispensable for survival in the host. *FEMS Yeast Res.* 9, 923–935
10. Valdés-Santiago, L. *et al.* (2012) Phenotypic comparison of samdc and spe mutants reveals complex relationships of

polyamine metabolism in *Ustilago maydis*. *Microbiology* 158, 674–684
11. Valdés-Santiago, L. *et al.* (2012) Polyamine metabolism in fungi with emphasis on phytopathogenic species. *J. Amino Acids* 2012, 1–13
12. Mueller, E. *et al.* (2001) Ornithine decarboxylase knockout in *Tapesia yallundae* abolishes infection plaque formation in vitro but does not reduce virulence toward wheat. *Mol. Plant-Microbe Interact.* 14, 1303–1311
13. Majumdar, R. *et al.* (2019) Contribution of maize polyamine and amino acid metabolism toward resistance against *Aspergillus flavus* infection and aflatoxin production. *Front. Plant Sci.* 10, 692
14. Rocha, R.O. *et al.* (2020) Spermine-mediated tight sealing of the *Magnaporthe oryzae* appressorial pore–rice leaf surface interface. *Nat. Microbiol.* 5, 1472–1480
15. Shi, Z. *et al.* (2019) Putrescine is an intraspecies and interkingdom cell-cell communication signal modulating the virulence of *Dickeya zeae*. *Front. Microbiol.* 10, 1950
16. Kim, S.H. *et al.* (2016) The essential role of spermidine in growth of *Agrobacterium tumefaciens* Is determined by the 1,3-diaminopropane moiety. *ACS Chem. Biol.* 11, 491–499
17. Lowe-Power, T.M. *et al.* (2017) Metabolomics of tomato xylem sap during bacterial wilt reveals *Ralstonia solanacearum* produces abundant putrescine, a metabolite that accelerates wilt disease. *Environ. Microbiol.* 20, 1330–1349
18. Michael, A.J. (2018) Polyamine function in archaea and bacteria. *J. Biol. Chem.* 293, 18693–18701
19. Vilas, J.M.M. *et al.* (2018) Modulation of plant and bacterial polyamine metabolism during the compatible interaction between tomato and *Pseudomonas syringae*. *J. Plant Physiol.* 231, 281–290
20. Wang, Y. *et al.* (2016) Spermidine inversely influences surface interactions and planktonic growth in *Agrobacterium tumefaciens*. *J. Bacteriol.* 198, 2682–2691
21. Peyraud, R. *et al.* (2016) A resource allocation trade-off between virulence and proliferation drives metabolic versatility in the plant pathogen *Ralstonia solanacearum*. *PLoS Pathog.* 12, e1005939
22. Gerlin, L. *et al.* (2021) Unravelling physiological signatures of tomato bacterial wilt and xylem metabolites exploited by *Ralstonia solanacearum*. *Environ. Microbiol.* Published online April 19. 2021. https://doi.org/10.1111/1462-2920.15535
23. Oota, M. *et al.* (2020) Identification of naturally occurring polyamines as root-knot nematode attractants. *Mol. Plant* 13, 658–665
24. Gardiner, D.M. *et al.* (2010) Early activation of wheat polyamine biosynthesis during *Fusarium* head blight implicates putrescine as an inducer of trichothecene mycotoxin production. *BMC Plant Biol.* 10, 1–13
25. Alcázar, R. *et al.* (2010) Polyamines: molecules with regulatory functions in plant abiotic stress tolerance. *Planta* 231, 1237–1249
26. Gonzalez, M.E. *et al.* (2021) Current status and perspectives on the role of polyamines in plant immunity. *Ann. Appl. Biol.* 178, 244–255
27. Seifi, H.S. and Shelp, B.J. (2019) Spermine differentially refines plant defense responses against biotic and abiotic stresses. *Front. Plant Sci.* 10, 117
28. Walters, D.R. (2003) Polyamines and plant disease. *Phytochemistry* 64, 97–107
29. Rossi, F.R. *et al.* (2015) Role of arginine decarboxylase (ADC) in *Arabidopsis thaliana* defence against the pathogenic bacterium *Pseudomonas viridiflava*. *Plant Biol.* 17, 831–839
30. Liu, C. *et al.* (2019) The polyamine putrescine contributes to $H_2O_2$ and RbohD/F-dependent positive feedback loop in arabidopsis PAMP-triggered immunity. *Front. Plant Sci.* 10, 894
31. Liu, C. *et al.* (2020) Putrescine elicits ROS-dependent activation of the salicylic acid pathway in *Arabidopsis thaliana*. *Plant Cell Environ.* 43, 2755–2768
32. Zeiss, D.R. *et al.* (2019) Metabolomic profiling of the host response of tomato (*Solanum lycopersicum*) following infection by *Ralstonia solanacearum*. *Int. J. Mol. Sci.* 20, 3945
33. Pottosin, I. *et al.* (2014) Cross-talk between reactive oxygen species and polyamines in regulation of ion transport across the plasma membrane: implications for plant adaptive responses. *J. Exp. Bot.* 65, 1271–1283
34. Das, K.C. and Misra, H.P. (2004) Hydroxyl radical scavenging and singlet oxygen quenching properties of polyamines. *Mol. Cell. Biochem.* 262, 127–133
35. Ha, H.C. *et al.* (1998) The natural polyamine spermine functions directly as a free radical scavenger. *Proc. Natl. Acad. Sci. U. S. A.* 95, 11140–11145
36. Thoma, I. *et al.* (2003) Cyclopentenone isoprostanes induced by reactive oxygen species trigger defense gene activation and phytoalexin accumulation in plants. *Plant J.* 34, 363–375
37. Tierens, K.F.M-J. *et al.* (2002) *Esa1*, an *Arabidopsis* mutant with enhanced susceptibility to a range of necrotrophic fungal pathogens, shows a distorted induction of defense responses by reactive oxygen generating compounds. *Plant J.* 29, 131–140
38. Jabs, T. *et al.* (1997) Elicitor-stimulated ion fluxes and $O_2$- from the oxidative burst are essential components in triggering defense gene activation and phytoalexin synthesis in parsley. *Proc. Natl. Acad. Sci. U. S. A.* 94, 4800–4805
39. Gupta, K. *et al.* (2016) Hydrogen peroxide and polyamines act as double edged swords in plant abiotic stress responses. *Front. Plant Sci.* 7, 1343
40. Mohanta, T.K. *et al.* (2018) Early events in plant abiotic stress signaling: interplay between calcium, reactive oxygen species and phytohormones. *J. Plant Growth Regul.* 37, 1033–1049
41. Ton, J. *et al.* (2009) The multifaceted role of ABA in disease resistance. *Trends Plant Sci.* 14, 310–317
42. Liu, C. and Alcázar, R. (2021) A new insight into the contribution of putrescine to defense in *Arabidopsis thaliana*. *Plant Signal. Behav.* 16, 1885187
43. O'Neill, E.M. *et al.* (2018) Phevamine A, a small molecule that suppresses plant immune responses. *Proc. Natl. Acad. Sci. U. S. A.* 115, E9514–E9522
44. Zeiss, D.R. *et al.* (2021) Hydroxycinnamate amides: intriguing conjugates of plant protective metabolites. *Trends Plant Sci.* 26, 184–195
45. Lowe, T.M. *et al.* (2015) Hydroxycinnamic acid degradation, a broadly conserved trait, protects *Ralstonia solanacearum* from chemical plant defenses and contributes to root colonization and virulence. *Mol. Plant-Microbe Interact.* 28, 286–297
46. Nambeesan, S. *et al.* (2012) Polyamines attenuate ethylene-mediated defense responses to abrogate resistance to *Botrytis cinerea* in tomato. *Plant Physiol.* 158, 1034–1045
47. Kaur, H. *et al.* (2010) R2R3-NaMYB8 regulates the accumulation of phenylpropanoid-polyamine conjugates, which are essential for local and systemic defense against insect herbivores in *Nicotiana attenuata*. *Plant Physiol.* 152, 1731–1747
48. Onkokesung, N. *et al.* (2012) MYB8 controls inducible phenolamide levels by activating three novel hydroxycinnamoyl-coenzyme A: polyamine transferases in *Nicotiana attenuata*. *Plant Physiol.* 158, 389–407
49. Xu, S. *et al.* (2017) Wild tobacco genomes reveal the evolution of nicotine biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 114, 6133–6138
50. Wang, W. *et al.* (2019) Polyamine catabolism in plants: a universal process with diverse functions. *Front. Plant Sci.* 10, 561
51. Arrebola, E. *et al.* (2011) Chemical and metabolic aspects of antimetabolite toxins produced by *Pseudomonas syringae* pathovars. *Toxins (Basel)* 3, 1089–1110
52. Bender, C.L. *et al.* (1999) *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. *Microbiol. Mol. Biol. Rev.* 63, 266–292
53. Mitchell, R.E. (1976) Isolation and structure of a chlorosis-inducing toxin of *Pseudomonas phaseolicola*. *Phytochemistry* 15, 1941–1947
54. Bachmann, A.S. *et al.* (1998) Inhibition of ornithine decarboxylase activity by phaseolotoxin: Implications for symptom production in halo blight of French bean. *Physiol. Mol. Plant Pathol.* 53, 287–299
55. Arrebola, E. *et al.* (2003) Mangotoxin: a novel antimetabolite toxin produced by *Pseudomonas syringae* inhibiting ornithine/arginine biosynthesis. *Physiol. Mol. Plant Pathol.* 63, 117–127
56. Shanker, S. *et al.* (2017) The virulence-associated protein HsvA from the fire blight pathogen *Erwinia amylovora* is a polyamine amidinotransferase. *J. Biol. Chem.* 292, 21366–21380

57. Zheng, X.Y. *et al.* (2012) Coronatine promotes *Pseudomonas syringae* virulence in plants by activating a signaling cascade that inhibits salicylic acid accumulation. *Cell Host Microbe* 11, 587–596
58. Melotto, M. *et al.* (2006) Plant stomata function in innate immunity against bacterial invasion. *Cell* 126, 969–980
59. Lou, Y.R. *et al.* (2016) Arabidopsis NATA1 acetylates putrescine and decreases defense-related hydrogen peroxide accumulation. *Plant Physiol.* 171, 1443–1455
60. Adio, A.M. *et al.* (2011) Biosynthesis and defensive function of $N^{\delta}$-Acetylornithine, a jasmonate-induced arabidopsis metabolite. *Plant Cell* 23, 3303–3318
61. Ceulemans, E. *et al.* (2021) Pathogen effectors: exploiting the promiscuity of plant signaling hubs. *Trends Plant Sci.* Published online March 2, 2021. https://doi.org/10.1016/j.tplants.2021.01.005
62. Wu, D. *et al.* (2019) A plant pathogen type III effector protein subverts translational regulation to boost host polyamine levels. *Cell Host Microbe* 26, 638–649
63. Denancé, N. *et al.* (2018) Two ancestral genes shaped the *Xanthomonas campestris* TAL effector gene repertoire. *New Phytol.* 219, 391–407
64. Kim, N.H. *et al.* (2013) Pepper arginine decarboxylase is required for polyamine and γ-aminobutyric acid signaling in cell death and defense response. *Plant Physiol.* 162, 2067–2083
65. Tarkowski, Ł.P. *et al.* (2020) γ-Aminobutyric acid and related amino acids in plant immune responses: emerging mechanisms of action. *Plant Cell Environ.* 43, 1103–1116
66. Xian, L. *et al.* (2020) A bacterial effector protein hijacks plant metabolism to support pathogen nutrition. *Cell Host Microbe* 28, 1–10
67. Hewezi, T. *et al.* (2010) *Arabidopsis* spermidine synthase is targeted by an effector protein of the cyst nematode. *Plant Physiol.* 152, 968–984
68. Rodríguez-Kessler, M. *et al.* (2008) Polyamine metabolism in maize tumors induced by *Ustilago maydis*. *Plant Physiol. Biochem.* 46, 805–814
69. Plett, J.M. *et al.* (2020) Mycorrhizal effector PaMiSSP10b alters polyamine biosynthesis in Eucalyptus root cells and promotes root colonization. *New Phytol.* 228, 712–727
70. Bailey, A. *et al.* (2000) Ornithine decarboxylase of *Stagonospora (Septoria) nodorum* is required for virulence toward wheat. *J. Biol. Chem.* 275, 14242–14247
71. Guevara-Olvera, L. *et al.* (1997) Cloning and disruption of the ornithine decarboxylase gene of *Ustilago maydis*: evidence for a role of polyamines in its dimorphic transition. *Microbiology* 143, 2237–2245
72. Michael, A.J. (2015) Biosynthesis of polyamines in eukaryotes, Archaea, and bacteria. In *Polyamines: A Universal Molecular Nexus for Growth, Survival, and Specialized Metabolism* (Tomonobu, K. and Hideyuki, S., eds), pp. 3–14, Springer
73. Tomar, P.C. *et al.* (2013) Cadaverine: a lysine catabolite involved in plant growth and development. *Plant Signal. Behav.* 8, 1559–2316
74. Cona, A. *et al.* (2006) Functions of amine oxidases in plant development and defence. *Trends Plant Sci.* 11, 80–88
75. Fincato, P. *et al.* (2011) Functional diversity inside the *Arabidopsis polyamine oxidase* gene family. *J. Exp. Bot.* 62, 1155–1168
76. Moschou, P.N. *et al.* (2008) Plant polyamine catabolism: the state of the art. *Plant Signal. Behav.* 3, 1061–1066
77. Liu, T. *et al.* (2014) *Oryza sativa* polyamine oxidase 1 back-converts tetraamines, spermine and thermospermine, to spermidine. *Plant Cell Rep.* 33, 143–151
78. Rocha, R.O. and Wilson, R.A. (2019) Essential, deadly, enigmatic: polyamine metabolism and roles in fungal cells. *Fungal Biol. Rev.* 33, 47–57
79. Tassoni, A. *et al.* (2000) Polyamine content and metabolism in *Arabidopsis thaliana* and effect of spermidine on plant development. *Plant Physiol. Biochem.* 38, 383–393
80. Kamada-Nobusada, T. *et al.* (2008) A putative peroxisomal polyamine oxidase, AtPAO4, is involved in polyamine catabolism in *Arabidopsis thaliana*. *Plant Cell Physiol.* 49, 1272–1282
81. Lou, Y.R. *et al.* (2020) Arabidopsis ADC1 functions as an $N^{\delta}$-acetylornithine decarboxylase. *J. Integr. Plant Biol.* 62, 601–613
82. Carper, S.W. *et al.* (1991) Spermidine acetylation in response to a variety of stresses in *Escherichia coli*. *J. Biol. Chem.* 266, 12439–12441
83. Tabor, C.W. (1968) The effect of temperature on the acetylation of spermidine. *Biochem. Biophys. Res. Commun.* 30, 339–342
84. Nguyen, A.Q.D. *et al.* (2015) Elimination of polyamine N-acetylation and regulatory engineering improved putrescine production by *Corynebacterium glutamicum*. *J. Biotechnol.* 201, 75–85

# Appendix V – Supplementary information on plant – pathogen metabolic model

### Table A: Sink reactions active in the model

| Sink reaction ID | Sink metabolite name | Flux value (µmol/g plant/day) | |
|---|---|---|---|
| | | *low density** | *high density** |
| Sink_xyl_M_5mtr_xyl | 5-Methylthio-D-ribose | 0,00064918 | 1,1603E-06 |
| Sink_xyl_M_co2_xyl | $CO_2$ | 2,03502953 | 0,00458446 |
| Sink_xyl_M_cobalt2_xyl | cobalt | 0 | -4,267E-08 |
| Sink_xyl_M_for_xyl | formate | 0 | 4,604E-07 |
| Sink_xyl_M_h2o_xyl | water | 1,43167347 | 0,00374621 |
| Sink_xyl_M_meoh_xyl | methanol | 1,9252E-07 | 3,4409E-10 |
| Sink_xyl_M_nh4_xyl | ammonia | 0,0146614 | 0 |
| Sink_xyl_M_oh_xyl | hydroxide | 0,91578274 | 0 |
| Sink_xyl_M_ptrc_xyl | putrescine | 0,00792738 | 8,1228E-05 |

*low density: $10^4$ cells/g of stem fresh weight / high density: $10^8$ cells/g of stem fresh weight

### Table B: Upper/lower bounds imposed on *R. solanacearum* metabolism

| Reaction ID | Lower Bound | Upper Bound |
|---|---|---|
| **R_NGAME_patho** | 2.37 | NA |
| **R_GLNtex_patho** | 0 | 3.84 |
| **R_ASNtex_patho** | 0 | 0.11 |
| **R_GLCtex_patho** | 0 | 0.967 |
| **R_ASPtex_patho** | 0 | 0.023 |
| **R_PHEtex_patho** | 0 | 0.035 |
| **R_TYRtex_patho** | 0 | 0.017 |
| **R_SUCRtex_patho** | 0 | 0.035 |
| **R_PROtex_patho** | 0 | 0.100 |
| **R_ARGtex_patho** | 0 | 0.055 |
| **R_LYStex_patho** | 0 | 0.070 |
| **R_THRtex_patho** | 0 | 0.044 |
| **R_VALtex_patho** | 0 | 0.074 |
| **R_ILEtex_patho** | 0 | 0.056 |
| **R_LEUtex_patho** | 0 | 0.069 |
| **R_FUMtex_patho** | 0 | 0 |
| **R_ALAtex_patho** | 0 | 0 |

Fluxes in mmol·g pathogen dry weight $^{-1}$·h$^{-1}$

**Other calibrations**

- **Putrescine excretion flux**

Imposed as additional equality constraint on the system

*R_PTRCtex + 0.05 R_GLCtex_patho + 0.038 R_GLNtex_patho = 0*

- **Additional maintenance cost on glucose uptake**

Imposed as ATP hydrolysis integrated in sucrose uptake reaction

2.94 M_atp_c + 2.94 M_h2o_c + k M_glc_D_xyl <--> 1.0 M_glc_D_p_patho + 2.94 M_adp_c + 2.94 M_pi_c + 2.94 M_h_c

k: value depends on plant/pathogen ratio

# Appendix VI – Growth experiment and macroscopic modeling on xylem-mimicking medium

**Bacterial strains and inoculum.** *R. solanacearum* strain used for this study is the wild-type strain GMI1000. The bacteria was plated on Phi medium (with $g.L^{-1}$: Bactopeptone: 10; Casamino acid: 1; Yeast Extract: 1), supplied with glucose 20% (5 $g.L^{-1}$) and TTC (0.05 $g.L^{-1}$) for 48 hours at 28°C. Then, a colony of bacteria was grown in minimal medium (with g.L-1: $FeSO_4$, $7H_2O$: 1.25*10-4, $(NH_4)_2,SO_4$: 0.5 $MgSO_4,7H_2O$: 0.05; $KH_2PO_4$: 3.4) supplemented with glutamate (100 mMC). pH was adjusted to 6.5 using KOH. 250ml shake flasks containing 50 mL of culture were used. Inocula were shaken at 180 rpm on an orbital shaker placed in an incubator at 28°C.

**Shake-flask growth experiment in xylem-mimicking media (XMM)**

GMI1000 was grown in 1L shake flasks shaken at 180 rpm on an orbital shaker placed in an incubator at 28°C. The starting volume culture was 250mL. Culture medium was the same minimal medium as the inocula supplemented with a mixture of difference organic substrates whose relative concentrations mimicks a tomato xylem (Table 1). Dynamic sampling was performed for monitoring biomass growth (OD measurement at 600nm using a spectrometer), substrate consumption (NMR analysis) and products excretion (NMR analysis).

Sampling time was chosen according to the kinetic evolution with intensive sampling during the exponential growth phase. After OD measurement, the collected volume for each point was filtered with a 0.22 µm micro-filter and conserved at -20°C for further analysis.

**Table A: Initial concentration of organic substrates in XMM**

| **Molecule** | **Concentration (mMC)** | **Percentage of organic carbon** | **Percentage in healthy xylem** (Chapter 3) |
|---|---|---|---|
| Glutamine | 34.1 ± 0.38 | 79.8 | 76.7 |
| Lysine | 1.89 ± 0.03 | 4.4 | 4.1 |
| Asparagine | 1.4 ± 0.01 | 3.3 | 3.3 |
| Arginine | 0.98 ± 0.03 | 2.3 | 2.1 |
| Proline | 0.94 ± 0.04 | 2.2 | 1.9 |
| Leucine | 0.74 ± 0.01 | 1.7 | 2.2 |
| Valine | 0.7 ± 0.01 | 1.6 | 1.5 |
| Isoleucine | 0.60 ± 0.004 | 1.4 | 1.2 |
| Phenylalanine | 0.54 ± 0.007 | 1.3 | 0.9 |
| Threonine | 0.38 ± 0.007 | 0.9 | 1.6 |
| Tyrosine | 0.27 ± 0.004 | 0.6 | 0.9 |
| Aspartate | 0.20 ± 0.002 | 0.5 | 0.1 |
| Glucose | 2.8 ± 0.26 | 6.6 | 0.5 |
| Sucrose | 0.68 ± 0.07 | 1.6 | 1.9 |

**NMR analysis of culture filtrate**. Small-size organic metabolites from the liquid culture in shake flaks were quantified by NMR. Samples were analyzed by 1D $^{1}$H NMR on MetaToul analytics platform (UMR5504, UMR792, CNRS, INRAE, INSA 135 Avenue de Rangueil 31077 Toulouse Cedex 04, France), using the Bruker Avance 500 MHz. The samples were kept all along the analysis at a temperature of 280°K because of the presence of glucose. TSP-d4 standard (Sodium 3-(trimethylsilyl)(1-13C,2H4)propanoate) was used as a reference. Resonances of metabolites were manually integrated using Bruker TopSpin software and the concentrations were calculated based on the number of equivalent protons for each integrated signal and on the TSP final concentration.

**Mathematical modeling of the kinetic growth in XMM**

Since glutamine is the most abundant carbon substrate present in XMM, and growth stops after glutamine exhaustion in the media, we assumed GMI1000 to follow a Monod-type growth. The others substrates were assumed to be used as complementary substrates, whose assimilation kinetic is proportional to glutamine assimilation. In overall, the model fits the experimental data, except for glucose. The assimilation kinetic for glucose was thus changed, and assumed to follow its own Monod-type kinetic. The results mathematical model is thus:

$$\begin{cases} \frac{dGln}{dt} = -k_{Gln} * \mu_{\underset{Gln}{max}} . \frac{Gln}{Gln + Ks_{Gln}} * BIOM \\ \frac{dGluc}{dt} = -k_{Gluc} * \mu_{\underset{Gluc}{max}} . \frac{Gluc}{Gluc + Ks_{Gluc}} * BIOM \\ \frac{dOtherSubstrate}{dt} = k_{othersubstrate} * -k_{Gln} * \mu_{\underset{Gln}{max}} . \frac{Gln}{Gln + Ks_{Gln}} * BIOM \\ \frac{dBIOM}{dt} = \mu_{\underset{Gln}{max}} . \frac{Gln}{Gln + Ks_{Gln}} * BIOM + \mu_{\underset{Gluc}{max}} . \frac{Gluc}{Gluc + Ks_{Gluc}} * BIOM \\ \frac{dputr}{dt} = k_{putr_{Gln}} . \mu_{\underset{Gln}{max}} . \frac{Gln}{Gln + Ks_{Gln}} * BIOM + k_{putr_{Gluc}} . \mu_{\underset{Gluc}{max}} . \frac{Gluc}{Gluc + Ks_{Gluc}} * BIOM \end{cases}$$

With:

- Gln the concentration of glutamine (mM)
- $k_{Gln}$ the stoichiometric coefficient linked to glutamine assimilation ($mM.g^{-1}.L$)
- $\mu_{\underset{Gln}{max}}$ the growth rate on glutamine boosted with other amino acids and sucrose ($h^{-1}$)
- $Ks_{Gln}$ the half-saturation constant of glutamine assimilation (mM)
- Gluc the concentration of glucose (mM)
- $k_{Gluc}$ the stoichiometric coefficient linked to glucose assimilation ($mM.g^{-1}.L$)
- $\mu_{\underset{Gluc}{max}}$ the growth rate on glucose ($h^{-1}$) when amino acids are present
- $Ks_{Gluc}$ the half-saturation constant of glucose assimilation (mM)
- OtherSubstrate the concentration of the other organic substrates present in the media (mM)
- $k_{othersubstrate}$ the assimilation constant of other organic substrates present in the media whose assimilation are proportional to the assimilation of glutamine
- BIOM the biomass dry weight concentration ($g.L^{-1}$)
- putr the concentration of putrescine excreted by the bacteria (mM)
- $k_{putr_{Gln}}$ the stoichiometric coefficient of excretion of putrescine due to growth on glutamine ($mM.g^{-1}.L$)
- $k_{putr_{Gluc}}$ the stoichiometric coefficient of excretion of putrescine due to growth on glucose ($mM.g^{-1}.L$)

Parameters of the model (Table 2) were estimated by using the scipy.optimize function of scipy library (version 1.1.1, Nelder-Mead Algorithm) in Python 3.6.

**Table B: Values and units of model's parameters**

| Parameter | Value | Unit |
|---|---|---|
| $k_{Gln}$ | 1.42E-02 | $mM.g^{-1}.L$ |
| $k_{putr_{Gln}}$ | 5.51E-04 | $mM.g^{-1}.L$ |
| $\mu_{max_{Gln}}$ | 3.16E-01 | $h^{-1}$ |
| $Ks_{Gln}$ | 8.45E-01 | mM |
| $k_{Gluc}$ | 0.01091875 | $mM.g^{-1}.L$ |
| $k_{putr_{Gluc}}$ | 0.00057051 | $mM.g^{-1}.L$ |
| $\mu_{max_{Gluc}}$ | 0.08851217 | $h^{-1}$ |
| $Ks_{Gluc}$ | 0.00027337 | mM |
| $k_{Phe}$ | 1.33E-04 | - |
| $k_{Val}$ | 2.74E-04 | - |
| $k_{Leu}$ | 2.54E-04 | - |
| $k_{Arg}$ | 2.03E-04 | - |
| $k_{Sucr}$ | 1.29E-04 | - |
| $k_{Tyr}$ | 6.37E-05 | - |
| $k_{Pro}$ | 3.70E-04 | - |
| $k_{Lys}$ | 2.60E-04 | - |
| $k_{Thr}$ | 1.63E-04 | - |
| $k_{Iso}$ | 2.07E-04 | - |
| $k_{Asn}$ | 4.24E-04 | - |
| $k_{Asp}$ | 8.55E-05 | - |

# Appendix VII – Supplementary Materials of "Genome-scale investigation of the metabolic determinants generating bacterial fastidious growth"

**Supplementary Table 1.** Comparison of the reconstructed network of *X. fastidiosa* CFBP 8418 with a draft network of X. fastidiosa 9a5c.

**Supplementary Table 2.** Comparison of the reconstructed network of *X. fastidiosa* CFBP 8418 with a draft network of *X. fastidiosa* Temecula1.

**Supplementary Text 1.** *Xylella fastidiosa* metabolic network curation

**Supplementary Text 3.** Constraints used on metabolic modeling of *Xylella fastidiosa*.

**Supplementary Text 4.** Mathematical modeling of metabolism. FlexFlux user guide.

**Supplementary Table 1**

Comparison of the reconstructed network of *X. fastidiosa* CFBP 8418 with a draft network of *X. fastidiosa* 9a5c.

| Reaction ID | Reaction name | Reaction formula | EC number | Pathway | Gene |
|---|---|---|---|---|---|
| *Reactions absent in 9a5c strain but identified in CFBP 8418* | | | | | |
| R_NNAMr | nicotinamidase | 1 M_h2o_c + 1 M_ncam_c -> 1 M_nh4_c + 1 M_nac_c | 3.5.1.19 | NAD NADP biosynthesis | ( XFCFBP8418_020780 ) |
| R_SHCHD2 | sirohydrochlorin dehydrogenase (NAD) | 1 M_nad_c + 1 M_dscl_c -> 1 M_scl_c + 1 M_h_c + 1 M_nadh_c | 1.3.1.76 | Porphyrin metabolism | ( XFCFBP8418_010970 ) |
| R_SHCHF | sirohydrochlorin ferrochelatase | 1 M_scl_c + 1 M_fe2_c -> 1 M_sheme_c + 3 M_h_c | 4.99.1.4 | Porphyrin metabolism | ( XFCFBP8418_010970 ) |
| R_UPP3MT | uroporphyrinogen methyltransferase | 2 M_amet_c + 1 M_uppg3_c -> 1 M_dscl_c + 2 M_ahcys_c + 1 M_h_c | 2.1.1.107 | Porphyrin metabolism | ( XFCFBP8418_010970 ) |
| R_MTAP | 5'-methylthioadenosine:orthophosphate methylthio-D-ribosyltransferase | 1 M_5mta_c + 1 M_pi_c -> 1 M_5mdr1p_c + 1 M_ade_c | 2.4.2.28 | Met Cys Ser metabolism | ( XFCFBP8418_021330 ) |
| *Reactions absent in CFBP 8418 strain but identified in 9a5c* | | | | | |
| R_ALHD4 | Aldehyde dehydrogenase (butanal, NAD) | 1 M_nad_c + 1 M_h2o_c + 1 M_btal_c -> 1 M_nadh_c + 1 M_1boh_c + 2 M_h_c | 1.2.1.57 | Alternate carbon metabolism | ( WP_010894208.1 ) or ( WP_010894291.1 ) or ( WP_038200764.1 ) |
| R_ACTNabc | R-acetoin transport via ABC system | 1 M_atp_c + 1 M_h2o_c + 1 M_actn_R_e -> 1 M_adp_c + 1 M_h_c + 1 M_pi_c + 1 M_actn_R_c | NA | Alternate carbon metabolism | ( WP_023906641.1 ) |
| R_ACTNabc1 | R-acetoin efflux via ABC system | 1 M_atp_c + 1 M_h2o_c + 1 M_actn_R_c -> 1 M_adp_c + 1 M_h_c + 1 M_pi_c + 1 M_actn_R_e | NA | Alternate carbon metabolism | ( WP_023906641.1 ) |
| R_CINNDO | Cinnamate dioxygenase | 1 M_cinnm_c + 1 M_h_c + 1 M_nadh_c + 1 M_o2_c -> 1 M_cenchddd_c + 1 M_nad_c | NA | Alternate carbon metabolism | ( WP_010893973.1 ) |
| R_PPPNDO | Phenylpropanoate dioxygenase | 1 M_h_c + 1 M_nadh_c + 1 M_o2_c + 1 M_pppn_c -> 1 M_cechddd_c + 1 M_nad_c | NA | Alternate carbon metabolism | ( WP_010893973.1 ) |
| R_POAACR | Peroxyaminoacrylate reductase | 1 M_nadh_c + 1 M_poaac_c -> 1 M_3amac_c + 1 M_h2o_c + 1 M_nad_c | NA | Nucleotide Salvage Pathway | ( WP_010892886.1 ) |
| R_EXOGLAS | Exoglucanase | 1 M_14bglucan_e + 5 M_h2o_e -> 6 M_cellb_e | 3.2.1.91 | Plant cell wall degradation | ( WP_010893773.1 ) |
| R_GLCDpp | glucose dehydrogenase (ubiquinone 8 as acceptor) | 1 M_glc_D_p + 1 M_h2o_p + 1 M_q8_c -> 1 M_h_p + 1 M_glcn_p + 1 M_q8h2_c | 1.1.5.2 | Oxidative phosphorylation | ( WP_010894211.1 ) |

**Supplementary Table 2**

Comparison of the reconstructed network of *X. fastidiosa* CFBP 8418 with a draft network of *X. fastidiosa* Temecula1.

| Reaction ID | Reaction name | Reaction formula | EC number | Pathway | Gene |
|---|---|---|---|---|---|
| *Reactions absent in Temecula1 strain but identified in CFBP 8418* | | | | | |
| R_DHPS2 | Dihydropteroate synthase | 1 M_4abz_c + 1 M_6hmhptpp_c -> 1 M_dhpt_c + 1 M_ppi_c | 2.5.1.15 | Folate biosynthesis and metabolism | ( XFCFBP8418_014930 ) |
| R_ADCL | 4-aminobenzoate synthase | 1 M_4adcho_c <-> 1 M_pyr_c + 1 M_4abz_c + 1 M_h_c | 4.1.3.38 | Folate biosynthesis and metabolism | ( XFCFBP8418_014810 ) |
| R_HSTPTr | Histidinol phosphate transaminase | 1 M_imacp_c + 1 M_glu_L_c <-> 1 M_akg_c + 1 M_hisp_c | 2.6.1.9 | His metabolism | ( XFCFBP8418_020310 ) |
| R_5DGLCNR | 5-dehydro-D-gluconate reductase | 1 M_5dglcn_c + 1 M_h_c + 1 M_nadph_c <-> 1 M_nadp_c + 1 M_glcn_c | 1.1.1.69 | Alternate carbon metabolism | ( XFCFBP8418_010080 ) |
| *Reactions absent in CFBP 8418 strain but identified in Temecula1* | | | | | |
| R_POAACR | Peroxyaminoacrylate reductase | 1 M_nadh_c + 1 M_poaac_c -> 1 M_3amac_c + 1 M_h2o_c + 1 M_nad_c | NA | Nucleotide Salvage Pathway | ( WP_010892886.1 ) |
| R_EXOGLAS | Exoglucanase | 1 M_14bglucan_e + 5 M_h2o_e -> 6 M_cellb_e | 3.2.1.91 | Plant cell wall degradation | ( WP_010893773.1 ) |

# *Xylella fastidiosa* metabolic network curation

## I. Curation process

The network was curated pathway by pathway, starting by central and canonical pathways.

For each reaction in a pathway:

- Name, ID, FormulaIds, FormulaNames, EC, pathways, type were verified using KEGG, MetaCyc and BiGG. For the reversibility, MetaCyc (particularly EcoCyc) was used as reference.
- Changes were explained in comments.
- If the prop2 score = 100, there was an orthology for all the necessary genes. The reaction was kept in the network except if it was a duplicate.
- If the prop2 score was below 100, literature and metabolic databases (KEGG, MetaCyc) were used to decide if the reaction was kept. BLASTP and Conserved Domains Databases were used to find if a *X. fastidiosa* gene could complete the enzymatic complex.
- Depending on the results of this database research, a score was attributed to the reaction to represent the probability of this reaction, and it was indicated if the reaction has to be deleted or kept.

After curation, the orthology was different than the one generated by the automatic reconstruction. So, a column “final genes” and “final orthology” were added. Pathways names were manually rewritten to standardize them in the network, in the column “pathways”. Automatically generated subsystems were kept in the column “subsystem”.

When an information was missing, the mention No_Assignment was attributed.

## II. Network guide

- **#Id**

Reaction ID usually come from BiGG databases. Some ID were manually written when this was a new reaction in accordance with the nomenclature.

- **FormulaIds**

The BiGG database nomenclature was used. The reversibility of the reaction was manually curated but is sometimes uncertain and should be treated cautiously.

- **EC**

EC number were verified using KEGG, as some of the EC numbers generated automatically were out-of-date. When there was no EC for a reaction, No_Assignment was attributed.

- **type**

Indicates if the reaction is internal (inside the cell) or a transport/exchange reaction, or external (outside the cell).

- **prop**

From the automatic reconstruction, indicates the proportion (%) of genes from the reference organism that were conserved.

- **prop2**

From the automatic reconstruction, indicates the proportion (%) of essential genes from the reference organism that were conserved. If the score is 100, there are all orthology evidences necessary for the reaction.

- **ref_no**

Indicates from each organism of the automatic reconstruction the orthology was found.

| ref_no | strain | model |
|---|---|---|
| 1 | *Escherichia coli* K-12 | iJO1366 |
| 2 | *Ralstonia solanacearum* GMI1000 | iRP1476 |
| 3 | *Pseudomonas aeruginosa* PAO1 | iMO1086 |
| 4 | *Ralstonia eutropha* H16 | RehMBEL1391 |
| 5 | *Bacillus subtilis* 168 | iYO844 |

- **quality score**

A score from 1 to 7 was given to assess the probability of the reaction. A score of 1 is for the reaction with high probability, and a score of 7 is for simple assumptions that would need to be studied more precisely.

Some reactions were generated using the tool fast automated reconstruction (Nucleic Acids Research, gky537, https://doi.org/10.1093/nar/gky537). As no precise information was given for the quality of the reaction, it was ranked score 3.

- **Quality score**

Our own quality score, adapted to the network.

| Situation | Score | Thiele and Palsson score associated | Validation level |
|---|---|---|---|
| Complete orthology (prop score 100) | 1 | 2 | **Validated** |
| Spontaneous reaction | 1 | 2 | **Validated** |
| Sufficient orthology (prop2 score 100) | 2 | 2 | **Validated** |
| Evidence of the reaction in publications | 3 | 2, 3, 4 | **Validated** |
| Reaction generated by fast automated reconstruction | 3 | 2, 3, 4 | **Validated** |
| Missing gene non-essential | 4 | 2 | **Validated** |
| Missing gene found in the target genome | 5 | 2 | **Validated** |
| prop2 score ≥ 50 and reaction connected to others | 6 | 1 | **Assumed** |
| reaction necessary to complete a pathway | 6 | 1 | **Assumed** |
| Arbitrary choice, modeling | 7 | 1 | **Assumed** |
| Potential reaction: no evidence and no connection to the network | 7 | 1 | **Assumed** |

- **Thiele and Palsson score**

Confidence score proposed by Thiele and Palsson, Nat Protoc., 2010, related to our quality score.

To choose between 2, 3, 4:

| | |
|---|---|
| 2 | Physiological data (indirect evidence from secretion product, defined medium…) |
| 3 | Genetic data (knock-out, knock-in, over-expression) |
| 4 | Biochemical data (direct evidence like protein purification, biochemical assays…) |

- **reference (PMID)**

The PubMed ID (PMID) was used each time a publication was used to cure a reaction: evidence of the reaction, characterization of a macromolecule, *X. fastidiosa* specificity for example.

- **final genes**

The final genes from *X. fastidiosa* generating the reaction are given. If a gene is unknown, No_Assignment was used.

- **final orthology**

The final ortholog from a reference organism (sometimes not in the list of the five reference organisms) that was used to build the relation between gene and reaction.

- **origin**

It is precised if the reaction was generated by the automatic reconstruction, was manually added or comes from fast automated reconstruction

- **pathways**

A standardized version of the pathways was manually added.

**Constraints used on metabolic modeling of *Xylella fastidiosa***

The following specific constraints were used to perform simulations.

Each carbon exchange reaction was set to zero, except the carbon source used in the simulation:

R_EX_gal_e, R_EX_asp_L_e, R_EX_pro_L_e, R_EX_tre_e, R_EX_man_e, R_EX_glyc_e, R_EX_xyl_D_e, R_EX_glu_L_e, R_EX_mal_L_e, R_EX_rib_D_e, R_EX_fru_e, R_EX_ac_e, R_EX_glc_D_e, R_EX_asn_L_e, R_EX_uri_e, R_EX_adn_e, R_EX_cit_e, R_EX_inost_e, R_EX_gly_glu_L_e, R_EX_ala_L_e, R_EX_acac_e, R_EX_lyx_L_e, R_EX_pyr_e, R_EX_dextrin_e, R_EX_4abut_e, R_EX_gam_e, R_EX_arg_L_e, R_EX_his_L_e, R_EX_orn_e, R_EX_chitin_polymer_e, R_EX_gln_L_e

As no carbon uptake is known for *X. fastidiosa*, a minimal L-glutamine import (*R_EX_gln_L_e*) was put as the objective function. Growth rate and protein excretion flux were estimated from a published growth assay on defined medium (XDM2*) (Leite, Andersen, and Ishida 2004). To simplify the system, the protein excretion flux was defined as an excretion of the protein LesA, described as a major excreted product of the bacterium (Nascimento et al. 2016). EPS excretion flux was estimated from a published quantification assay (Navarrete and De La Fuente 2014). ATP maintenance was assumed to be proportional to the growth rate, so was estimated to be 3.073 $mmol.g_{DW}^{-1}.d^{-1}$, taking as reference that *R. solanacearum* ATP maintenance is 8.39 $mmol.g_{DW}^{-1}.h^{-1}$ for a growth rate of 0.439 $h^{-1}$.

Flux values were converted by day unit instead of hour unit since values are small. It led to the following constraints:

| **Reaction** | **Flux ($mmol.g_{DW}^{-1}.d^{-1}$)** |
|---|---|
| R_DM_BIOMASS_c | 0.1608 |
| R_DM_EPS_XF_e | 1.296 |
| R_DM_LesA_e | 0.002405 |
| R_ATPM | 3.0731 |

**References**

Leite, Breno, Peter Craig Andersen, and Maria Lucia Ishida. 2004. “Colony Aggregation and Biofilm Formation in Xylem Chemistry-Based Media for Xylella Fastidiosa.” *FEMS Microbiology Letters* 230(2): 283–90.

Nascimento, Rafael et al. 2016. “The Type II Secreted Lipase/Esterase LesA Is a Key Virulence Factor Required for Xylella Fastidiosa Pathogenesis in Grapevines.” *Scientific Reports* 6(January): 1–17. http://dx.doi.org/10.1038/srep18598.

Navarrete, Fernando, and Leonardo De La Fuente. 2014. “Response of Xylella Fastidiosa to Zinc: Decreased Culturability, Increased Exopolysaccharide Production, and Formation of Resilient Biofilms under Flow Conditions.” *Applied and Environmental Microbiology* 80(3): 1097–1107.

# Mathematical modeling of metabolism

## 1. Quasi-steady state assumption [1]

A mass balance on the cell metabolism implies the formulation of equation 1.

$$\frac{dM}{dt} = K \cdot v \cdot B$$
$$v_i \geq 0 \; if \; v_i \; irreversible, i \in \{1; \dots; r\} \qquad (1)$$

M: metabolites concentrations vector
K: stoichiometric matrix of the metabolite network (size m × r)
v: reactions fluxes vector
B: biomass
Usual units are mM for M, $mM \cdot h^{-1} \cdot g_{dry\ weight}$ for v and $g_{dry\ weight}$ for B.

Biomass is defined in equation 2. In a metabolic network, it will be artificially represented by a biomass equation. [2]

$$B = M_{w,biomass\ metabolites} \cdot [\text{Biomass metabolites}] \qquad (2)$$

$M_{w,biomass\ metabolites}$: molecular weight of metabolites constituting biomass
[Biomass metabolites]: metabolites constituting biomass
Metabolites constituting biomass are amino acids, cofactors, lipids and nucleic acids.

Metabolites can be separated into external metabolites consisting of substrates and products, internal metabolites and biomass. (eq. 3)

$$\frac{d\begin{pmatrix} S \\ C \\ P \\ B \end{pmatrix}}{dt} = \begin{pmatrix} K_S \\ K_C \\ K_P \\ K_B \end{pmatrix} \cdot v \cdot B$$
$$v_i \geq 0 \; if \; v_i \; irreversible, i \in \{1; \dots; r\} \qquad (3)$$

v: reaction flux vector
B: biomass
S: extracellular substrate concentration vector
C: internal metabolite concentration vector
P: excreted product concentration vector
$K_S$, $K_C$, $K_P$, $K_B$: subset of the complete stoichiometric matrix K corresponding to, respectively, extracellular substrates S, internal metabolites C, excreted products P and biomass B

Since the mathematical expression of v and the kinetics parameters related are often not known due to lack of experimental data, a quasi-stationary-state approximation (QSSA) is assumed: the metabolic concentration of internal metabolites are assumed constant (dC/dt=0). Biologically speaking, internal metabolites are assumed to be consumed immediately as they are produced. Only external metabolites and macromolecules constituting biomass such as membrane lipids and proteins are assumed to have slower dynamics leading to accumulation. This approximation is usually verified at the exponential phase of a bacterial growth. It allows to obtain the following system of equations (eq. 4).

$$\begin{cases} \dfrac{d\begin{pmatrix} S \\ P \\ B \end{pmatrix}}{dt} = \begin{pmatrix} K_S \\ K_P \\ K_B \end{pmatrix} \cdot v \cdot B \\ \dfrac{dC}{dt} = K_c \cdot v \cdot B \end{cases} \quad v_i \geq 0 \; if \; v_i \; irreversible, \; i \in \{1; \dots; r\} \qquad (3')$$

$$\xrightarrow[\frac{dC}{dt}=0]{} \quad \begin{cases} \dfrac{d\begin{pmatrix} S \\ P \\ B \end{pmatrix}}{dt} = \begin{pmatrix} K_S \\ K_P \\ K_B \end{pmatrix} \cdot v \cdot B \\ 0 = K_C \cdot v \end{cases} \quad v_i \geq 0 \; if \; v_i \; irreversible, \; i \in \{1; \dots; r\} \qquad (4)$$

## 2. Flux Balance Analysis [1]

The metabolic fluxes are now the solution v of the system $\mathbf{K_c} \cdot \mathbf{v} = \mathbf{0}$. However, this system is an under-determined system: in realistic metabolic model, there are more reactions than compounds (r>m), which leads to an infinity of solutions v. To reach a point within the solution space, an objective function **Z=f(v)** must be maximized or minimized. Generally, the objective function is assumed linear of the form $\mathbf{Z=c^Tv}$ with c the vector of weights, representing the contribution of each reaction to the objective function. An optimization solver will then be used to determine the metabolic fluxes upon a minimization or maximization objective function. This resolution of the system is called **Flux Balance Analysis (FBA)**.

Different scenarios of objective function can be used depending on the biological question investigated. Frequently, it is considered that microorganisms evolved by maximizing their growth rates ($v_B$), so the hypothesis that metabolic fluxes are shaped by growth maximization is reasonable. The scenario 1 is then widely used. In this scenario, a constant value is given to the substrate uptake flux ($v_S$) to prevent an unlimited biomass production. This value could be determined experimentally or approximated from literature. It is also possible to put as objective the minimization of the uptake rate and give a constant value (experimentally determined) for biomass growth (scenario 2).

Generally, FBA method gives a biomass flux superior than the experimental one, because there are unintegrated reactions such as housekeeping functions. An artificial ATP maintenance reaction is usually added in a metabolic network to unify fluxes between experimental and modeling data. ATP maintenance reaction consists in ATP hydrolysis: ATP → ADP + P. The flux value of this will be specified in the optimization problem.

$$\max v_B \quad \begin{cases} Kc \cdot v = 0 \\ v_i \geq 0 \text{ if } v_i \text{ irreversible} \\ i \in \{1; \ldots; r\} \\ v_S = constant \end{cases}$$

**Scenario 1, FBA**

$$\min v_S \quad \begin{cases} Kc \cdot v = 0 \\ v_i \geq 0 \text{ if } v_i \text{ irreversible} \\ i \in \{1; \ldots; r\} \\ v_B = constant \end{cases}$$

**Scenario 2, FBA**

The resolution of the system gives metabolic fluxes satisfying a maximization of biomass synthesis (or minimization of substrate uptake). However, some fluxes can take several values to reach this optimal biomass synthesis, which are not always biologically relevant. This is particularly the case in the presence of cycles in the metabolic network. Another formulation of FBA can then be formulated to find the minimal fluxes sustaining a maximal biomass flux (or minimal substrate uptake), with the first objective value integrated as a new constraint in a second optimization problem. An example is given in scenarios 3 and 4.

$$\max v_B \quad \begin{cases} Kc \cdot v = 0 \\ v_i \geq 0 \text{ if } v_i \text{ irreversible} \\ i \in \{1; \ldots; r\} \\ v_S = constant \end{cases}$$

$$\min \sum_i |v_i| \quad \begin{cases} Kc \cdot v = 0 \\ v_i \geq 0 \text{ if } v_i \text{ irreversible} \\ i \in \{1; \ldots; r\} \\ v_S = constant \\ v_B = v_{B,solved} \end{cases}$$

**Scenario 3, FBA**

$$\min v_S \quad \begin{cases} Kc \cdot v = 0 \\ v_i \geq 0 \text{ if } v_i \text{ irreversible} \\ i \in \{1; \ldots; r\} \\ v_B = constant \end{cases}$$

$$\min \sum_i |v_i| \quad \begin{cases} Kc \cdot v = 0 \\ v_i \geq 0 \text{ if } v_i \text{ irreversible} \\ i \in \{1; \ldots; r\} \\ v_S = constant \\ v_S = v_{S,solved} \end{cases}$$

**Scenario 4, FBA**

### 3. Flux Variability Analysis [3]

A drawback of FBA is that sometimes, several solutions v can satisfy the objective function. **Flux Variability Analysis (FVA)** allows finding the minimal and maximal flux that can carry each reaction while satisfying the optimal objective value found in FBA (e.g. maximal biomass flux or minimal substrate uptake rate).

The value determined through FBA is used as a new constraint, and FBA is computed for each metabolic flux, a first time by maximizing the metabolic flux and the second time by minimizing it.

To relax the system and enlarge the possibilities of fluxes variations, a deviation from the optimal flux can also be authorized: flux = optimal flux ± ε optimal flux (e.g ε = 1%). FVA scenario is presented below.

min or max $v_k$

$$Kc \cdot v = 0$$

$$v_i \geq 0 \text{ if } v_i \text{ irreversible}$$

$$i \in \{1; \dots; r\}$$

$$v_S = constant$$

for i in 1 to r

min $v_i$

max

$$Kc \cdot v = 0$$

$$v_i \geq 0, i \in \{1; \dots; r\}$$

$$v_s = constant$$

$$v_{k,solved} - \varepsilon v_{k,solved} < v_k < v_{k,solved} + \varepsilon v_{k,solved}$$

**FVA**

### 4. Gene Deletion Study [4]

In genome-scale metabolic networks, metabolic reactions are usually associated with the enzymatic complex catalyzing the reaction. It is then possible to determine if a gene is essential (*in silico*) to sustain growth or perform any other metabolic function. The algorithm to perform this essay is the following:

For each gene in the metabolic network:

- It is determined for which reactions the gene is necessary.

In the example below, the gene A is necessary for the metabolic reaction R1 but not for R2.

| Reaction | Enzymatic complex |
|---|---|
| R1 | ( Gene A and Gene B ) or ( Gene A and Gene C ) |
| R2 | ( Gene A ) or ( Gene D ) |

- For each reaction the gene is necessary, the flux of the reaction is set to zero.

On the previous example, the flux for R1 will be constrained to zero.

- FBA is then, performed with, for example, biomass maximization as objective. If no solution is obtained, then the gene is essential (*in silico*) to sustain growth. Otherwise the gene is facultative.

Similarly, the essentiality of reactions can be assessed by performing FBA with setting one by one metabolic fluxes to zero.

### 5. *In silico* comparison of carbon substrates

FBA can be alternatively performed on a range of identified carbon substrates. To compare rigorously the maximal growth rate generated by a given substrate, the uptake rates must be normalized by their number of carbon par molecule. For example, if we put 20 mmol·$g_{DW}^{-1}$·$h^{-1}$ for glutamine (5 carbons), 16.67 mmol·$g_{DW}^{-1}$·$h^{-1}$ will be put for glucose (6 carbon), to have the same carbon uptake rate (100 mmol C·$g_{DW}^{-1}$.$h^{-1}$).

The relative growth rates can be compared by computing the ratio:

$$Relative\ growth\ rate_{substrate\ i}\ (\%) = \frac{v_{biomass,substrate\ i}}{v_{biomass,substrate\ with\ highest\ biomass\ flux}} \times 100$$

## 6. Efficiency analysis

The efficiency of all the reactions available in a metabolic network to produce a molecule/macromolecule of interest, or biomass, can be analyzed using FBA.

The maximization of the molecule/macromolecule of interest production flux, or biomass flux, is put as objective. Additional constraints in the model (as ATP maintenance) are removed.

FBA solution fluxes include the $CO_2$ excretion, which represent the carbon losses from metabolic reactions.

The percentage of carbon loss upon the production, and subsequently the yield of production, can be determined:

$$Carbon\ loss\ (\%) = \frac{v_{CO_2,excreted}}{v_{substrate\ uptake} \times Number\ of\ carbon\ in\ substrate} \times 100$$

$$Yield\ of\ production\ (\%) = 100 - Carbon\ loss\ (\%)$$

**References**


1. Orth JD, Thiele I, Palsson BØ. What is flux balance analysis? Nat Biotechnol. 2010 Mar 1;28(3):245–8.
2. Feist AM, Palsson BO. The biomass objective function. Curr Opin Microbiol. 2010;13(3):344–9.
3. Mahadevan R, Schilling CH. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab Eng. 2003 Oct 1;5(4):264–76.
4. Segrè D, Vitkup D, Church GM. Analysis of optimality in natural and perturbed metabolic networks. Proc Natl Acad Sci U S A. 2002 Nov 12;99(23):15112–7.

# FlexFlux user guide

Guide to perform the simulations presented in
"Gerlin et al., *Genome-scale investigation of the metabolic determinants generating bacterial fastidious growth*"

If you want more information about the mathematical systems used in metabolic modeling (FBA, FVA, efficiency…) please consult the supplementary file "Mathematical modeling of metabolism".

## 1. Installing CPLEX.

The optimization software package IBM ILOG CPLEX Optimization Studio is used to solve linear programming problems in the study.

The software is free for students and academics upon registration to the IBM website, and can be easily downloaded there: https://my15.digitalexperience.ibm.com/b73a5759-c6a6-4033-ab6b-d9d4f9a6d65b/dxsites/151914d1-03d2-48fe-97d9-d21166848e65/technology/data-science.

## 2. Downloading and installing FlexFlux

FlexFlux is an open-source tool for metabolic flux and regulatory analysis, which is described there: http://lipm-bioinfo.toulouse.inra.fr/flexflux/documentation.html.

The webpage http://lipm-bioinfo.toulouse.inra.fr/flexflux/installation.html explains all the steps to download and install FlexFlux on Linux or Windows.
Briefly:

- Java 7 or higher must be installed on the computer
  http://www.oracle.com/technetwork/java/javase/downloads/index.html
- FlexFlux executable must be downloaded here:
  http://lipm-bioinfo.toulouse.inra.fr/flexflux/download/FlexFlux2.2.1.zip
- On Windows, FlexFlux is installed by unzipping the downloaded file in the folder of your choice.
- On Linux, you must have previously installed a solver, and then, you just need to run the commands from the folder where you put FlexFlux/.

  ```
  chmod +x Flexflux.sh
  chmod +x GraphicalFlexflux.sh
  ```

- The procedure to link CPLEX with FlexFlux is explained here:
  http://lipm-bioinfo.toulouse.inra.fr/flexflux/installation.html

## 3. Using FlexFlux

To perform metabolic modeling simulations (FBA, FVA, Gene Deletion Study…) using FlexFlux, you only need to:

1. Have the **SBML file of your metabolic network**
2. Have the **constraint file** specifying the constraints and the optimization problem to solve
3. Type a command line.

Alternatively, the graphical version of FlexFlux can be used using the command line GraphicalFlexflux.sh (Linux) or GraphicalFlexflux (Windows).

### 3.1. Performing FBA with substrate uptake minimization

To perform FBA on *X. fastidiosa*, you need the file glnConstraints.tab given in the folder flexflux/FBA/xfas in the github repository https://github.com/lgerlin/xfas-metabolic-model. The file specifies the two objectives functions: first, the minimization of glutamine (the carbon substrate) uptake and then minimization of the fluxes sum. All the other carbon substrates uptake apart from glutamine are constrained to 0, and ATP maintenance (R_ATPM), EPS production (R_DM_EPS_XF_e), virulence protein production (R_DM_LesA_e), biomass (R_DM_BIOMASS_c) are put to values determined by literature (see our Supplementary File 7 "Constraints used on metabolic modeling of *Xylella fastidiosa*").

The SBML file is given in the same folder.

To perform FBA, you need to place the constraints and SBML files in the same folder than the FlexFlux folder, and then use the command line:

- On Linux

```
Flexflux.sh FBA -cons glnConstraints.tab -out FBA_xfas_gln.txt -s XF_network.xml -sol CPLEX -ext -plot
```

- On Windows

```
Flexflux FBA -cons glnConstraints.tab -out FBA_xfas_gln.txt -s XF_network.xml -sol CPLEX -ext -plot
```

Alternatively, you can execute the command from the folder were the SBML and constraints file are located, and replace Flexflux or Flexflux.sh by /path/Flexflux or /path/Flexflux.sh, with /path/ the complete path to the Flexflux folder. Inversely, you can execute the command from the folder where Flexflux is located, and replace glnConstraints.tab and XF_network.xml by /path/glnConstraints.tab and /path/XF_network.xml, with /path/ the complete path to the Flexflux folder.

The result of the simulation (flux values) will be displayed graphically (as the argument "-plot" is activated in the command line). It will also be given in the file FBA_xfas_gln.txt. The objective value given is the sum of fluxes. The value of the glutamine uptake flux can be found at R_EX_gln_L_e line.

Similarly, FBA can be performed on the other metabolic network *R. solanacearum* with the constraint file available in flexflux/FBA/rsol (constraint values were put according to *R. solanacearum* metabolic model study by Peyraud et al., 2016) and the following command line:

- On Linux

```
Flexflux.sh FBA -cons glnConstraints.tab -out FBA_rsol_gln.txt -s rsolGMI1000_20190128_forModelling.xml -sol CPLEX -ext -plot
```

- On Windows

```
Flexflux.sh FBA -cons glnConstraints.tab -out FBA_rsol_gln.txt -s rsolGMI1000_20190128_forModelling.xml -sol CPLEX -ext -plot
```

### 3.2. Performing FBA to compare substrates

In the study, biomass production rates are compared for all *X. fastidiosa* carbon sources identified and being modeled.

Constraint files were generated for each substrate on flexflux/FBA/xfas_alternative-substrates. The objective function is now biomass maximization, and a value of substrate uptake was given. The value is normalized to provide the same value of carbon uptake (100 mmol $C \cdot h^{-1} \cdot g_{dry\ weight}^{-1}$). The minimization of the sum of fluxes was removed as only the biomass flux is studied there.

A similar command than the previous one can be used. glnConstraints.tab must be replaced by the appropriate substrate file (e.g glcConstraints.tab to study glucose) in the command line and the output file must also be renamed depending on the substrate (e.g FBA_xfas_glc.txt to study glucose).

For example, for glucose:

- On Linux

```
Flexflux.sh FBA -cons glcConstraints.tab -out FBA_xfas_glc.txt -s XF_network.xml -sol CPLEX -ext -plot
```

- On Windows

```
Flexflux FBA -cons glcConstraints.tab -out FBA_xfas_glc.txt -s XF_network.xml -sol CPLEX -ext -plot
```

### 3.3. Performing FBA to determine network efficiency (metabolic yields)

To determine the metabolic losses associated with biomass production on *X. fastidiosa*, you need the file biomassConstraints.tab given in the folder flexflux/FBA/xfas_efficiency-analysis in the github repository https://github.com/lgerlin/xfas-metabolic-model. The SBML file is also provided in the same folder.

The following command line is used:

- On Linux

```
Flexflux.sh FBA -cons biomassConstraints.tab -out FBA_xfas_biomass_efficiency.txt -s XF_network.xml -sol CPLEX -ext -plot
```

- On Windows

```
Flexflux FBA -cons biomassConstraints.tab -out FBA_xfas_biomass_efficiency.txt -s XF_network.xml -sol CPLEX -ext -plot
```

Similarly, metabolic losses associated with EPS or virulence protein can be determined. biomassConstraints.tab must be replaced by epsConstraints.tab or proteinConstraints.tab, and FBA_xfas_biomass_efficiency.txt must be replaced by FBA_xfas_eps_efficiency.txt or FBA_xfas_protein_efficiency.txt.

The same simulations can be performed on *R. solanacearum*, with the same command lines. The constraint and network files can be found in the folder flexflux/FBA/rsol_efficiency-analysis in the github repository https://github.com/lgerlin/xfas-metabolic-model. In the command lines, XF_network.xml must be replaced by rsolGMI1000_20190128_forModelling.xml.

### 3.4. Performing FBA to study the impact of FBPase activity on growth

The impact of the absence of inefficient FBPase is examined in the study. To compare optimal and suboptimal FBPase flux, constraint and SBML files are provided in flexflux/FBA/xfas_fbpase.

Considering an optimal FBPase, the following command line is used:

- On Linux

```
Flexflux.sh FBA -cons fbpoptConstraints.tab -out FBA_xfas_gln_fbpopt.txt -s XF_network.xml -sol CPLEX -ext -plot
```

- On Windows

```
Flexflux FBA -cons fbpoptConstraints.tab -out FBA_xfas_gln_fbpopt.txt -s XF_network.xml -sol CPLEX -ext -plot
```

Considering a suboptimal FBPase, the following command line is used:

- On Linux

```
Flexflux.sh FBA -cons fbpsuboptConstraints.tab -out FBA_xfas_gln_fbpsubopt.txt -s XF_network.xml -sol CPLEX -ext -plot
```

- On Windows

```
Flexflux FBA -cons fbpsuboptConstraints.tab -out FBA_xfas_gln_fbpsubopt.txt -s XF_network.xml -sol CPLEX -ext –plot
```

### 3.5. Performing FVA

To perform FVA on *X. fastidiosa*, you need the file glnConstraints.tab given in the folder flexflux/FVA/xfas in the github repository https://github.com/lgerlin/xfas-metabolic-model. The constraints are similar to the previous files. To have an appropriate comparison with other bacteria, constraints related to virulence (protein and EPS production) were removed. The ATP maintenance and glutamine uptake values chosen are identical to the ones determined experimentally for *R. solanacearum*. The objective function consists in maximizing biomass.

The SBML file is given in the same folder.

To perform FVA, you need to place the constraints and SBML files in the same folder than the FlexFlux folder, and then use the command line:

- On Linux

```
Flexflux.sh FVA -cons glnConstraints.tab -out FVA_xfas_gln.txt -lib 1 -s XF_network.xml -sol CPLEX -ext -plot
```

- On Windows

```
Flexflux FVA -cons glnConstraints.tab -out FVA_xfas_gln.txt -lib 1 -s XF_network.xml -sol CPLEX -ext -plot
```

The argument -lib 1 authorizes a deviation of the optimality of 1%.

FVA can also be performed on *R. solanacearum* and *E. coli* with the files given in flexflux/FVA/rsol and flexflux/FVA/ecol, and the following command line:

**For *R. solanacearum***

- On Linux

```
Flexflux.sh FVA -cons glnConstraints.tab -out FVA_rsol_gln.txt -lib 1 -s rsolGMI1000_20190128_forModelling.xml -sol CPLEX -ext -plot
```

- On Windows

```
Flexflux FVA -cons glnConstraints.tab -out FVA_rsol_gln.txt -lib 1 -s rsolGMI1000_20190128_forModelling.xml -sol CPLEX -ext –plot
```

**For *E. coli***

- On Linux

```
Flexflux.sh FVA -cons constraints.txt -out FVA_ecol.txt -lib 1 -s ecoli-orth2011.xml -sol CPLEX -ext –plot
```

- On Windows

```
Flexflux FVA -cons constraints.txt -out FVA_ecol.txt -lib 1 -s ecoli-orth2011.xml -sol CPLEX -ext –plot
```

### 3.6. Performing Gene Deletion Study (KO)

Gene Deletion Study or KO (knock-out) is performed in two different environmental conditions: glucose medium or protein-rich medium. The method is the same, only the constraint file is changed.

To perform Gene Deletion Study on *X. fastidiosa*, you need the file constraints.tab given in the folder flexflux/KO/xfas_glc in the github repository https://github.com/lgerlin/xfas-metabolic-model. The constraints are similar to the previous files. To have an appropriate comparison with other bacteria, constraints related to virulence (protein and EPS production) were removed. The ATP maintenance is identical to the one used for *R. solanacearum*, converted in this appropriate unit ($h^{-1}$ replaced by $day^{-1}$). Glucose uptake values for *R. solanacearum* and *X. fastidiosa* were assumed to be proportional (using the number of mol of carbon) to the uptake rate for glutamine previously determined for the two organisms. The objective function consists in maximizing biomass.

The SBML file is given in the same folder.

To perform Gene Deletion, you need to place the constraints and SBML files in the same folder than the FlexFlux folder, and then use the command line:

- On Linux

```
Flexflux.sh KO -mode 1 -cons constraints.tab -out KO_genes.txt -s XF_network.xml -sol CPLEX -ext –plot
```

- On Windows

```
Flexflux KO -mode 1 -cons constraints.tab -out KO_genes.txt -s XF_network.xml -sol CPLEX -ext –plot
```

Reaction knock-out can also be performed: -mode 1 must be replaced by –mode 0, and the output file name must be replaced by KO_reactions.txt.

To test Gene Deletion Study and reaction knock-out on protein-rich environment, the constraints.tab file must be replaced by the file in the folder flexflux/KO/xfas_protein-rich.

Gene Deletion Study and reaction knock-out can also be performed on ***R. solanacearum***. The command lines are similar and the appropriate constraints are available in flexflux/KO/rsol_glc or flexflux/KO/rsol_protein-rich. For Gene Deletion Study:

- On Linux

```
Flexflux.sh KO -mode 1 -cons constraints.tab -out KO_genes.txt -s rsolGMI1000_20190128_forModelling.xml -sol CPLEX -ext –plot
```

- On Windows

```
Flexflux KO -mode 1 -cons constraints.tab -out KO_genes.txt -s rsolGMI1000_20190128_forModelling.xml -sol CPLEX -ext –plot
```

Similarly, reaction knock-out can also be performed: -mode 1 must be replaced by –mode 0, and the output file name must be replaced by KO_reactions.txt.

The simulations can also be performed on ***E. coli*** in the two environments, with the constraints and SBML files available at flexflux/KO/ecol_glc or flexflux/KO/ecol_protein-rich. For Gene Deletion Study, the command lines are:

- On Linux

```
Flexflux.sh KO -mode 1 -cons constraints.txt -out KO_genes.txt -s ecoli-orth2011.xml -sol CPLEX -ext –plot
```

- On Windows

```
Flexflux KO -mode 1 -cons constraints.txt -out KO_genes.txt -s ecoli-orth2011.xml -sol CPLEX -ext -plot
```

Similarly, reaction knock-out can also be performed: -mode 1 must be replaced by –mode 0, and the output file name must be replaced by KO_reactions.txt.

**Author**: Léo Gerlin

**Title**: Study of metabolic interactions between plant and xylem-colonizing bacteria using constraint-based modeling

**Supervisors**: Stéphane Genin, Caroline Baroukh

**Abstract**:

Several pathogenic bacteria that cause devastating diseases in plants colonize the xylem vessels. Xylem sap is often described as being poor in organic carbon, which is paradoxical since some bacteria can reach high cellular densities in this environment. Furthermore, recent studies revealed a close relationship between the regulation of virulence and metabolism in these bacteria. Thus, an integrative analysis of plant host and bacterial metabolisms should help understanding the contribution of bacterial nutrition to virulence once in the plant. To this end, we used constraint-based metabolic modeling. This methodology relies on the formulation of linear optimization problem and predicts/quantifies the fluxes of matter from substrate assimilation to biomass production. As model organisms, we used a bacterium from the *Ralstonia solanacearum* species complex (*Ralstonia pseudosolanacearum* GMI1000), the causal agent of bacterial wilt, and its natural host tomato. The first part of the project aimed at building a multi-organ metabolic model of the plant, supported by the acquisition of experimental data, and integrating the fluxes of matter of xylem. This model was shown to be a relevant tool to predict xylem composition upon environmental or physiological conditions. Physiological and metabolic dynamics of bacterial wilt on tomato were then characterized through an experimental study. It allowed quantifying the key parameters of the interaction. Notably, a tipping point towards disease happens after a strong bacterial proliferation and a drain of glutamine and asparagine in the xylem. These parameters were used to build and calibrate a metabolic model of the *R. solanacearum* – tomato interaction that predicted: i) the bacterial density threshold detrimental to plant growth, ii) the maximal bacterial density that can be reached in the plant through xylem fluxes, iii) the respective contribution of sugars and glutamine in the interaction. The model revealed that the bacterial density limiting plant growth was close to the tipping point observed experimentally, suggesting that resource exploitation was a key feature required for disease. In addition, the xylem fluxes could sustain a rapid growth of *R. solanacearum* until very high densities, indicating that xylem should not be considered as a 'poor environment' for infectious bacteria. Finally, the modeling results suggest a critical role for glutamine. Linking plant nutrition and cultivation condition to glutamine flux in xylem appears a promising perspective to restrict the effect of xylem-colonizing bacteria on crops. To have a more global overview of the role of metabolism in xylem-colonizing bacteria lifestyles, the metabolic model of another bacterium, *Xylella fastidiosa* (subsp. *multiplex* CFBP8418) was built and manually curated using literature and experimental databases. I demonstrated that *X. fastidiosa* specificities (slow growth, environmental specialization) are reflected in its metabolic capacities (reduced and weakly robust compared to *R. solanacearum*) as well as in the absence of some metabolic reactions/pathways suspected to contribute to fast growth. The reconstruction of metabolic models thus appears to be a relevant tool to better characterize the different lifestyles of bacterial plant pathogens.

**Keywords**: modeling, xylem, *Ralstonia solanacearum*, metabolism, tomato, *Xylella fastidiosa*

**Auteur** : Léo Gerlin

**Titre** : Etude des interactions métaboliques entre plante et bactéries colonisatrices du xylème en utilisant la modélisation par contraintes

**Directeurs de thèse** : Stéphane Genin, Caroline Baroukh

**Résumé** :

Parmi les bactéries pathogènes de plantes les plus préoccupantes, plusieurs colonisent les vaisseaux du xylème. Le xylème est décrit comme un environnement pauvre en sources carbonées, ce qui paraît paradoxal puisque certaines bactéries y atteignent de hautes densités cellulaires. Par ailleurs, des études récentes ont mis en lumière l'étroite relation entre la régulation de la virulence et du métabolisme chez ces bactéries. Ces constats nous poussent à adopter une analyse intégrative des métabolismes de la plante hôte et de la bactérie, pour mieux comprendre la contribution de la nutrition bactérienne *in planta* dans l'interaction avec l'hôte. Nous nous proposons d'utiliser pour cela la modélisation métabolique par contraintes. Cette méthodologie repose sur la formulation de problèmes d'optimisation linéaire et permet de prédire et de quantifier l'ensemble des flux matière empruntés entre l'assimilation d'un substrat et la production de biomasse. Nous avons utilisé comme modèles la bactérie *Ralstonia solanacearum* (formellement *Ralstonia pseudosolanacearum* GMI1000), agent causal du flétrissement bactérien, et la tomate, son hôte naturel. La première partie de ces travaux a donc consisté à construire un modèle métabolique multi-organe de la plante, soutenu par l'acquisition de données expérimentales, et intégrant les flux matière du xylème. Ce modèle s'est révélé un outil pertinent pour prédire la composition du xylème à partir des conditions environnementales et physiologiques de la plante. Ensuite, une étude expérimentale de la dynamique physiologique et métabolique du flétrissement bactérien sur tomate a permis de quantifier des paramètres clés de l'interaction. Notamment, un point de basculement vers la maladie a été caractérisé, intervenant après une forte prolifération de la bactérie et un épuisement de la glutamine et de l'asparagine dans le xylème. Ces paramètres ont été utilisés pour développer un modèle métabolique de l'interaction *R. solanacearum* – tomate calibré sur de nombreuses données expérimentales, et de prédire : i) la densité bactérienne qui sera nuisible à la croissance de la plante, ii) la densité bactérienne maximale pouvant proliférer dans la plante grâce aux flux du xylème, iii) la contribution de l'assimilation des sucres et de la glutamine dans l'interaction. Il a ainsi été révélé que la densité bactérienne limitant la croissance de la plante était proche de celle du point de basculement obtenu expérimentalement, suggérant que l'utilisation des ressources est un paramètre clé dans le basculement vers la maladie. Aussi, les flux du xylème sont en mesure de soutenir une croissance rapide de *R. solanacearum* jusqu'à des densités très élevées, soulevant donc qu'en prenant en compte des données quantitatives, le xylème n'est pas à considérer comme pauvre pour la bactérie. Enfin, un rôle particulièrement critique de la glutamine est suggéré par le modèle. Relier la nutrition de la plante et ses conditions de culture au flux de glutamine dans le xylème apparaît donc une piste prometteuse pour limiter l'effet de bactéries colonisatrices du xylème sur les cultures. Afin de parvenir à une vision plus globale du rôle du métabolisme dans le mode de vie des bactéries colonisatrices du xylème, le modèle métabolique d'une autre bactérie, *Xylella fastidiosa* (sous-espèce *multiplex* CFBP8418), a été reconstruit et manuellement curé en utilisant la littérature et des données expérimentales. Il a ainsi été montré que les particularités de *X. fastidiosa* (croissance très lente, spécialisation environnementale) se répercutaient dans ses capacités métaboliques (restreintes et peu robustes par rapport à *R. solanacearum*), et dans l'absence de plusieurs réactions ou voies suspectées de contribuer à une croissance rapide. La reconstruction de modèles métaboliques constitue donc un outil pertinent pour comprendre et expliquer les différences de styles de vie entre bactéries pathogènes de plantes.

**Mots clés** : modélisation, xylème, *Ralstonia solanacearum*, métabolisme, tomate, *Xylella fastidiosa*