



UNIL | Université de Lausanne

Unicentre

CH-1015 Lausanne

<http://serval.unil.ch>

Year : 2021

Environmental activity of the inoculant *Pseudomonas veronii* 1YdBTEX2 during bioaugmentation in natural and polluted soils

Morales-Otero Marian

Morales-Otero Marian, 2021, Environmental activity of the inoculant *Pseudomonas veronii* 1YdBTEX2 during bioaugmentation in natural and polluted soils

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive <http://serval.unil.ch>

Document URN : urn:nbn:ch:serval-BIB_4151C844762F1

Droits d'auteur

L'Université de Lausanne attire expressément l'attention des utilisateurs sur le fait que tous les documents publiés dans l'Archive SERVAL sont protégés par le droit d'auteur, conformément à la loi fédérale sur le droit d'auteur et les droits voisins (LDA). A ce titre, il est indispensable d'obtenir le consentement préalable de l'auteur et/ou de l'éditeur avant toute utilisation d'une oeuvre ou d'une partie d'une oeuvre ne relevant pas d'une utilisation à des fins personnelles au sens de la LDA (art. 19, al. 1 lettre a). A défaut, tout contrevenant s'expose aux sanctions prévues par cette loi. Nous déclinons toute responsabilité en la matière.

Copyright

The University of Lausanne expressly draws the attention of users to the fact that all documents published in the SERVAL Archive are protected by copyright in accordance with federal law on copyright and similar rights (LDA). Accordingly it is indispensable to obtain prior consent from the author and/or publisher before any use of a work or part of a work for purposes other than personal use within the meaning of LDA (art. 19, para. 1 letter a). Failure to do so will expose offenders to the sanctions laid down by this law. We accept no liability in this respect.



UNIL | Université de Lausanne

Faculté de biologie
et de médecine

Département de Microbiologie Fondamentale

**Environmental activity of the inoculant *Pseudomonas veronii*
1YdBTEX2 during bioaugmentation in natural and polluted soils**

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Marian MORALES

Master of Science en Ingénierie Environnementale de l'Université de Valle (Colombie)

Jury

Prof. Tadeusz Kawecki, Président
Prof. Jan van der Meer, Directeur de thèse
Prof. Dietmar Pieper, expert
Prof. Nico Boon, expert

Lausanne
(2021)



UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Ecole Doctorale
Doctorat ès sciences de la vie

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président·e	Monsieur	Prof.	Tadeusz	Kawecki
Directeur·trice de thèse	Monsieur	Prof.	Jan Roelof	van der Meer
Expert·e·s	Monsieur	Prof.	Dietmar	Pieper
	Monsieur	Prof.	Nico	Boon

le Conseil de Faculté autorise l'impression de la thèse de

Madame Marian Derly Morales Otero

Master in engineering, Universidad del Valle, Colombie

intitulée

**Environmental activity of the inoculant
Pseudomonas veronii 1YdBTEX2 during
bioaugmentation in natural and polluted soils**

Date de l'examen : 4 août 2021

Date d'émission de l'imprimatur : Lausanne, le 10 août 2021

pour le Doyen
de la Faculté de biologie et de médecine

Prof. Niko GELDNER
Directeur de l'Ecole Doctorale

*Beautiful is what we see,
More beautiful is what we know,
Most beautiful, by far, is what we don't.
– Nicolas Steno*

To my parents...

A mis padres...

TABLE OF CONTENTS

SUMMARY.....	iii
RÉSUMÉ.....	vi
GENERAL INTRODUCTION.....	1
CHAPTER 1.....	15
The genome of the toluene-degrading <i>Pseudomonas veronii</i> strain 1YdBTEX2 and its differential gene expression in contaminated Sand	
CHAPTER 2.....	39
Genome-wide gene expression changes of <i>Pseudomonas veronii</i> 1YdBTEX2 during invasion in polluted soils	
CHAPTER 3.....	61
Identification of genes involved in soil adaptation in <i>Pseudomonas veronii</i> 1YdBTEX2	
GENERAL DISCUSSION.....	97
CURRICULUM VITAE.....	105
Acknowledgements.....	107

SUMMARY

Bioaugmentation uses the capacities of specific bacterial strains inoculated into sites to enhance pollutant biodegradation. Bioaugmentation mainly involves introducing bacteria that deploy their metabolic properties and adaptation potential to survive and propagate in the contaminated environment using the target pollutant as a carbon, nutrient or energy source. Most of our understanding of biodegradation activity comes from experiments carried out under laboratory conditions. However, attempts to demonstrate the potential for bioaugmentation of biodegraders under *in situ* conditions have shown both success and failures, without clear understanding of the underlying reasons or mechanisms. We hypothesized that a better understanding is necessary of the processes occurring during the invasion of inoculants into complex environments, such as polluted soils. Only then can we identify the bottlenecks for bioaugmentation and fully exploit the natural diversity of biodegrader candidates. In this context, the overall aim of this work was to rationally study the soil survival capacity and contrast the metabolic and physiological strategies of the BTEX-degrading bacterium *Pseudomonas veronii* 1YdBTEX2 during adaptation and invasion to natural non-sterile and contaminated soils, with that in polluted liquid suspended cultures or standardized porous materials.

In the general introduction, I present the state of the art in soil bioremediation and bioaugmentation, and degradation of BTEX compounds in particular.

In the first research chapter, I focus on understanding the adaptive response of *P. veronii* 1YdBTEX2 during transition from liquid batch culture to contaminated soil. For this, I analyzed the short-term (1 h) changes in genome-wide gene expression in non-sterile sand compared to liquid medium, both in presence or absence of toluene. In a collaborative effort, we improved the genome sequence of *P. veronii* 1YdBTEX2, and showed that it actually covers three individual replicons with a total size of 8 Mb, with a large proportion of genes with unknown functions. One-hour exposure to toluene, both in soil and liquid, triggered massive transcription (up to 208-fold induction) of multiple gene clusters, such as toluene degradation pathway(s), chemotaxis, and toluene efflux pumps. This clearly underlines their key role in the adaptive response to toluene. In comparison to liquid medium, cells in soil drastically changed the expression of genes involved in membrane functioning (e.g., lipid composition, lipid metabolism, cell fatty acid synthesis), osmotic stress response (e.g., polyamine or trehalose synthesis, uptake of potassium), and putrescine metabolism, highlighting the adaptive mechanisms of *P. veronii* to readjust its metabolism to sand.

In the next chapter, I focused on understanding the potential changes in the *P. veronii* global transcriptome during actual growth and survival (stationary phase) in contaminated soils. In addition, I aimed to investigate differences in behaviour among three different non-

sterile soils and a historically contaminated material with (polycyclic) aromatic hydrocarbons. I showed that *P. veronii* established itself in all except one soil, at the expense of toluene. The strain also grew in its absence of toluene but to a lower population density. This indicates its capacity to survive in real field conditions under contamination stress. I compared the transcriptomic responses during transition phase, exponential growth and stationary phase among the soils and to the behaviour induced in regular liquid culture or inert silica matrix. The transcriptomic analysis revealed that *P. veronii* displayed a versatile global program, characterized by commonly shared and soil-specific strategies, and signs of nutrient limitations in later growth phases, which explain its behaviour in the soil environments. We also observed a strikingly conserved core metabolic expression in the exponential phase irrespective of the growth environment. The non-growth in one soil that carried the highest resident background microbial community may be the result of predation by higher loads of resident protists, and/or of possible competition for substrate by resident microbes. In the absence of sizeable *P. veronii* population, however, I did not manage to isolate sufficient RNA for transcriptomics, and signs of predatory or competition stress could thus not be uncovered.

Finally, in the third chapter, I focused on characterizing the genes that might be important for growth and survival of *P. veronii* in (contaminated) soils. For this I used a transposon-insertion scanning approach. With help I generated two independent genome-wide insertion libraries of *P. veronii* through conjugation of a non-replicating hyperactive mini-Tn5 delivering an *aph* gene insertion (Kanamycin resistance gene). I inoculated the libraries then either in liquid suspended culture media or in two different soils under presence of toluene. Libraries were propagated for 50 generations in the same environments. Deep sequencing of amplified boundaries between the inserted *aph*- and host genes, served to quantify the relative abundances of insertions over time and condition. Fitness importance of genes for soil growth and survival was assessed from de- or increases of the relative abundance of insertions over generation times, and compared to random insertion models. Apart from typically considered 'essential' genes that were mostly absent already in the starting libraries, our data showed only a small number of genes implicated in fitness loss, that were overlapping between both soils and absent from liquid growth. These pointed to crucial functions of urea and short-chain fatty acid metabolism, oxidative stress defense, and nutrient/element transport for soil growth. Our analysis also showed that mutants inactivating flagella biosynthesis and motility had a strong fitness gain in soils. Long-term soil growth and survival of *P. veronii* in soils is thus a function of a variety of independent genes, but not necessarily of a coherent 'soil-specific' growth program.

The results obtained in my thesis enabled to draw a global picture of the natural behavior, strategies, and capabilities of *P. veronii* to grow and survive under environmentally relevant conditions, such as posed by non-sterile and polluted soils.

By comparing transcriptomic responses in different soils and materials, together with the transposon mutagenesis scanning at different growth phases, I am confident to have covered a broad range of conditions and scenarios that help our understanding of the mechanisms necessary for strain adaptation and survival upon inoculation. The integration of the results obtained throughout the above-described chapters showed that *P. veronii* is a robust strain. Its fast adaptability is explained by the large and redundant genome, which encodes several mechanisms to maintain its main central core metabolism despite the environment it is inoculated and to adjust efficiently to the constantly changing soil conditions. Finally, this work also shows evidence of possible reasons for which *P. veronii* would not be able to invade complex microbiomes.

RÉSUMÉ

La bioaugmentation utilise les capacités de certaines bactéries inoculées dans un site contaminé pour améliorer la dégradation de polluants. Ce procédé implique l'introduction des bactéries qui peuvent exercer leurs propriétés métaboliques et d'adaptation pour survivre et prospérer dans le site contaminé en utilisant le polluant comme source de carbone, nutriment et énergie. La plupart des connaissances à propos de l'activité de biodégradation proviennent des expériences menées en conditions de laboratoire. Cependant, les tentatives pour démontrer le potentiel pour la bioaugmentation de ces bactéries sur site ont montré des réussites, mais aussi des échecs, sans vraiment que l'on comprenne les raisons ou mécanismes de ces résultats.

Nous croyons nécessaire une meilleure compréhension des processus qui ont lieu pendant l'invasion des bactéries inoculées dans des environnements complexes comme le sont les sites contaminés. De cette façon, on pourrait identifier les « problèmes » de la bioaugmentation et pouvoir ainsi exploiter au mieux la diversité naturelle des bactéries capables de dégrader des polluants. Dans ce contexte, le principal objectif de ce travail était d'étudier de façon rationnelle la capacité de survie et les stratégies métaboliques et physiologiques de la bactérie *Pseudomonas veronii* 1YdBTEX2 (qui dégradent les BTEX). On les a comparé pendant l'adaptation et l'invasion de cette bactérie dans les sols naturels non stériles et contaminés, avec celle des cultures en suspension dans des liquides pollués ou des matériaux poreux standardisés.

Dans l'introduction général, je présente l'état des connaissances en bioremédiation de sols et le concept du bioaugmentation. En particulier je décris la dégradation de composants BTEX (Benzène, Toluène, Éthylbenzène et Xylènes), qui sont des substances volatiles très toxiques.

Dans le premier chapitre de recherche, je me focalise sur la compréhension des réponses adaptatives de *P. veronii* 1YdBTEX2 pendant la transition de cultures liquides à des sols contaminés. Dans cette optique, j'ai étudié les changements à court terme (1 heure) de l'expression de gènes des bactéries inoculées dans le sable non stérile, versus des cultures liquides ; dans les deux cas, en présence ou absence de toluène. Au cours d'un travail collaboratif, nous avons amélioré la séquence génomique de *P. veronii* 1YdBTEX2 et nous avons montré qu'en réalité, il est composé de trois réplicons individuels avec une taille de 8 Mb et qui inclut une large proportion de gènes avec des fonctions inconnues. Une heure d'exposition au toluène, en culture liquide ou au sol, déclenche une transcription massive (induction jusqu'à 208 fois) de multiples groupes de gènes, comme codant par exemple les voies métaboliques de dégradation du toluène, la chimiotaxie ou les pompes d'efflux de toluène. Ces résultats montrent clairement le rôle clé d'une réponse adaptative au toluène. Comparativement au milieu liquide, les cellules dans le sol changent drastiquement l'expression des gènes impliqués dans le fonctionnement de la membrane (comme par exemple, la composition et métabolisme des

graisses et la synthèse des acides gras dans la cellule), la réponse au stress osmotique (par exemple, la synthèse de polyamines ou tréhalose, et l'absorption de potassium) et le métabolisme de la putrescine, mettant en évidence les mécanismes adaptatives de *P. veronii* pour réajuster son métabolisme au sable.

Dans le chapitre suivant je tente de comprendre les changements du transcriptome global de *P. veronii* pendant la croissance et survie à long terme (phase stationnaire) dans des sols contaminés. En outre, j'ai cherché à étudier les différences de comportement dans 3 types de sols non stériles et un « sol » historiquement contaminé avec des hydrocarbures aromatiques polycycliques. J'ai constaté que *P. veronii* arrive à s'établir dans tous, sauf un, au détriment du toluène. La souche croit aussi en absence de toluène, mais avec une densité de population moindre. Ces résultats montrent une capacité de survie dans des conditions réelles de terrain, sous le stress de la contamination. J'ai comparé les réponses du transcriptome pendant la phase de transition (court terme) ainsi que la croissance exponentielle et stationnaire dans le sol avec le comportement induit dans les cultures liquides ou dans une matrice de silice inerte. L'analyse révèle un programme global et versatile de *P. veronii*, caractérisé par des stratégies communes ou spécifiques au sol et des évidences de limitation de nutriments dans les phases de croissance tardives, ce qui explique son comportement dans les environnements du sol. Nous avons aussi observé une expression métabolique centrale remarquablement conservée dans la phase exponentielle, sans distinction de l'environnement de croissance. L'absence de croissance dans un des sols, qui contenait une communauté complexe de microbes, est peut être le résultat de prédation de la part d'une abondante présence de protistes résidants et/ou une possible compétition pour le substrat par les microbes déjà présents. En raison de la quantité insuffisante de *P. veronii* pour extraire suffisamment d'ARN pour l'analyse du transcriptome, des signes de prédation ou de stress compétitif n'ont pas pu être élucidés.

Finalement, dans le quatrième chapitre, j'ai caractérisé les gènes qui pourraient être importants pour la croissance et la survie de *P. veronii* dans les sols (contaminés). Pour ce faire, j'ai utilisé la technique des insertions massives de transposons. De cette manière, on a pu générer deux bibliothèques indépendantes d'insertions de transposons dans tout le génome de *P. veronii* à travers l'utilisation d'un transposon mini-Tn5 hyperactive non répliquatif avec un gène (*aph*) qui confère de la résistance à la kanamycine. Ces bibliothèques ont été inoculées dans un milieu de culture liquide et dans deux sols différents en présence de toluène et on les a laissées se propager pendant 50 générations dans le même environnement. Le séquençage profond des limites amplifiées entre l'insertion du gène *aph* et les gènes de la bactérie ont servi à quantifier la relative abondance des insertions en fonction du temps et des conditions de croissance. L'importance de la valeur adaptative des gènes pour la croissance et la survie au sol a été évaluée à partir de la diminution ou de l'augmentation de l'abondance relative des insertions au

cours des générations, et comparée à des modèles d'insertion aléatoire. En excluant les gènes typiquement considérés comme essentiels qui ont été majoritairement absents dans les bibliothèques de départ, nos résultats montrent seulement un petit nombre de gènes impliqués dans la perte de valeur adaptative communs aux deux sols et absents de la culture liquide. Ces résultats indiquent des fonctions cruciales du métabolisme de l'urée et des acides gras à chaîne courte, de la défense contre le stress oxydatif et du transport des nutriments/éléments pour la croissance au sol. Notre analyse a aussi montré que les mutants qui inactivent la biosynthèse des flagelles et la motilité ont un fort gain de valeur adaptative dans le sol. La croissance et survie à long terme de *P. veronii* dans le sol est plutôt en fonction d'une variété des gènes indépendants et pas nécessairement d'un programme cohérent spécifique au sol.

Les résultats obtenus dans ma thèse permettent de dessiner un tableau global du comportement naturel, des stratégies et de la capacité de *P. veronii* à croître et à survivre dans des conditions environnementales particulières telles que celles des sols non stériles et pollués. En comparant les réponses transcriptomiques dans différents sols et matériaux et le scan de mutants par transposon en différentes phases de croissance, je pense avoir couvert une ample variété de conditions et de scénarii qui aident à notre compréhension des mécanismes nécessaires à l'adaptation de la souche et à la survie après inoculation. L'intégration de l'ensemble des résultats montrent que *P. veronii* est une souche robuste, sa prompte adaptabilité est expliquée par son large et redondant génome qui code plusieurs mécanismes pour maintenir son métabolisme principal, malgré l'environnement où il est inoculé, en s'ajustant de manière efficace aux conditions changeantes du sol. Finalement, ce travail met en évidence des possibles raisons pour lesquelles *P. veronii* ne serait pas en mesure d'envahir des microbiomes complexes.

GENERAL INTRODUCTION







BTEX soil pollution

Soil is an irreplaceable natural resource that sustains life on the planet, challenged by food and energy demands of an increasing population. Healthy soil harbors a vast microbial population arranged in microbial communities characterized by the assembly of multiple species encompassing a remarkable genetic and phenotypic diversity that varies within (at least) each cm scale (Bardgett and van der Putten, 2014; Sessitsch et al., 2001). Microbial communities are essential for soil formation and fertility. In particular, microorganisms regulate all the main biogeochemical cycles on the planet and have a pivotal role in plant growth (Ward and Singh, 2004). Yet, the relation between microbial genetic diversity, community structure, and soil function is not completely understood (Borer and Or, 2021; Nannipieri et al., 2017).

Soil fertility and diversity are threatened by world-wide chemical pollution, erosion as well as land usage. In general terms, the main sources of soil contamination worldwide are petroleum products, pesticides, urban waste, chemical and biological warfare (Sales da Silva et al., 2020). These products are not only harmful for human health – even at low concentrations, but also drastically change soil physicochemical properties and microbiological characteristics affecting the balance of natural soil ecosystems (Adebambo et al., 2020; Ahmed et al., 2019; Ashraf et al., 2014; Wongbunmak et al., 2020; Yang et al., 2017). Petroleum hydrocarbon pollution in soil forms a complex mixture of aliphatic and aromatic compounds, including polycyclic aromatic Hydrocarbons (i.e., PAHs) and volatile organic compounds from the BTEX group (benzene, toluene, ethyl benzene, and the isomeric o-, m- and p-xylenes) (Sales da Silva et al., 2020).

BTEX compounds are the major aromatic hydrocarbons present in petroleum products like gasoline, heating oil, diesel and lubricant (Table 1). They are used in the manufacture of chemicals, plastic, solvents, cosmetics and pharmaceutical products (Correa et al., 2012; Dehghani et al., 2018). Their release into the environment is frequently associated with oil spills, whether by accident or due to anthropogenic activities (Cunningham et al., 2001; Lee et al., 2019; Mohan et al., 2020; Wongbunmak et al., 2020). Regardless of the source of BTEX pollution, the ecotoxicity of compounds belonging to this group requires special attention. They can persist for a long time in soil or ground water due to their polarity, long half-life cycle, strong adsorption, and recalcitrant nature (Ali et al., 2020), characteristics that make them difficult to remediate.

Table 1. BTEX characteristics

Compound	Benzene	Toluene	Ethylbenzene	<i>m</i> -Xylene	<i>o</i> -Xylene	<i>p</i> -Xylene
Chemical structure						
Molecular formula	C ₆ H ₆	C ₇ H ₈	C ₈ H ₁₀	C ₈ H ₁₀	C ₈ H ₁₀	C ₇ H ₈
Molecular weight(g mol ⁻¹)	78	92	106	106	106	106
Water solubility (mg L ⁻¹)	1700	515	152	–	175	198
MCL (mg L ⁻¹)	0.005	1	0.7	10	10	10

Adapted from (Mohammadi et al., 2020). MCL = Maximal contaminant limit (legal threshold limit on the amount of a substance that is allowed in public water systems under the Safe Drinking Water Act – SDWA).

Soil bioremediation

The functional recovery of contaminated soil is essential to sustain its future role as fertile ground for agriculture and soil life. The development of more efficient, environmentally friendly, and low-cost technologies has been and will be the subject of countless research efforts (Aqeel et al., 2014; Fuentes et al., 2014; Ritter and Scarborough, 1995). Several treatments have been proposed. They can be grouped in three categories: physical, chemical and biological methods, which can be carried out *in situ* or *ex situ*.

Physical methods encompass technologies that separate the organic chemicals and transfer the contaminants from the soil/sediment to another medium such as air or water (Li et al., 2018). The physical separation methods require a capture step in which the contaminant is concentrated for recovery, reuse, or a more cost-effective ultimate disposal. The separation can be made by injecting gas/oxygen, water with surfactants, by the generation electrochemical gradients, or thermal decontamination (Fox, 1996). Available *in situ* physical treatment technologies include electrokinetic separation (Mohan et al., 2020), fracturing (Wong and Alfaro, 2001), soil flushing (Befkadu and Chen, 2018), and soil vapor extraction (Rahbeh and Mohtar, 2007), among others. The great disadvantages of these methods include their high-energy consumption, mechanical problems in the numerous tools and equipment required in this process (e.g., blockages and clogging). This reflects in the high cost of the systems setting-up making them financially impossible, especially when large areas of soil have to be treated. Besides, one of the possible undesired byproducts generated with physical treatments is that the soil structure itself may become damaged. For instance, heating treatments have been more efficient in removing volatile organic compounds (VOC) than any other physical methods. However, the treatments' efficiency is compromised by soil characteristics such as water content and porosity, and more energy is required in heating soils with high moisture content, making it a costly endeavor. Secondly, heating soils to temperatures above 100 °C to remove the VOC

causes degradation reactions leading to the creation of undesirable byproducts of reduced volatility together with the destruction of soil minerals and organic matter, which can potentially impair the ability to restore soils and ecosystems to their original state (Vidonish et al., 2016).

Chemical treatments, on the other hand, focus on the chemical oxidation of organic contaminants to destroy or transform them into less toxic or biodegradable compounds. In this regard, advanced oxidation processes (AOPs) such as fenton oxidation, persulfate oxidation and ozonation have been investigated for their potential to treat diesel–fuel-polluted soils (Ayoub et al., 2010; Cheng et al., 2016; Li et al., 2014; Sherwood and Cassidy, 2014). Other commonly used chemical methods include leaching, cement curing or plasma oxidation (Li et al., 2017). The main drawback of these treatments is the generation of secondary pollution, which increases the already high cost but also the complexity of the remediation (Johnson and Affam, 2019; Lu et al., 2010; Rivas et al., 2009; Wang et al., 2021). Overall, both physical and chemical methods for soil remediation are neither economically sustainable nor environmentally friendly at large scale.

The remediation of polluted soil by microorganisms is an attractive alternative to the traditional ‘muck, suck and truck’ physical-chemical processes. Bioremediation has been the subject of investigation for decades (Lovley, 2003); and, therefore, is by no means a new concept. Two major approaches have been used for biological treatment. The first one, known as bioremediation, consists of stimulating native microbial communities present in contaminated sites. This can be done by adding optimum levels of essential nutrients required for microbial metabolism (Lee et al., 2019), or by introducing sufficient oxygen or other electron acceptors. The general objective is to engage and stimulate the autochthonous microbial community in breaking down the contaminants. Whereas this is a highly efficient working process for regular contaminants (e.g., petroleum oils), its efficiency is limited and restricted to the existing metabolic diversity of the native microbes to effectively use the pollutant as an energy and carbon source. In other words, in absence of appropriate xenobiotic degradation pathways in the resident community, bioremediation is unlikely to provoke to desired effect. A second approach, known as bioaugmentation, however, consists of adding nonnative microorganisms or consortia especially suited to degrading the target pollutant to enhance the pollutant biodegradation *in situ* (Szulc et al., 2014). The foreigner strain(s) is(are) intended to survive, colonize and propagate at the expense of the degraded pollutant.

Consistent with the astonishing bacterial metabolic diversity, a wide variety of bacterial strains able to partially or entirely mineralize different petroleum hydrocarbon fractions have been isolated directly from contaminated sites (Table 2). More than 100 genera and more than 200 species of petroleum-degrading bacteria have been found in aquatic and soil environments (Fuentes et al., 2014). These mainly include bacteria and molds. Those strains

are of particular interest because they have the potential to be used in the decontamination of fuel polluted environments. The environmental isolates are often well characterized in the laboratory. In this way, specialized pathways to use BTEX as their sole carbon and energy source have been described (Lima-Morales et al., 2016) and summarized in Figure 1. Additionally, the study of organic-solvent-tolerant bacteria also revealed several mechanisms to withstand toxicity. They include the *trans* isomerization of fatty acids and the increased biosynthesis of phospholipids, which provide the strains with a more rigid cell membrane less susceptible to the structural disturbances caused by organic solvents (Junker and Ramos, 1999). Another common characteristic is the presence of multiple multidrug efflux pump systems (Ramos et al., 2002; Sardesai and Bhosle, 2002) and the release of outer membrane vesicles (OMVs) from the cell surface (Ramos et al., 2002). The latter leads to a rapid increase in cell surface hydrophobicity and mediates the formation of cell aggregates and biofilms formation (Eberlein et al., 2018) thus protecting the cell from external damage caused by toxic compounds.

Table 2. Petroleum-degrading microorganisms

Classification	Hydrocarbon degrading genera
Bacteria	<i>Achromobacter, Acinetobacter, Alcaligenes, Arthrobacter, Actinomycetes, Bacillus, Brevibacterium, Brevibacillus, Cycloclasticus, Citrobacter freundii, Chromobacterium, Corynebacterium, Enterobacter cloacae, Micrococcus, Mycobacterium, Micromonospora, Nocardia, Serratia marcescens, Sphingomonas, Staphylococcus, Stenotrophomonas, Streptomyces parvus, Sphingomonas, Spirillum, Ochrobactrum, Oleispira, Oleiphilus, Pseudomonas, Rhodococcus, Vibrio</i>
Fungi	<i>Trichoderma, Penicillium, Aspergillus, Mortierella, Rhodotorula, Selenotila, Trichosporon, Cladosporium, Cunninghamella, Fusarium, Gliocladium, Mucor, Verticillium, Phanerochaete</i>
Algae	<i>Anabaena, Oscillatoria, Nostoc</i>

Adopted from (Wang et al., 2021)

To the best of our knowledge, the use of microorganisms for remediation is the method of choice for treating BTEX-contaminated soil *in situ*. Biodegradation is widely regarded as a cost-effective and environmentally friendly strategy, although, attempts to demonstrate the potential for bioaugmentation in soils have resulted in successes and failures (Ławniczak et al., 2020; Radwan et al., 2019).

One of the most challenging aspects in bioaugmentation is the survival of strains after they are introduced into soil (Mrozik and Piotrowska-Seget, 2010; Thompson et al., 2005). It has been observed that the number of exogenously added microorganisms significantly decreases shortly after inoculation (Ławniczak et al., 2020; Mrozik and Piotrowska-Seget, 2010). Thus, the success of bioaugmentation efforts does not depend solely on the metabolic capacity and the tolerance to the toxicity of the inoculated strain. The successful establishment and efficient removal of the target pollutant is influenced by abiotic factors such as pH, water content, oxygen,

nutrient availability, and soil salinity (Al Disi et al., 2017). These factors also affect the bioavailability of the target compound. Another critical aspect that may negatively impact bioaugmentation, widely discussed but little addressed in the literature, is the interaction between autochthonous and introduced microorganisms, such as predation, competition for nutrients and niches, and other antagonistic interactions with autochthonous microbial soil populations (Ławniczak et al., 2020; Villaverde et al., 2019). Nevertheless, it is of a common understanding that there is a general lack of knowledge on how environmental conditions as a whole influence the activity of introduced bacteria. Thus, it is clear that a detailed understanding of the physiological reactions and the limitations of the strains to colonize polluted environments may help to improve the success of their rational manipulation and overcome factors that limit efficient bioaugmentation.

Although we have learned key aspect of environmental isolates in laboratory conditions, more studies testing the performance of the strains in 'real' environments have to be carried out. Combined approaches (e.g., culture-depend methods, DNA based method, modeling, etc) are required to understand the bottlenecks of xenobiotic degradation (e.g., BTEX, heavy metals, PHA's) and bacteria adaptation. It is necessary to elucidate the strain constraints to better understand how bacteria survive in and respond to complex environments such as natural and polluted soils.

Use of omics-tools to understand bacterial adaptation to harmful conditions: the case of solvent tolerance and soil inoculations

Microorganisms can speed up environmental restoration through bioaugmentation. However, the promise of bioaugmentation has several aspects that limit its implementation. In general, microbial processes the strains deploy to degrade or transform contaminants in laboratory conditions might not work in the field, or bioaugmentation strategies successful in one location might not work in another. How to overcome such drawbacks is not obvious, but a more detailed understanding of the limitations of strain as well as the mechanisms controlling the growth and activity in contaminated environments is needed. Additionally, knowledge regarding microbial dynamics upon inoculation and during bioaugmentation together with the interactions with biotic and abiotic factors is a gap that has to be filled to move bioremediation applications one step forward.

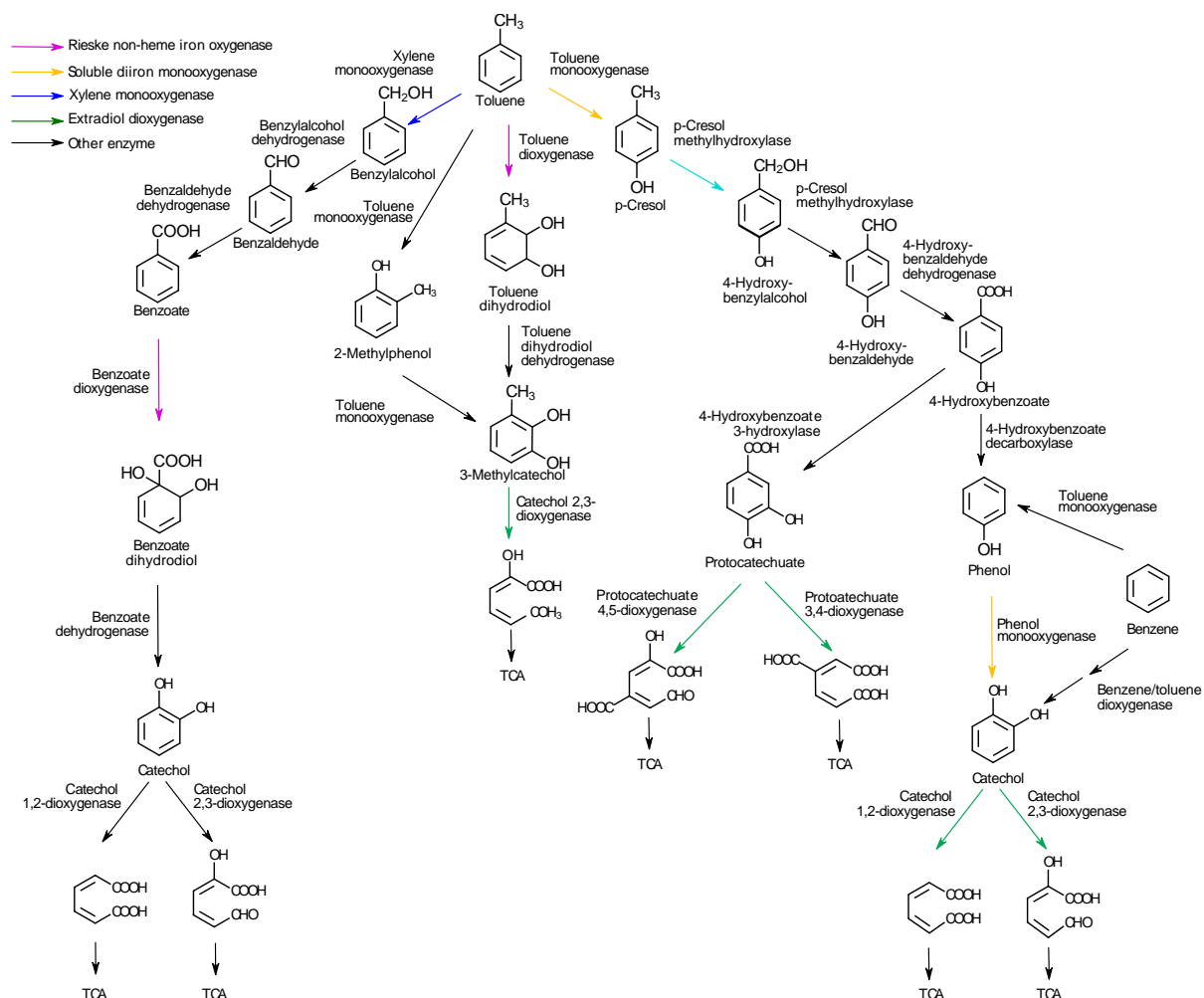


Figure 1. Major pathways for the aerobic metabolism of toluene and benzene. Activation is usually achieved by Rieske nonheme iron oxygenases typically catalyzing a dioxygenation. Rearomatization is then catalyzed by dihydrodiol dehydrogenases. Alternatively, benzene and toluene can be activated through subsequent monooxygenations catalyzed by soluble diiron monooxygenases (toluene monooxygenase and/or phenol monooxygenase) to form the respective catechols. Toluene can be additionally activated through oxidation of the methylsubstituent initiated by xylene monooxygenase or through toluene 4-monooxygenation with 4-methylphenol as an intermediate. Central di- or trihydroxylated intermediates are subject to ring cleavage by intradiol dioxygenases or extradiol dioxygenases. Ring cleavage products are channelled to the tricarboxylic acid cycle via central reactions. Adopted from (Duarte et al., 2017; Lima-Morales et al., 2016)

The isolation and characterization of organisms responsible for bioremediation using culture-dependent techniques is the first effort to survey for their metabolic capabilities, along with other physiological aspects that are likely to control growth in the polluted environments. However, before the application of molecular techniques, it was impossible to define whether potential isolates are actually key bioremediation players.

More recently, a number of DNA-based molecular techniques have been used to study microorganisms responsible for bioremediation (Lovley, 2003; Wang et al., 2009). The

tremendous advance of high-throughput sequencing technologies has improved our capacity to look into the species composition and gene content of microbial communities. They allow identifying with high precision the diversity of microbes within a community or over time, based on sequencing and comparison of 16S rRNA gene amplicons (Widder et al., 2016). Through direct sequencing of purified community DNA, the gene content of the most abundant community members can be identified and sometimes assembled in metagenomic and computational strategies. Finally, to determine whether and which genes in the community are actively being transcribed and translated into functional proteins, metatranscriptomic, proteomics, and metabolomic analyses can be carried out. With the production of transposon mutants it is possible to screen for genes involved in fitness during soil bioaugmentation. The application of omic-tools for the taxonomic and functional aspects of the microbial communities from contaminated sites has led to discovering new bacterial strains that otherwise were not accessible by using the traditional culturing techniques. However, the general molecular and functional response of cells to a given environment has been poorly explored and most studies have been limited to laboratory conditions.

One example of the use of combine approaches to better understand bacteria soil adaptation and biodegradation is *Sphingomonas wittichii* RW1, a candidate strain proposed for the targeted degradation of dioxins and DBFs in soils, which has been subjected to multiple studies. Laboratory experiments revealed how RW1 metabolized chlorinated aromatic hydrocarbons (Nam et al., 2012; Wilkes et al., 1996; Wittich et al., 1992), but it was not until the genome was sequenced and published (Miller et al., 2010) that functional studies were carried out.

A first study attempted to expose strains to defined water stress conditions, similar as are expected to appear during soil adaptation, by using polyethylene glycol 8000 (PEG8000) and sodium chloride additions to liquid media. Transcriptomic profiling of RW1 showed that the response to the generated matric and solute stress shared a small number of induced genes, including the induction of genes involved in trehalose and exopolysaccharide biosynthesis and the reduced expression of genes for flagellar assembly, despite the not apparent effect on the growth rate (Johnson et al., 2011). A complementary study showed that global transcriptomic reactions of RW1 upon actual inoculation in non-saturated sand are significantly different to those under defined salt or matric stress in liquid batch cultures even when using the same carbon substrate, although the specific growth rates were similar under both conditions (Coronado et al., 2012; Moreno-Forero and van der Meer, 2015). A genome-wide transposon scanning of RW1 identified essential genes for long-term survival and genes conferring fitness advantage in soil. The analysis suggested that genes involved in nutrient scavenging, energy production, and motility are important for short-term adaptation, whereas fatty acid

metabolism and oxidative stress are pivotal functions for long-term survival. Interestingly, this approach also allowed identifying genes whose mutations could be associated with fitness gain given the significant increase of this mutant in the meta-population (Roggo et al., 2013). These results clearly showed that strains intended for bioaugmentation respond specifically to soil conditions, which are different from responses in a liquid suspended medium. This demonstrates the importance of conducting studies in conditions as close to reality as possible. This work also strongly suggested that soil-specific responses are not as easily distinguished by simple comparison to liquid culture controls, but may need specific controls for soil matrix effects. Interestingly, strain RW1 expressed several genes for motility consistently lower and genes for polysaccharide biosynthesis consistently higher in sand, matrix and solute stress (Moreno-Forero and van der Meer, 2015).

Studying stress responses of isolates upon inoculation in hostile environments (like polluted soils) is necessary for understanding their adaptation, which may help to further optimize biotechnological efforts, such as bioaugmentation. In this respect, several studies measured the transcriptomic response of the toluene-degrading *Pseudomonas putida* KT2440 to osmotic and oxidative stress or metal toxicity in laboratory conditions, similar to what can be found in soil environments. These studies showed upregulation of genes encoding for efflux systems, universal stress proteins, cell envelope homeostasis, and redox enzymes (Bojanovič et al., 2017; Peng et al., 2018; Svenningsen et al., 2015), common traits to cope with particular stressors.

Transcriptomic studies also allowed a better understanding of the tradeoff between the metabolic and stress-response programs induced by *P. putida* when using toluene, a membrane-damaging toxic compound, as a sole carbon source. Notably, after exposure to toluene a strong and specific metabolic response at the level of aromatic pathways was concomitant with a general stress response at several levels, reflected by the induction of genes known to respond to membrane damage, oxidative stress, and misfolding of soluble proteins (Domínguez-Cuevas et al., 2006; Molina-Santiago et al., 2017). Similar to *S. wittichii*, KT2440 avoided energy-intensive processes such as flagellar biosynthesis under short-term toluene exposure. In contrast, another study using transposon mutagenesis showed that *P. putida* DOT-T1E mutants with insertions in the flagellar export apparatus (e.g., *flhB* and *fliP*) were not only deprived of motility but also were hypersensitive to toluene (Segura et al., 2001), suggesting a connection between the FliP and FlhB proteins and toluene tolerance and the need of an intact flagellar export system for toluene tolerance.

Literature shows a wide-range of studies looking at the response to toxic compounds by tolerant bacteria. What is missing are studies on the performance of strains for bioaugmentation under realistic conditions. My project contributes to this by investigating how bacteria survive

and adapt to natural and polluted soils in the presence of a toxic compound (i.e., toluene) during the different growth phases under near-environmental conditions compared with laboratory controls.

Pseudomonas veronii as a model

Pseudomonas veronii strain 1YdBTEX2 is a Gram-negative, motile bacteria isolated from a jet-fuel contaminated soil (Hradcany, Czech Republic) (Junca and Pieper, 2004). The strain is a candidate for bioaugmentation due to its capacity of using benzene, toluene, ethylbenzene, m- and p-xylene (BTEX) as sole carbon and energy sources (de Lima-Morales et al., 2015). The strain was an abundant member of the soil bacterial community in the contaminated site suggesting that it is well suited to survive in heavily polluted soil. These characteristics made 1YdBTEX2 a good model for the study of I) toluene catabolic pathway regulation under controlled and soil conditions and II) the adaptive mechanisms to survive in (contaminated) soils in the presence of monoaromatic solvents.

Project aims and thesis layout

Our capacity to link microbial composition and functioning is still very limited. We believe that we can obtain better insights about the processes driving emerging functionalities and self-organization of complex microbial communities by studying and characterizing the behaviour of an introduced strain into a (non-sterile, polluted) environment.

Thus, our central hypothesis here was that inoculant survival, growth and activity are dependent on biological reactions and interactions between the inoculant and the stress response associated to the toxic compound and the inoculant and the resident community; and within the physico-chemical setting of its (soil) growth environment. If we better understand those reactions and how they determine the inoculant survival, we may become better at predicting the success of the inoculants to degrade the intended compound(s) in the system. For this reason, we propose to first describe and characterize the reactions of inoculants in a real complex system.

My thesis project strived for a complete characterization and interpretation of all the essential functions enabling the survival, growth and activity of strain 1YdBTEX2 in soil polluted with aromatic compounds. We attempted to identify critical factors of its activity in soil microcosms to improve the conditions for a better establishment and propagation of the strain in real scenarios. We followed the behaviour of *P. veronii* upon inoculation in microcosms consisting of both natural soils artificially polluted with toluene and real polluted materials. We used, mainly, genomic, metatranscriptomic and transposon mutagenesis approaches to study

the metabolic and physiological adaptive responses across different physiological states of the strain during the bioaugmentation process.

The specific aims of the present project were thus the following:

- 1) Characterize the reaction of *P. veronii* 1YdBTEX2 upon inoculation in a soil contaminated or not with toluene, by measuring its genome-wide transcriptome differences.
- 2) Study the mechanisms of *P. veronii* to adapt to different (polluted) soil environments with increasing biological complexity.
- 3) Identify genes that are essential or give fitness advantage to survival in soils with increasing biological complexity polluted or not with toluene.
- 4) Understand the 'environmental survival' capacity of *P. veronii* 1YDBTEX2 in soils polluted with toxic compounds by measuring its genome-wide transcriptome differences across its growth phases.

In the **first chapter**, we characterized the 1YdBTEX2 genome and described its gene expression changes during adaptation to a soil environment. We focused here on the immediate transcriptional response of strain 1YdBTEX2 to sand polluted with toluene in comparison with liquid control. In the **second chapter**, we studied the specific and different cellular reactions of *P. veronii* during adaptation, growth and survival in a variety of natural clean and contaminated non-sterile soils, in order to better understand the strain's metabolic adaptation from its global transcriptome. We studied growth of the strain after inoculation in different soils with increasing biological complexity, and deduced reactions from differentially expressed genes or gene functions. From this we assessed their importance for adaptive response to different environments. In the **third chapter**, we performed a genome-wide transposon insertion scanning approach to identify the contribution of individual genes and operons to the survival and fitness of the strain in two artificially contaminated soils in comparison to liquid culture.

References

1. Adebambo, T.H., Fox, D.T., and Otitolaju, A.A. (2020). Toxicological study and genetic basis of BTEX susceptibility in *Drosophila melanogaster*. *Front. Genet.* *11*, 1275.
2. Ahmed, N., Ok, Y.S., Jeon, B.-H., Kim, J.R., Chae, K.-J., and Oh, S.-E. (2019). Assessment of benzene, toluene, ethyl-benzene, and xylene (BTEX) toxicity in soil using sulfur-oxidizing bacterial (SOB) bioassay. *Chemosphere* *220*, 651–657.
3. Ali, N., Dashti, N., Khanafer, M., Al-Awadhi, H., and Radwan, S. (2020). Bioremediation of soils saturated with spilled crude oil. *Sci. Rep.* *10*, 1116.
4. Aqeel, M., Jamil, M., and Yusoff, I. (2014). Soil Contamination, risk assessment and remediation. In *Environmental Risk Assessment of Soil Contamination*, (InTech), M.J. Maah, ed. (Rijeka: IntechOpen), pp. 1–56.
5. Ashraf, M.A., Maah, M.J., and Yusoff, I. (2014). Soil Contamination, risk assessment and remediation. In *Environmental Risk Assessment of Soil Contamination*, (London, UK: IntechOpen), pp. 1–56.
6. Ayoub, K., van Hullebusch, E.D., Cassir, M., and Bermond, A. (2010). Application of advanced oxidation processes for TNT removal: A review. *J. Hazard. Mater.* *178*, 10–28.
7. Bardgett, R.D., and van der Putten, W.H. (2014). Belowground biodiversity and ecosystem functioning. *Nature* *515*, 505–511.
8. Befkadu, A.A., and Chen, Q. (2018). Surfactant-enhanced soil washing for removal of petroleum hydrocarbons from contaminated soils: a review. *Pedosphere* *28*, 383–410.
9. Bojanovič, K., D'Arrigo, I., and Long, K.S. (2017). Global transcriptional responses to osmotic, oxidative, and imipenem stress conditions in *Pseudomonas putida*. *Appl. Environ. Microbiol.* *83*, e03236-16.
10. Borer, B., and Or, D. (2021). Spatiotemporal metabolic modeling of bacterial life in complex habitats. *Curr. Opin. Biotechnol.* *67*, 65–71.
11. Cheng, M., Zeng, G., Huang, D., Lai, C., Xu, P., Zhang, C., and Liu, Y. (2016). Hydroxyl radicals based advanced oxidation processes (AOPs) for remediation of soils contaminated with organic compounds: A review. *Chem. Eng. J.* *284*, 582–598.
12. Coronado, E., Roggo, C., Johnson, D.R., and van der Meer, J.R. (2012). Genome-wide analysis of salicylate and dibenzofuran metabolism in *Shingomonas wittichii* RW1. *Front. Microbiol.* *3*, 1–13.
13. Correa, S.M., Arbilla, G., Marques, M.R.C., and Oliveira, K.M.P.G. (2012). The impact of BTEX emissions from gas stations into the atmosphere. *Atmos. Pollut. Res.* *3*, 163–169.
14. Cunningham, J.A., Rahme, H., Hopkins, G.D., Lebron, C., and Reinhard, M. (2001). Enhanced in situ bioremediation of BTEX-contaminated groundwater by combined injection of nitrate and sulfate. *Environ. Sci. Technol.* *35*, 1663–1670.
15. Dehghani, M., Fazlzadeh, M., Sorooshian, A., Tabatabaee, H.R., Miri, M., Baghani, A.N., Delikhoon, M., Mahvi, A.H., and Rashidi, M. (2018). Characteristics and health effects of BTEX in a hot spot for urban pollution. *Ecotoxicol. Environ. Saf.* *155*, 133–143.
16. Domínguez-Cuevas, P., González-Pastor, J.-E., Marqués, S., Ramos, J.-L., and de Lorenzo, V. (2006). Transcriptional tradeoff between metabolic and stress-response programs in *Pseudomonas putida* KT2440 cells exposed to toluene. *J. Biol. Chem.* *281*, 11981–11991.
17. Duarte, M., Nielsen, A., Camarinha-Silva, A., Vilchez-Vargas, R., Bruls, T., Wos-Oxley, M.L., Jauregui, R., and Pieper, D.H. (2017). Functional soil metagenomics: elucidation of polycyclic aromatic hydrocarbon degradation potential following 12 years of in situ bioremediation. *Environ. Microbiol.* *19*, 2992–3011.
18. Eberlein, C., Baumgarten, T., Starke, S., and Heipieper, H.J. (2018). Immediate response mechanisms of Gram-negative solvent-tolerant bacteria to cope with environmental stress: cis-trans isomerization of unsaturated fatty acids and outer membrane vesicle secretion. *Appl. Microbiol. Biotechnol.* *102*, 2583–2593.
19. Fox, R.D. (1996). Physical/Chemical Treatment of Organically Contaminated Soils and Sediments. *J. Air Waste Manage. Assoc.* *46*, 391–413.
20. Fuentes, S., Méndez, V., Aguila, P., and Seeger, M. (2014). Bioremediation of petroleum hydrocarbons: catabolic genes, microbial communities, and applications. *Appl. Microbiol. Biotechnol.* *98*, 4781–4794.
21. Johnson, O.A., and Affam, A.C. (2019). Petroleum sludge treatment and disposal: A review. *Environ. Eng. Res.* *24*, 191–201.
22. Johnson, D.R., Coronado, E., Moreno-Forero, S.K., Heipieper, H.J., and van der Meer, J.R. (2011). Transcriptome and membrane fatty acid analyses reveal different strategies for responding to

- permeating and non-permeating solutes in the bacterium *Sphingomonas wittichii*. *BMC Microbiol.* *11*, 250.
23. Junker, F., and Ramos, J.L. (1999). Involvement of the cis/trans isomerase Cti in solvent resistance of *Pseudomonas putida* DOT-T1E. *J. Bacteriol.* *181*, 5693–5700.
 24. Ławniczak, Ł., Woźniak-Karczewska, M., Loibner, A.P., Heipieper, H.J., and Chrzanowski, Ł. (2020). Microbial degradation of hydrocarbons-basic principles for bioremediation: A review. *Molecules* *25*, 856.
 25. Lee, Y., Lee, Y., and Jeon, C.O. (2019). Biodegradation of naphthalene, BTEX, and aliphatic hydrocarbons by *Paraburkholderia aromaticivorans* BN5 isolated from petroleum-contaminated soil. *Sci. Rep.* *9*, 860.
 26. Li, D.-C., Xu, W.-F., Mu, Y., Yu, H.-Q., Jiang, H., and Crittenden, J.C. (2018). Remediation of petroleum-contaminated soil and simultaneous recovery of oil by fast pyrolysis. *Environ. Sci. Technol.* *52*, 5330–5338.
 27. Li, R., Liu, Y., Mu, R., Cheng, W., and Ognier, S. (2017). Evaluation of pulsed corona discharge plasma for the treatment of petroleum-contaminated soil. *Environ. Sci. Pollut. Res.* *24*, 1450–1458.
 28. Li, X., Cao, X., Wu, G., Temple, T., Coulon, F., and Sui, H. (2014). Ozonation of diesel-fuel contaminated sand and the implications for remediation end-points. *Chemosphere* *109*, 71–76.
 29. Lima-Morales, D., Jáuregui, R., Camarinha-Silva, A., Geffers, R., Pieper, D.H., and Vilchez-Vargas, R. (2016). Linking microbial community and catabolic gene structures during the adaptation of three contaminated soils under continuous long-term pollutant stress. *Appl. Environ. Microbiol.* *82*, 2227–2237.
 30. Lovley, D.R. (2003). Cleaning up with genomics: applying molecular biology to bioremediation. *Nat. Rev. Microbiol.* *1*, 35–44.
 31. Lu, M., Zhang, Z., Qiao, W., Guan, Y., Xiao, M., and Peng, C. (2010). Removal of residual contaminants in petroleum-contaminated soil by Fenton-like oxidation. *J. Hazard. Mater.* *179*, 604–611.
 32. Miller, T.R., Delcher, A.L., Salzberg, S.L., Saunders, E., Detter, J.C., and Halden, R.U. (2010). Genome sequence of the dioxin-mineralizing bacterium *Sphingomonas wittichii* RW1. *J. Bacteriol.* *192*, 6101 LP – 6102.
 33. Mohammadi, L., Rahdar, A., Bazrafshan, E., Dahmardeh, H., Susan, M.A., and Kyzas, G.Z. (2020). Petroleum hydrocarbon removal from wastewaters: a review. *Processes* *8*.
 34. Mohan, H., Lim, J.-M., Cho, M., Park, Y.-J., Seralathan, K.-K., and Oh, B.-T. (2020). Remediation of BTEX and Cr(VI) contamination in soil using bioelectrochemical system—an eco-friendly approach. *Environ. Sci. Pollut. Res.* *27*, 837–845.
 35. Molina-Santiago, C., Udaondo, Z., Gómez-Lozano, M., Molin, S., and Ramos, J.-L. (2017). Global transcriptional response of solvent-sensitive and solvent-tolerant *Pseudomonas putida* strains exposed to toluene. *Environ. Microbiol.* *19*, 645–658.
 36. Moreno-Forero, S.K., and van der Meer, J.R. (2015). Genome-wide analysis of *Sphingomonas wittichii* RW1 behaviour during inoculation and growth in contaminated sand. *ISME J.* *9*, 150–165.
 37. Mroziak, A., and Piotrowska-Seget, Z. (2010). Bioaugmentation as a strategy for cleaning up of soils contaminated with aromatic compounds. *Microbiol. Res.* *165*, 363–375.
 38. Nam, I.-H., Kim, Y.-M., Murugesan, K., and Chang, Y.-S. (2012). A Catabolic activity of *Sphingomonas wittichii* RW1 in the biotransformation of carbazole. *Water, Air, Soil Pollut.* *223*, 943–949.
 39. Nannipieri, P., Ascher, J., Ceccherini, M.T., Landi, L., Pietramellara, G., and Renella, G. (2017). Microbial diversity and soil functions. *Eur. J. Soil Sci.* *68*, 12–26.
 40. Peng, J., Miao, L., Chen, X., and Liu, P. (2018). Comparative transcriptome analysis of *Pseudomonas putida* KT2440 revealed its response mechanisms to elevated levels of zinc stress. *Front. Microbiol.* *9*, 1669.
 41. Radwan, S.S., Al-Mailem, D.M., and Kansour, M.K. (2019). Bioaugmentation failed to enhance oil bioremediation in three soil samples from three different continents. *Sci. Rep.* *9*, 19508.
 42. Rahbeh, M.E., and Mohtar, R.H. (2007). Application of multiphase transport models to field remediation by air sparging and soil vapor extraction. *J. Hazard. Mater.* *143*, 156–170.
 43. Ramos, J.L., Duque, E., Gallegos, M.-T., Godoy, P., Ramos-Gonzalez, M.I., Rojas, A., Teran, W., and Segura, A. (2002). Mechanisms of solvent tolerance in gram-negative bacteria. *Annu. Rev. Microbiol.* *56*, 743–768.
 44. Ritter, W.F., and Scarborough, R.W. (1995). A review of bioremediation of contaminated soils and groundwater. *J. Environ. Sci. Heal. Part A Environ. Sci. Eng. Toxicol.* *30*, 333–357.
 45. Rivas, J., Gimeno, O., de la Calle, R.G., and Beltrán, F.J. (2009). Ozone treatment of PAH

- contaminated soils: operating variables effect. *J. Hazard. Mater.* *169*, 509–515.
46. Roggo, C., Coronado, E., Moreno-Forero, S.K., Harshman, K., Weber, J., and van der Meer, J.R. (2013). Genome-wide transposon insertion scanning of environmental survival functions in the polycyclic aromatic hydrocarbon degrading bacterium *Sphingomonas wittichii* RW1. *Environ. Microbiol.* *15*, 2681–2695.
 47. Sales da Silva, I.G., Gomes de Almeida, F.C., Padilha da Rocha e Silva, N.M., Casazza, A.A., Converti, A., and Asfora Sarubbo, L. (2020). Soil bioremediation: overview of technologies and trends. *Energies* *13*.
 48. Sardessai, Y., and Bhosle, S. (2002). Tolerance of bacteria to organic solvents. *Res. Microbiol.* *153*, 263–268.
 49. Segura, A., Duque, E., Hurtado, A., and Ramos, J.L. (2001). Mutations in genes involved in the flagellar export apparatus of the solvent-tolerant *Pseudomonas putida* DOT-T1E strain impair motility and lead to hypersensitivity to toluene shocks. *J. Bacteriol.* *183*, 4127 LP – 4133.
 50. Sessitsch, A., Weilharter, A., Gerzabek, M.H., Kirchmann, H., and Kandeler, E. (2001). Microbial population structures in soil particle size fractions of a long-term fertilizer field experiment. *Appl. Environ. Microbiol.* *67*, 4215–4224.
 51. Sherwood, M.K., and Cassidy, D.P. (2014). Modified Fenton oxidation of diesel fuel in arctic soils rich in organic matter and iron. *Chemosphere* *113*, 56–61.
 52. Svenningsen, N.B., Pérez-Pantoja, D., Nickel, P.I., Nicolaisen, M.H., de Lorenzo, V., and Nybroe, O. (2015). *Pseudomonas putida* mt-2 tolerates reactive oxygen species generated during matrix stress by inducing a major oxidative defense response. *BMC Microbiol.* *15*, 202.
 53. Szulc, A., Ambrożewicz, D., Sydow, M., Ławniczak, Ł., Piotrowska-Cyplik, A., Marecik, R., and Chrzanowski, Ł. (2014). The influence of bioaugmentation and biosurfactant addition on bioremediation efficiency of diesel-oil contaminated soil: Feasibility during field studies. *J. Environ. Manage.* *132*, 121–128.
 54. Thompson, I.P., Van Der Gast, C.J., Ciric, L., and Singer, A.C. (2005). Bioaugmentation for bioremediation: the challenge of strain selection. *Environ. Microbiol.* *7*, 909–915.
 55. Vidonish, J.E., Zygourakis, K., Masiello, C.A., Sabadell, G., and Alvarez, P.J.J. (2016). Thermal treatment of hydrocarbon-impacted soils: a review of technology innovation for sustainable remediation. *Engineering* *2*, 426–437.
 56. Villaverde, J., Láiz, L., Lara-Moreno, A., González-Pimentel, J.L., and Morillo, E. (2019). Bioaugmentation of PAH-contaminated soils with novel specific degrader strains isolated from a contaminated industrial site. Effect of hydroxypropyl- β -cyclodextrin as PAH bioavailability enhancer. *Front. Microbiol.* *10*, 2588.
 57. Wang, S., Wang, D., Yu, Z., Dong, X., Liu, S., Cui, H., and Sun, B. (2021). Advances in research on petroleum biodegradability in soil. *Environ. Sci. Process. Impacts* *23*, 9–27.
 58. Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* *10*, 57–63.
 59. Ward, O.P., and Singh, A. (2004). Soil bioremediation and phytoremediation — an overview. In *Applied Bioremediation and Phytoremediation*. Soil Biology, A. Singh, and O.P. Ward, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 1–12.
 60. Widder, S., Allen, R.J., Pfeiffer, T., Curtis, T.P., Wiuf, C., Sloan, W.T., Cordero, O.X., Brown, S.P., Momeni, B., Shou, W., et al. (2016). Challenges in microbial ecology: building predictive understanding of community function and dynamics. *ISME J.* *10*, 2557–2568.
 61. Wilkes, H., Wittich, R., Timmis, K.N., Fortnagel, P., and Francke, W. (1996). Degradation of chlorinated dibenzofurans and dibenzo-*p*-dioxins by *Sphingomonas sp.* strain RW1. *Appl. Environ. Microbiol.* *62*, 367–371.
 62. Wittich, R.M., Wilkes, H., Sinnwell, V., Francke, W., and Fortnagel, P. (1992). Metabolism of dibenzo-*p*-dioxin by *Sphingomonas sp.* strain RW1. *Appl. Environ. Microbiol.* *58*, 1005–1010.
 63. Wong, R.C.K., and Alfaro, M.C. (2001). Fracturing in low-permeability soils for remediation of contaminated ground. *Can. Geotech. J.* *38*, 316–327.
 64. Wongbunmak, A., Khiawjan, S., Suphantharika, M., and Pongtharangkul, T. (2020). BTEX biodegradation by *Bacillus amyloliquefaciens* subsp. *plantarum* W1 and its proposed BTEX biodegradation pathways. *Sci. Rep.* *10*, 17408.
 65. Yang, Y., Li, J., Xi, B., Wang, Y., Tang, J., Wang, Y., and Zhao, C. (2017). Modeling BTEX migration with soil vapor extraction remediation under low-temperature conditions. *J. Environ. Manage.* *203*, 114–122.

CHAPTER 1

The genome of the toluene-degrading *Pseudomonas veronii* strain 1YdBTEX2 and its differential gene expression in contaminated Sand

Marian Morales¹, Vladimir Sentschilo¹, Claire Bertelli^{2,7}, Andrea Komljenovic^{3,7}, Nadezda Kryuchkova-Mostacci^{3,7}, Audrey Bourdilloud⁴, Burkhard Linke⁵, Alexander Goesmann⁵, Keith Harshman⁶, Francisca Segers⁴, Fabien Delapierre⁴, Damien Fiorucci⁴, Mathieu Seppey⁴, Evgeniya Trofimenko⁴, Pauline Berra⁴, Athimed El Taher⁴, Chloé Loiseau⁴, Dejan Roggero⁴, Madeleine Sulfiotti⁴, Angela Etienne⁴, Gustavo Ruiz Buendia⁴, Loïc Pillard⁴, Angelique Escoriza⁴, Roxane Moritz⁴, Cedric Schneider⁴, Esteban Alfonso⁴, Fatma Ben Jeddou⁴, Oliver Selmoni⁴, Gregory Resch¹, Gilbert Greub², Olivier Emery¹, Manupriyam Dubey¹, Trestan Pillonel², Marc Robinson-Rechavi^{3,7}, Jan Roelof van der Meer^{1*}

Departments of ¹Fundamental Microbiology, ³Ecology and Evolution, ⁴Master in Molecular Life Sciences, ⁶Lausanne Genomic Technologies Facility, Center for Integrative Genomics, University of Lausanne, 1015 Lausanne, Switzerland

2) Institute of Microbiology, University Hospital Center and University of Lausanne, 1011 Lausanne, Switzerland.

5) Bioinformatics and Systems Biology, Justus-Liebig-University, 35392 Gießen, Germany.

7) SIB Swiss Institute for Bioinformatics, 1015 Lausanne, Switzerland.

Previously published in PLOS ONE 11(11):e0165850. <https://doi.org/10.1371/journal.pone.0165850>

Author Contributions

1. **Conceptualization:** MM VS CB MRR JvdM.
2. **Data curation:** MM CB VS JvdM.
3. **Formal analysis:** MM VS CB AK NKM AB KH FS FD DF M. Seppey ET PB AET CL DR M. Sulfiotti A. Etienne GRB LP A. Escoriza RM CS EA FBJ OS GR GG OE MD TP MRR JvdM.
4. **Investigation:** MM VS CB AK NKM AB KH FS FD DF M. Seppey ET PB AET CL DR M. Sulfiotti A. Etienne GRB LP A. Escoriza RM CS EA FBJ OS GR GG OE MD TP MRR JvdM.
5. **Methodology:** MM AK VS MD MRR JvdM.
6. **Writing – original draft:** MM JvdM.
7. **Writing – review & editing:** MM CB VS AK FS GRB RM MRR JvdM

RESEARCH ARTICLE

The Genome of the Toluene-Degrading *Pseudomonas veronii* Strain 1YdBTEX2 and Its Differential Gene Expression in Contaminated Sand

Marian Morales¹, Vladimir Sentchilo¹, Claire Bertelli^{2,7}, Andrea Komljenovic^{3,7}, Nadezda Kryuchkova-Mostacci^{3,7}, Audrey Bourdilloud⁴, Burkhard Linke⁵, Alexander Goesmann⁵, Keith Harshman⁶, Francisca Segers⁴, Fabien Delapierre⁴, Damien Fiorucci⁴, Mathieu Seppey⁴, Evgeniya Trofimenko⁴, Pauline Berra⁴, Athimed El Taher⁴, Chloé Loiseau⁴, Dejan Roggero⁴, Madeleine Sulfiotti⁴, Angela Etienne⁴, Gustavo Ruiz Buendia⁴, Loïc Pillard⁴, Angélique Escoriza⁴, Roxane Moritz⁴, Cedric Schneider⁴, Esteban Alfonso⁴, Fatma Ben Jeddou⁴, Oliver Selmoni⁴, Gregory Resch¹, Gilbert Greub², Olivier Emery¹, Manupriyam Dubey¹, Trestan Pillonel², Marc Robinson-Rechavi^{3,7}, Jan Roelof van der Meer^{1*}



OPEN ACCESS

Citation: Morales M, Sentchilo V, Bertelli C, Komljenovic A, Kryuchkova-Mostacci N, Bourdilloud A, et al. (2016) The Genome of the Toluene-Degrading *Pseudomonas veronii* Strain 1YdBTEX2 and Its Differential Gene Expression in Contaminated Sand. PLoS ONE 11(11): e0165850. doi:10.1371/journal.pone.0165850

Editor: Pankaj Kumar Arora, MJP Rohilkhand University, INDIA

Received: September 7, 2016

Accepted: October 18, 2016

Published: November 3, 2016

Copyright: © 2016 Morales et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files. The curated genome of *P. veronii* strain 1YdBTEX2 was deposited in the European Nucleotide Archive under bioproject number PRJEB11417. The RNA-seq reads were deposited in the NCBI Short Read Archive under accession number SRP077862.

Funding: The study was funded by SystemsX.ch (<<http://SystemsX.ch>>), the Swiss Initiative in Systems Biology (<<http://www.systemsx.ch>>), to

1 Department of Fundamental Microbiology, University of Lausanne, Lausanne, Switzerland, 2 Institute of Microbiology, University Hospital Center and University of Lausanne, Lausanne, Switzerland, 3 Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland, 4 Master in Molecular Life Sciences, University of Lausanne, Lausanne, Switzerland, 5 Bioinformatics and Systems Biology, Justus-Liebig-University, Gießen, Germany, 6 Lausanne Genomic Technologies Facility, Center for Integrative Genomics, University of Lausanne, 1015 Lausanne, Switzerland, 7 SIB Swiss Institute for Bioinformatics, Lausanne, Switzerland

* janroelof.vandermeer@unil.ch

Abstract

The natural restoration of soils polluted by aromatic hydrocarbons such as benzene, toluene, ethylbenzene and *m*- and *p*-xylene (BTEX) may be accelerated by inoculation of specific biodegraders (bioaugmentation). Bioaugmentation mainly involves introducing bacteria that deploy their metabolic properties and adaptation potential to survive and propagate in the contaminated environment by degrading the pollutant. In order to better understand the adaptive response of cells during a transition to contaminated material, we analyzed here the genome and short-term (1 h) changes in genome-wide gene expression of the BTEX-degrading bacterium *Pseudomonas veronii* 1YdBTEX2 in non-sterile soil and liquid medium, both in presence or absence of toluene. We obtained a gapless genome sequence of *P. veronii* 1YdBTEX2 covering three individual replicons with a total size of 8 Mb, two of which are largely unrelated to current known bacterial replicons. One-hour exposure to toluene, both in soil and liquid, triggered massive transcription (up to 208-fold induction) of multiple gene clusters, such as toluene degradation pathway(s), chemotaxis and toluene efflux pumps. This clearly underlines their key role in the adaptive response to toluene. In comparison to liquid medium, cells in soil drastically changed expression of genes involved in membrane functioning (e.g., lipid composition, lipid metabolism, cell fatty acid synthesis), osmotic stress response (e.g., polyamine or trehalose synthesis, uptake of

Marian Morales, Manupriyam Dubey and Jan R. van der Meer. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. There was no additional external funding received for this study.

Competing Interests: The authors have declared that no competing interests exist.

potassium) and putrescine metabolism, highlighting the immediate response mechanisms of *P. veronii* 1YdBTEX2 for successful establishment in polluted soil.

Introduction

Contaminated soils pose an environmental and public health challenge for human societies. Pollution with crude oil or petrochemicals, such as jet fuels, solvents or polycyclic aromatic hydrocarbons is frequent and can lead to increased human exposure to carcinogens [1–3]. Pollution also leads to severe decrease in species diversity and loss of ecosystem functions [4–6]. Bacteria able to use aromatic compounds as sources of carbon and energy could potentially eliminate harmful organic contaminants and help reclaim land. Stimulation of *in-situ* community growth by landfarming or other bioremediation techniques can be an effective measure to speed up the degradation of organic contaminants in soil [5, 7, 8]. Alternatively, inoculation of specifically enriched communities or purified strains to contaminated sites has been proposed to enhance biodegradation rates (i.e., bioaugmentation). Bioaugmentation, however, is not common practice, because of inconsistent achievements [5, 9]. Accordingly, we and others have proposed that the lack of success of inoculation efforts may be due to our poor understanding of the factors determining invasion and establishment of non-native bacteria into existing communities [10–12] and in the real environment [8, 9, 13]. Therefore, in this study we determined the genome-wide gene expression changes in a bacterial strain relevant for bioaugmentation during a transition from typical laboratory cultivation to nonsterile contaminated sand. We chose to study the transition phase upon inoculation into contaminated sand, because, based on previous experience with other strains [11, 14, 15], we expected that the inoculant would display the highest expression changes and that we would see induction of the key genes involved in sand adaptation and contaminant degradation. It is clear, however, that cellular reactions during the transition phase are not necessarily equivalent to those during an actual bioaugmentation process.

The bacterial inoculant used for this study is *Pseudomonas veronii* 1YdBTEX2, a strain isolated from a jet-fuel contaminated soil (Hradcany, Czech Republic) [16] and capable of using benzene, toluene, ethylbenzene, *m*- and *p*-xylene (BTEX) as sole carbon and energy sources. Since the strain was an abundant member of the bacterial community in the contaminated site, we expected that it would be well-suited to survive toxicity of monoaromatic solvents, which can potentially be detrimental to bacterial cells [17] and adapt to a variety of contaminated soils. A draft genome sequence of *P. veronii* 1YdBTEX2 has been recently published in 63 scaffolds, a largely fragmented state [18], which makes genome-wide expression studies more difficult to interpret. Thus, we first obtained a full and gapless genome sequence of strain 1XdBTEX2 using long read PacBio SMRT sequencing. The genome was re-annotated using automated pipelines and manual curation. Subsequently, genome-wide expression patterns were studied after 1-hour exposure of exponentially growing cells of strain 1YdBTEX2 to either toluene or succinate, either in liquid suspension or in nonsterile sandy soil. Gene expression was quantified by RNA-sequencing using Illumina technology and mapping of reads onto the curated genome. PacBio sequencing, manual curation, gene expression experiments and preliminary RNA-sequencing analysis were performed within the context of a Master class [19]. Our study provides valuable insights into the cellular functions involved in the early adaptation response to organic solvents in sandy soil.

Materials and Methods

Isolation of DNA, sequencing and assembly of the *P. veronii* genome

P. veronii strain 1YdBTEX2 was cultured at 30°C with 180 rpm rotary shaking on 21C minimal medium (MM) [20] with 10 mM succinate. Cells were harvested in exponential phase at a culture turbidity of 0.15 at 600 nm by centrifugation of 10 ml culture for 6 min at 5000 x g, washed once with MM by 30 s vortexing and centrifugation as above. Total DNA was extracted using a Power Soil DNA Isolation kit (MoBio Laboratories, Carlsbad, CA, USA) and further concentrated by ethanol-sodium acetate precipitation and 75% ethanol washing. DNA pellets were briefly dried at room temperature, dissolved in sterile water after which the DNA quality was analyzed on a 2100 Electrophoresis Bioanalyzer Instrument (Agilent Technologies, Santa Clara, CA, USA). Subsequently, 7.5 µg of DNA was fragmented at 4100 rpm for 1 min with a Covaris g-TUBE device (Covaris Ltd, Brighton, UK) and 5 µg was used for preparing sequencing libraries (SMRTbell template prep kit 1.0, Pacific Biosciences, Menlo Park, CA, USA). DNA sequencing was performed on a PacBio RSII instrument (Pacific Biosciences) at the Lausanne Genomic Technologies Facility using three v3 SMRT™ Cells, P4-C2 chemistry and 180 min recording time. A total of 22,335 sequence reads with a mean length of 7,392 nt were *de novo* pre-assembled into 9 contigs using the Hierarchical Genome Assembly Process [21], yielding an average genome coverage of 73 times. Contigs were compared with each other and with an already existing assembly of the *P. veronii* 1YdBTEX2 genome (NCBI accession number AOUH01000001, 63 scaffolds) to detect possible sequence overlap, by using MAUVE [22]. Contigs were manually reassembled using Phrap and Consed [23] resulting in three final separate contigs. Ambiguous contig boundaries were further verified by PCR and manually corrected, when necessary.

Annotation

The genome of *P. veronii* 1YdBTEX2 was annotated with the automated pipeline of GenDB [24] and using Prodigal for gene prediction [25]. Annotations were further selectively and manually refined using key word searches and by comparison to SWISSPROT/UNIPROT databases [26] using Basic Local Alignment Search Tool (BLAST) [27]. Functional pathways were interpreted using the Kyoto Encyclopedia of Genes and Genomes (KEGG) [28, 29]. The following processes and genome elements were specifically examined: aromatic compound degradation, putative transporters and efflux pumps, chemotaxis and flagella, anaerobic and aerobic respiration, transposable elements, secretion systems, toxin-antitoxin systems and heavy metal resistances. Putative prophages in the *P. veronii* genome were identified using PHAST [30], whereas putative genomic islands were predicted using Islandviewer [31] followed by manual searching of repetitive boundary sequences. The web-based tool TA finder [32, 33] was used to detect possible toxin-antitoxin systems. Protein allocations to Cluster of Orthologous Groups were predicted using the Integrated Microbial Genomes & Microbiomes facilities [34].

Global transcriptional reactions of *P. veronii* to toluene and/or sand exposition

The genome-wide transcription response of *P. veronii* strain 1YdBTEX2 to toluene and/or sandy soil in comparison to liquid medium was analyzed using RNA-sequencing. Cells were inoculated ($\sim 10^7$ ml⁻¹) from exponentially growing liquid batch cultures on 10 mM succinate, as follows: 100 ml culture was centrifuged for 10 min at 4000 x g to collect the cells. Cells were resuspended in 15 ml of MM with 0.5 mM succinate, to avoid them going into starvation phase

during the 1 h assay. Aliquots of 100 μ l cell suspension were immediately transferred to 15 ml polypropylene screw cap centrifuge tubes containing either 2 g of sand with toluene (Sa-To), or 2 ml of MM with 5 mM succinate (Li-Su) or with toluene (Li-To). Further aliquots were complemented with up to 5 mM succinate and added to 2 g of sand (Sa-Su). Sand was collected in Spring 2015 from a beach near Lake Geneva ($^{\circ}$ 31.072'N, 6° 34.733'E) and dried at room temperature for 7 days. Toluene was dosed through the vapor phase by pipetting 100 μ l into a 1 ml micropipette tip, sealed on one end and inserted in the centrifuge tubes, which were closed during incubation. All treatments were started in quadruplicates and incubated for 1 h at 30°C at 160 rpm. Sand microcosms contained 4.8% gravimetric water content.

After 1 h, the cells were extracted by adding 5 ml of sterile saline solution (0.9% NaCl) to the tube and vortexing for 30 seconds. In the case of sandy microcosms, larger particles were allowed to sediment for a few seconds. Supernatants were filtered over 0.22 μ m PVDF Durapore membrane filters (Merck Millipore AG) by vacuum suction. The filter with the cells was removed, folded to have the cell layer facing inside, and immediately frozen in liquid nitrogen. Frozen filters were then placed into a Bead tube (MOBIO RNA PowerSoil[®] Total RNA Isolation Kit, MoBio Laboratories) and stored at -80°C until RNA isolation.

RNA isolation

RNA was extracted using the RNA PowerSoil Total RNA Isolation Kit (MoBio Laboratories). Frozen filters were removed from -80°C and coarsely crushed inside the tube with help of an RNase-free forceps, before continuing the bead-beating protocol as recommended by the manufacturer (MoBio Laboratories). The RNA pellet was resuspended in a final volume of 20 μ l of RNase-free water. Contaminating DNA was removed first by treatment with DNA-free Removal Kit (Ambion, ThermoFisher Scientific, Waltham, MA, USA), followed by TURBO DNase (Ambion). RNA was then purified using the RNeasy MinElute Cleanup kit (QIAGEN, Valencia, CA, USA).

RNA quantity and quality in the purified solutions was verified by reading 260 nm absorbance and the absorbance ratios at 260/280 nm and 260/230 nm, by using a NanoDrop spectrophotometer (ThermoFisher Scientific). RNA was migrated on an Agilent 2100 Bioanalyser (Agilent Technologies) to verify the presence of intact 16S and 23S rRNA. Genomic DNA contamination was checked by PCR using specific primers for a unique region in chromosome 2. If necessary, DNaseI treatment was repeated until no further DNA contamination was detected. RNA samples were then depleted from ribosomal rRNAs by using the Ribo-Zero rRNA Removal Kit Bacteria protocol (Epicentre, Madison, WI, USA). Subsequently, the RNA was reverse-transcribed, tagged on both ends, indexed and amplified by PCR using the ScriptSeq[™] v2 Bacteria kit and ScriptSeq[™] Index PCR primers set 1 (Epicentre). The resulting directional RNAseq libraries were sequenced using single 100-nt read chemistry on a Illumina HiSeq 2500 platform (Illumina, Inc., San Diego, USA) at the Lausanne Genomic Technologies Facility.

Data analysis

Read mapping, sorting and formatting of raw reads was done with Bowtie2 [35] and Samtools [36], using the finalized gapless *P. veronii* 1YdBTEX2 genome sequence as reference. Mapped reads were counted with HTSeq [37], then further processed and analysed with edgeR [38]. Only reads counted more than once per million in at least three replicates were kept. After normalisation of the counts, differential gene expression between pairwise conditions was tested in a modified Fischer exact test (as implemented in edgeR). ANOVA (also as implemented within

edgeR) was used to detect differential gene expression with interpretation groups “carbon source” (toluene or succinate) and “environment” (sand or liquid).

Genes were called significantly differentially expressed between two conditions at a false-discovery rate of <0.05 and a fold-change >2 , and were subsequently interpreted by using Gene Ontology (GO) analysis [39]. GO terms of *P. veronii* genes were inferred using the program BLAST2GO [40]. The web-based bioinformatics tool GOEAST [41] was then used to analyse GO data sets of significantly differentially expressed genes in each pair-wise comparison, under the TopGO “Weight” algorithm [42].

A network for the predicted aromatic compound metabolism of *P. veronii* 1YdBTEX2 was manually created in Cytoscape (version 3.2.1) [43] with nodes representing metabolites and edges representing gene expression data.

Gene marker exchange and nitrate respiration assay

In order to delete the *nar* genes, we amplified a 0.68 kb region upstream of *narL* (PVE_r1g2542) and a 0.8 kb region downstream of *narI* (PVE_r1g2549) using primers that introduce XmaI (5' ttttttttCCCGGGatgctcaccatcccc 3') and EcoRI (5' ttttttttGAATTCcaaccgcttgatcccc 3'), or XbaI (5' ttttttttTCTAGAggcagcttgagggttagatg 3'), and XmaI (5' ttttttttCCCGGGgcaaagcatgctcagcc 3') sites, respectively. PCR products were size-verified by agarose gel electrophoresis, purified using a Nucleospin Gel and PCR cleanup kit (Macherey-Nagel AG, Oensingen, Switzerland), ligated into pGEM-T-Easy (Promega AG, Dübendorf, Switzerland) and transformed into *Escherichia coli* DH5alpha. Plasmids were purified using a Nucleospin Plasmid kit (Macherey-Nagel) and sequenced for verification, after which inserts were recovered by digestion with XbaI and XmaI (Promega), or with XmaI and EcoRI, and ligated with plasmid pJP5603 I-SceI v2 [44], pre-digested with XbaI and EcoRI. *E. coli* DH5alpha-lambda-pir transformants were recovered on nutrient agar plates with $50 \mu\text{g ml}^{-1}$ of kanamycin. A plasmid with the correct inserts was purified and retransformed into *E. coli* S17-1 lambda-pir, from which it was transferred into *P. veronii* 1YdBTEX2 by conjugation as described elsewhere [45]. *P. veronii* recombinants were selected on MM agar plates containing $50 \mu\text{g ml}^{-1}$ kanamycin with toluene vapor as carbon source, and verified by PCR for the appropriate integration of the construct. Plasmid pSW-2 harbouring the gene encoding the SceI restriction enzyme was delivered into one of the *P. veronii* recombinants by a second conjugation, selecting for transconjugants on MM plates with 10 mM succinate, $10 \mu\text{g ml}^{-1}$ of gentamicin and 1 mM *m*-toluate (to induce *SceI* expression). *P. veronii* clones resistant to gentamicin but sensitive to kanamycin were recovered and verified for deletion of the *nar* genes by PCR. One clone (*P. veronii* Δ nar, strain 5239) was stored and tested for growth on nutrient broth (NB) medium in presence or absence of 1 g l^{-1} nitrate under aerobic or microaerophilic conditions. Aliquots of 0.1 ml stationary phase cultures of *P. veronii* 1YdBTEX2 or Δ nar grown on NB were transferred to 15 ml glass vials with screw-caps, containing either 7.5 ml or 15 ml of NB supplemented with 1 g l^{-1} potassium nitrate. The vials were incubated at 30°C for 16 h without shaking. Cell growth and the production of gas bubbles was inspected.

Database submission. The curated genome of *P. veronii* strain 1YdBTEX2 was deposited in the European Nucleotide Archive under bioproject number PRJEB11417. The RNA-seq reads were deposited in the NCBI Short Read Archive under accession number SRP077862.

Results

P. veronii 1YdBTEX2 genome features

Sequencing, assembly and manual refinement indicated that the genome of *P. veronii* strain 1YdBTEX2 consists of three replicons: a chromosome of 6,791,257 bp (named *chromosome 1*), a megaplasmid or second chromosome of 844,748 bp (named *chromosome 2*) and a plasmid of

373,858 bp (Fig 1). The plasmid has an extensive region of 23.5 kb (position 358719 spanning across to position 8288) highly similar to chromosome 1 (position 6746898 to 6770393, 99% identity). Chromosome 1 of *P. veronii* 1YdBTEX2 is closely related to *Pseudomonas fluorescens* SBW25 (AM181176.4), *Pseudomonas* sp. strain TKP (CP006852.1) and *Pseudomonas trivialis* IHBB745 (CP011507.1) (Fig 2, Figure A in S1 File), but aligns well with other *Pseudomonas* genomes such as *Pseudomonas syringae* pv. *syringae* B728a, *Pseudomonas putida* KT2440 or *Pseudomonas knackmussii* B13 (Fig 1A). In comparison to its closest chromosome relatives, *P. veronii* chromosome 1 has a large inversion of its central 2 Mb region (Fig 2, Figure A in S1 File). In contrast, there are no closely related sequences to chromosome 2 and the plasmid in public databases, although parts of other known plasmids can be detected (Fig 1B and 1C).

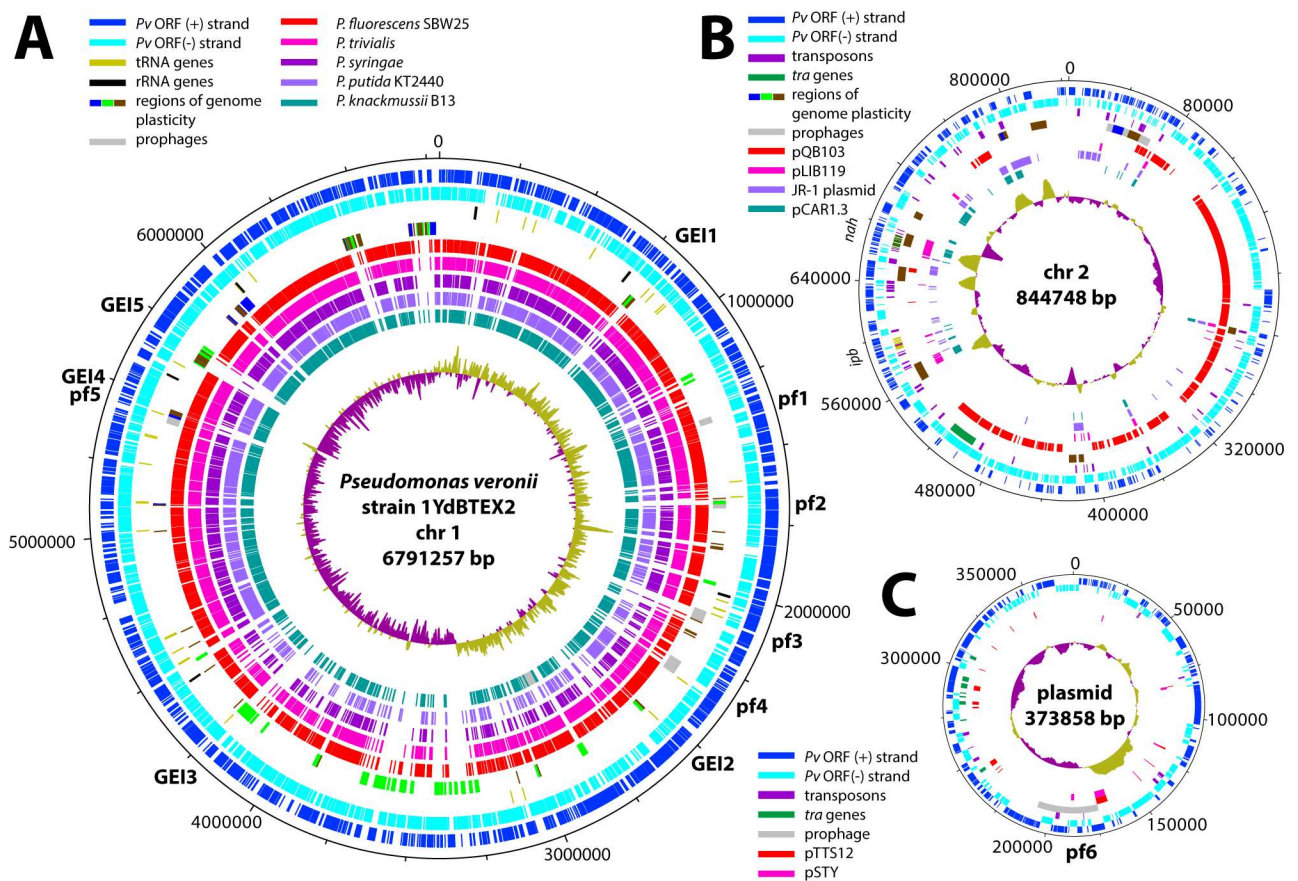


Fig 1. Circular maps of the replicons encompassing the *P. veronii* 1YdBTEX2 genome. (A) Chromosome 1 (chr1) with indication of possible genomic islands (GEI) and prophages (pf). The outermost circles show the location and orientation of predicted coding regions (blue and cyan), followed by tRNA (olive green) and rRNA genes (black), predicted regions of genome plasticity (blue-green-brown) islands and prophages (grey). The inner circles represent BLASTN comparisons with the close relatives *P. fluorescens* SBW25 (red, Acc. No. AM181176.4), *P. trivialis* strain IHBB745 (deep pink, CP011507.1), *P. syringae* pv. *syringae* B728a (dark purple, CP000075.1), *P. putida* KT2440 (light purple, AE015451.1) and *P. knackmussii* B13 (persian green, HG322950). GC skew (dark magenta and yellow green) is shown in the most central circle. (B) As A, but for the chromosome 2 replicon (chr2). Inner circles, from outwards to inwards, predicted transposons (dark purple) and *tra* genes (green), regions of genome plasticity (blue-green-brown) and prophages (grey), followed by BLASTN comparisons to *P. fluorescens* SBW25 plasmid pQB103 (red, AM235768.1, NC_009444.1), *Pseudomonas stutzeri* strain 19SMN4 plasmid pLIB119 (deep pink, CP007510.1), *Pseudomonas mandelii* JR-1 plasmid (dark purple, CP005961.1) and *Pseudomonas resinovorans* NBRC 106553 plasmid pCAR1.3 (Persian green, AP013069.1). (C) As B, but for the plasmid replicon. The inner circles represent the BLASTN comparisons with *P. putida* S12 plasmid pTTS12 (red, CP009975.1), and *Pseudomonas* sp. VLB120 plasmid pSTY (purple, CP003962.1). Plots generated with DNAPlotter [46].

doi:10.1371/journal.pone.0165850.g001

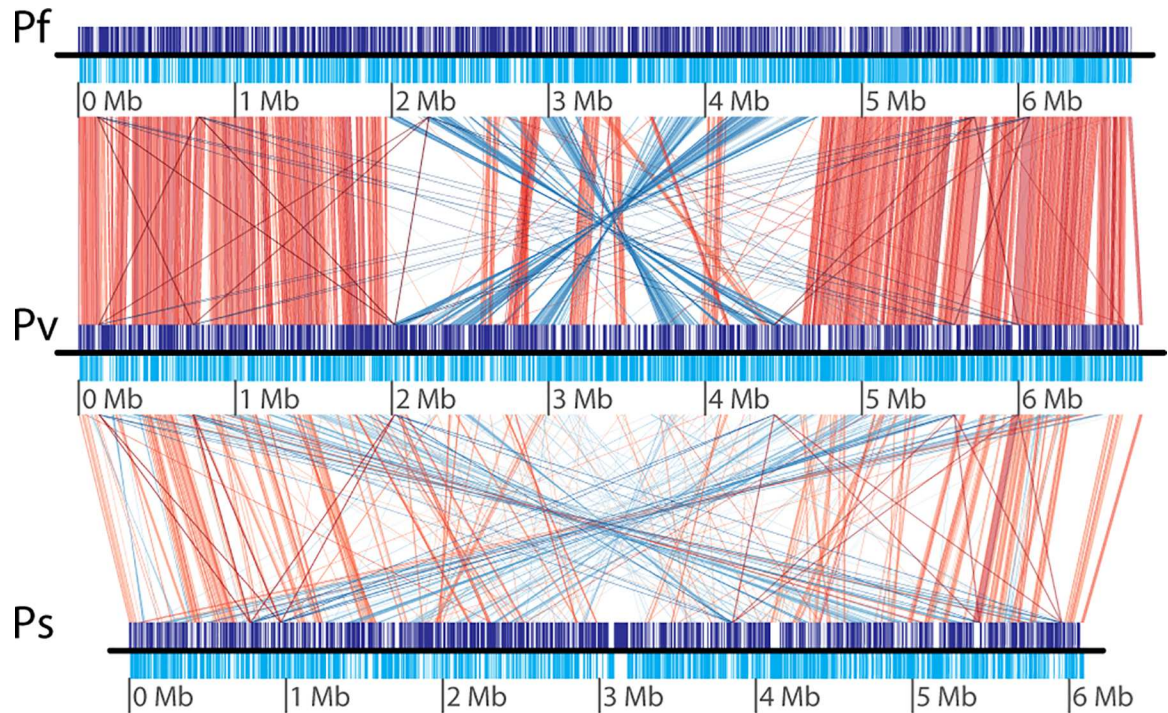


Fig 2. Comparison of the linearized *P. veronii* chromosome 1 replicon (Pv) to *P. fluorescens* SBW25 (Pf, Acc. No. AM181176.4) and *P. syringae* pv. *syringae* B728a (Ps, CP000075.1). Purple and cyan blocks show the location and orientation of coding regions on the positive and negative strands, respectively. Red and blue lines indicate direct and inverted colinear regions between both replicons, respectively, using a threshold of percentage nucleotide identity of 75%, maximum e-value for the region comparison of 1×10^{-10} , and minimum overlap length for display of 1 kb. Plot generated with GenoPlotR [47].

doi:10.1371/journal.pone.0165850.g002

PHAST analysis identified 6 putative prophages in the *P. veronii* genome, five of which occur on chromosome 1 and one on the plasmid (Fig 1, *pf*, Table A in S2 File). Four of the prophages seemed intact and complete (pf1, 3, 4 and pf6, Fig 1) according to manual annotation and inspection. Putative prophage pf5 overlaps with a predicted genomic island (GEI4). IslandViewer results suggested 47 regions of genome plasticity (RGP) on chromosome 1 (Table B in S2 File). Five of those (GEI1-5, Fig 1A, Table B in S2 File) carry an integrase gene nearby or encompass a gene for a tRNA, and have a boundary repeat sequence, which are typical characteristics for a genomic island [48]. However, no genes encoding clear conjugation systems (e.g., type IV secretion system) could be identified within those five GEIs. Notably, there is a hyper-variable region on chromosome 1 between position 3.2 and 3.8 Mb (Fig 1A), however, without characteristic gene markers for genomic islands. Eleven (chromosome 2) and one (plasmid) additional RGP were detected by IslandViewer on the other two replicons (Fig 1B and 1C). As these two replicons contain numerous genes encoding transposable elements or integrases (Fig 1B and 1C, Table B in S2 File), these may well have been implicated in creating the observed plasticity.

A total of 62.4% of genes on chromosome 1 could be assigned to clusters of orthologous groups (COGs), but only 24.2% and 30.0% of the genes on chromosome 2 and the plasmid, respectively (Figure B in S1 File), highlighting the large number of proteins of unknown function encoded in those two replicons. Compared to chromosome 1, chromosome 2 and the plasmid have a higher proportion of genes belonging to COG groups X (Mobilome: prophages and transposons), group L (Replication, recombination and repair) and M (plasmid, Cell wall,

Figure B in [S1 File](#)). Chromosome 2 carries no genes for particularly obvious basic cellular functions, except two genes for sigma factors (PVE_r2g0923, *rpoS*; PVE_r2g0399, sigma-32), two genes for ribosomal proteins, and a second copy of the gene for DNA polymerase III beta-subunit (PVE_r2g0243).

All genes required for aerobic respiration, such as NADH dehydrogenase, succinate dehydrogenase, cytochrome *c* oxidase, cytochrome *c* reductase, cytochrome *bd* complex, and an F-type ATPase were found encoded on chromosome 1 (Table C in [S2 File](#)). In addition, a superoperonic cluster encoding the enzymes required for a complete denitrification pathway (*nar*, *nir*, *nor* and *nos* genes, position 2644528 to 2685333) is located on chromosome 1, enabling the strain to grow under microaerophilic and anaerobic conditions with nitrate (Table D in [S2 File](#), [Fig 3A and 3B](#)).

A complete set of genes for flagellar biosynthesis (e.g., *fli*, *flh* and *flg*) and classical chemotaxis (e.g., *cheA*, *Y*, *Z*, *W* and *B*) is encoded on chromosome 1 in a single large cluster (PVE_r1g4269-4321). Additional chemotaxis components are also encoded, which might be associated with swarming or twitching motility (e.g., PVE_r1g4595-4604 and PVE_r1g3912-3921, Table E in [S2 File](#)). *P. veronii* chromosome 1 further encodes a complete type I secretion system, a SecA-dependent type II system, putatively two type III systems, and three or four type VI secretion systems (Table F in [S2 File](#)). No complete type IV secretion system is encoded on chromosome 1, but both chromosome 2 and the plasmid carry genes encoding proteins with weak similarity to type IV secretion systems and may thus constitute atypical conjugative systems (Table F in [S2 File](#)). A large variety of putative toxin/antitoxin pairs were predicted by the web-based tool TA finder [33], 17 of which are found on chromosome 1, nine on chromosome 2 and two on the plasmid, covering known (e.g., *vapBC*, *higAB*, *relEB*, *hipAB*, *hicAB*) as well as less well-known families (Table G in [S2 File](#)).

Putative heavy metal resistances encoded by *P. veronii* include copper (*copRSCD*, *copAB*, both on chromosome 1, 2 and plasmid), mercury (*merRTPCA*, on chromosome 2), tellurite (*terZABCDE*, chromosome 2), chromium-zinc-cobalt (*czcRABCD*, chromosome 1), and arsenic (*arsRBCH*, twice on chromosome 1) (Table H in [S2 File](#)). Complete gene clusters for catabolism of 3-hydroxyphenylpropionate, toluene (see below), salicylate, phenol, anthranilate, vanillate and 4-hydroxybenzoate were detected in the genome of *P. veronii* 1YdBTEX2,

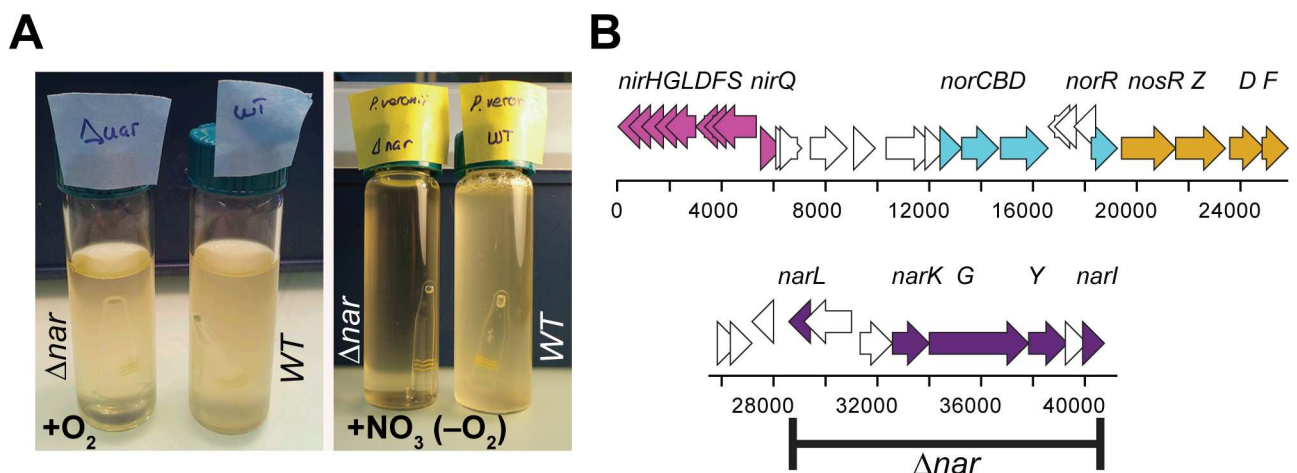


Fig 3. Overview of denitrification capacity of *P. veronii* 1YdBTEX2. (A) Overnight growth of *P. veronii* 1YdBTEX2 wild type (WT) and the Δnar mutant in presence (+O₂, left) or absence of air but with 15 mM nitrate supplemented medium (+NO₃, -O₂, right panel) conditions. Note the gas formation in the right panel of the WT incubation. (B) Gene regions predicted for denitrification in the *P. veronii* 1YdBTEX2 chromosome 1 with trivial gene names indicated. Black bar represents the deleted region in *P. veronii* Δnar .

doi:10.1371/journal.pone.0165850.g003

scattered on all three replicons (Figure C in [S1 File](#)), allowing both *meta*- and *ortho*-cleavage of catechol intermediates. In contrast, no genes for benzoate 1,2-dioxygenase were found and *P. veronii* 1YdBTEX2 does not grow on benzoate.

Analysis of the immediate catabolic response of *P. veronii* 1YdBTEX2 to toluene

P. veronii strain 1YdBTEX2 is mostly known for its capacity to degrade aromatic solvents such as BTEX [[15](#), [16](#), [49](#)]. Toluene and benzene are degraded through the same catabolic pathway but the enzymes potentially catalyzing these reactions are encoded in at least three to four gene clusters, several of which seem redundant [[49](#)]. Degradation of BTEX is thought to proceed via a multi-component aromatic ring dioxygenase, followed by a benzene-dihydrodiol dehydrogenase yielding catechol intermediates, and a *meta*-cleavage pathway, finally producing acetyl-CoA ([Fig 4D](#)) [[49](#)]. One of the involved gene clusters (*ipbAaAbAcAdBCEGFHD*) is localized on chromosome 2 ([Fig 1B](#), [Fig 4A](#)), and includes both the genes for the aromatic ring dioxygenase (*ipbAaAbAcAd*) and a set of *meta*-cleavage pathway genes (*ipbCEGFHD*). However, the gene for benzene-dihydrodiol dehydrogenase (*ipbB*), is preceded by a transposon insertion and contains two non-sense mutations that interrupt correct translation ([Fig 4A](#)). The *P. veronii* 1YdBTEX2 genome encodes two further *meta*-cleavage pathways, both located on chromosome 2, within what seem to be a phenol degradation pathway (*dmpRBCDEFGH*, [Fig 4B](#)) and a salicylate catabolic pathway (*nahRGTHINLOMKJ*, [Fig 4C](#)). PVE_r2g0805, also located on chromosome 2, may be encoding an alternative benzene-dihydrodiol dehydrogenase (on average 56% amino acid similarity to the three *ipbB* protein fragments).

To study which genes from those pathways are specifically induced upon exposure to toluene, which could indicate their implication in BTEX metabolism, a controlled RNA-seq transcriptome analysis was carried out. Four conditions were tested: cells incubated in liquid medium (Li) or in sand (Sa), and exposed to toluene (To) or to succinate (Su). RNA-seq analysis showed good replicate clustering for the liquid conditions and replicates grouped closely together in the Principal Component Analysis (both succinate and toluene, [Fig 5A](#)). A higher variability was observed among sand replicates, perhaps due to the higher heterogeneity of sand compared to liquid medium ([Fig 5A](#)).

Two-way ANOVA showed a total of 1114 genes with at least two-fold differential expression as a function of carbon source (false discovery rate < 0.05, $P < 0.01$), of which 319 (183 up and 136 down) were unique for carbon source and the remainder overlapping with the factor *environment*, ([Fig 5B](#)). About 50% of those encode conserved yet hypothetical proteins. The *ipbAaAbAcAd* genes, the *dmp* gene region and genes for a presumed BTEX efflux pump (*ttgGHI*, PVE_r2g0377-0383) showed among the highest observed transcript increase in presence of toluene (up to 54-fold, [Fig 5C](#)). The gene for C₄-dicarboxylate transport (PVE_r1g4585) was among those whose expression decreased the most under exposure to toluene (down by 27-fold).

To analyze the expression changes of *P. veronii* genes potentially implicated in aromatic compound metabolism more in general, we constructed a network with metabolites as nodes and known or predicted *P. veronii* gene functions as edges ([Fig 6](#), Figure C in [S1 File](#)). Upon exposure to toluene there was a clear change in the expression level of the genes suspected to be implicated in BTEX metabolism, whereas expression of genes coding for other aromatic metabolic pathways remained unaltered ([Fig 6](#), Figure D in [S1 File](#)). A clear exception was the gene for the putative alternative benzene-dihydrodiol dehydrogenase (PVE_r2g0805), whose expression was unaffected by the change from succinate to toluene ([Fig 6](#), compare Li-Su with Sa-To). This gene (PVE_r2g0805) has previously been proposed to be recruited and to replace

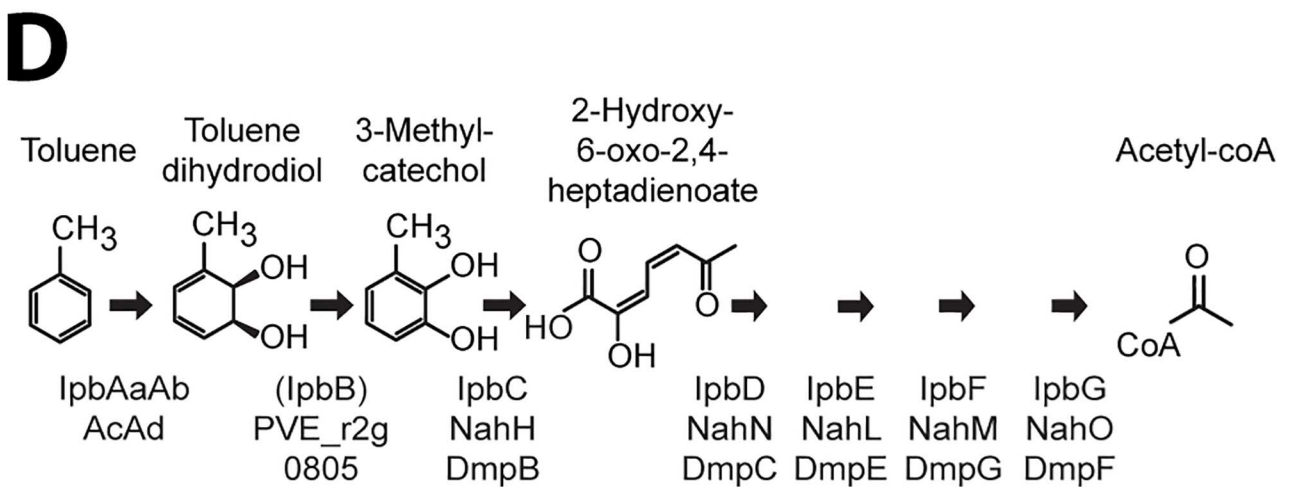
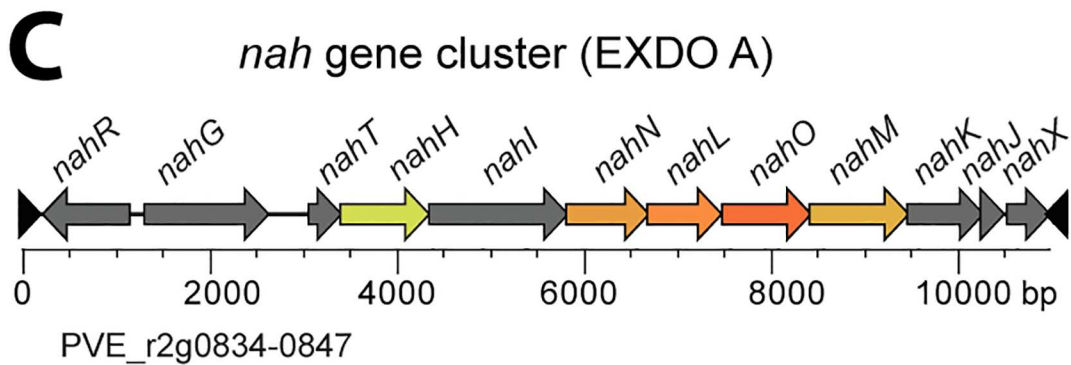
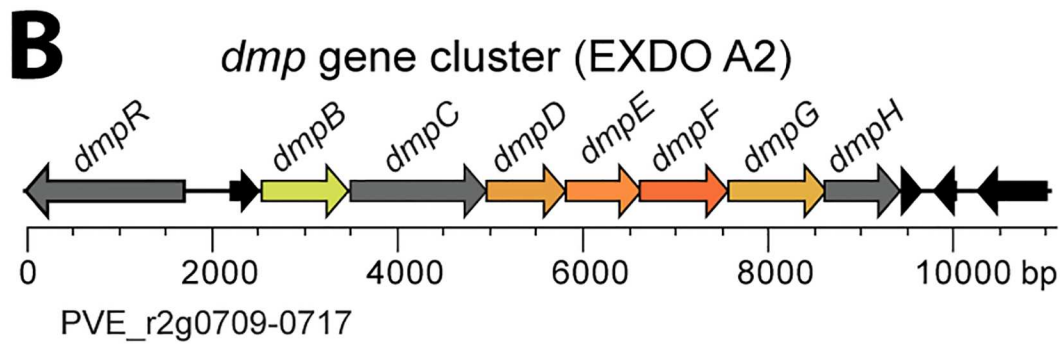
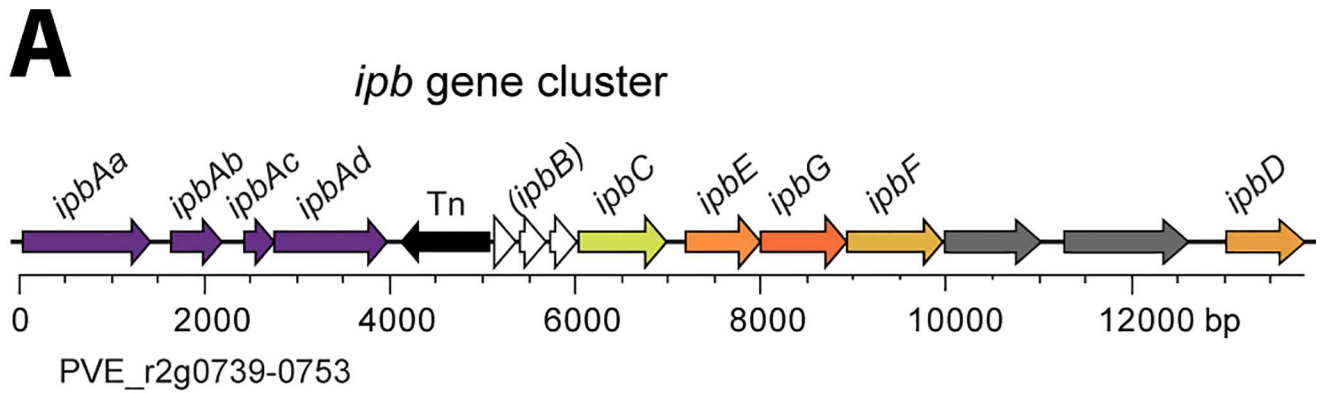


Fig 4. Organization of the clusters encoding genes involved in toluene degradation and catechol meta-cleavage pathways by *P. veronii* 1YdBTEX2. (A) Isopropylbenzene gene cluster (*ipb*). (B) *dmp*-like cluster (previously named EXDO A2). (C) *nah*-like meta-cleavage pathway genes (previously named EXDO A). Genes represented by arrows scaled to size (bp, base-pair) and relevant gene names as well as locus_tag (e.g., PVE_r2g0834) are indicated. Paralogous genes are colored similarly. Note the interruption in the *ipbB* reading frame by two stop codons (white arrows). Tn, transposable element (black arrows). (D) Proposed metabolic pathway for toluene to acetyl-CoA conversion with relevant chemical structures. Black arrows represent a single enzymatic step with the proposed involved (redundant) gene products indicated below.

doi:10.1371/journal.pone.0165850.g004

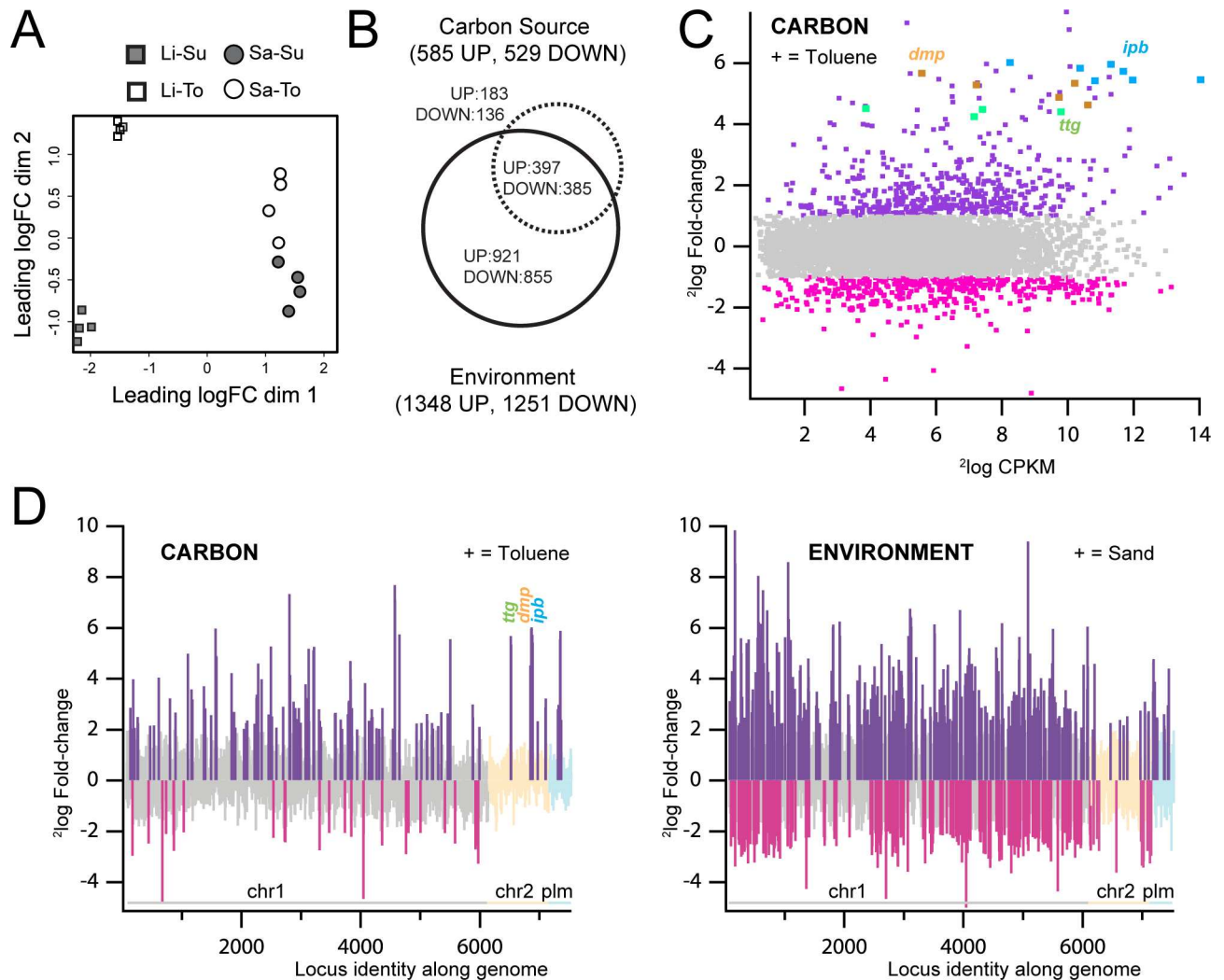


Fig 5. Genome-wide gene expression differences in *P. veronii* 1YdBTEX2 after 1 h exposure to different carbon sources or growth environment. (A) Two-dimensional Principal Component Analysis of quadruplicate global RNA-sequencing data sets of *P. veronii* 1YdBTEX2 incubated in liquid medium with succinate (Li-Su), or toluene (Li-To), or in sand with succinate (Sa-Su) or with toluene (Sa-To). (B) Venn diagram with the number of unique and common genes significantly differentially expressed (2-way ANOVA, $^2\log$ -fold-change [\log FC] >1, false-discovery rate [FDR] <0.05, P <0.01) as result of change of carbon source (succinate to toluene) or environment (liquid to sand). (C) Smear-plot of global gene expression intensity ($^2\log$ CPKM) versus expression changes ($^2\log$ fold change) compared between cells incubated with toluene (Li-To) versus succinate (Li-Su); in grey, genes not statistically differentially expressed (\log FC <1, FDR >0.05, P >0.01); magenta, genes with lower, and dark purple, genes with higher expression in presence of toluene (+). Blue, *ipb* genes; yellow, *dmp* genes; green, *ttg* genes (toluene efflux pump). (D) Gene expression changes as an effect of carbon source (succinate versus toluene, left) or of environment (liquid versus sand, right), and plotted as function of genomic location (chromosome 1, chr1; chromosome 2, chr2 and plasmid, plm; organized according to locus_tag number). Bars indicate $^2\log$ -fold change. Dark purple, statistically significantly higher expressed genes in presence of toluene (+, left) or sand (+, right); cyan, lower expressed genes in pink. Positions of the *ipb*, *dmp* and *ttg* genes are highlighted.

doi:10.1371/journal.pone.0165850.g005

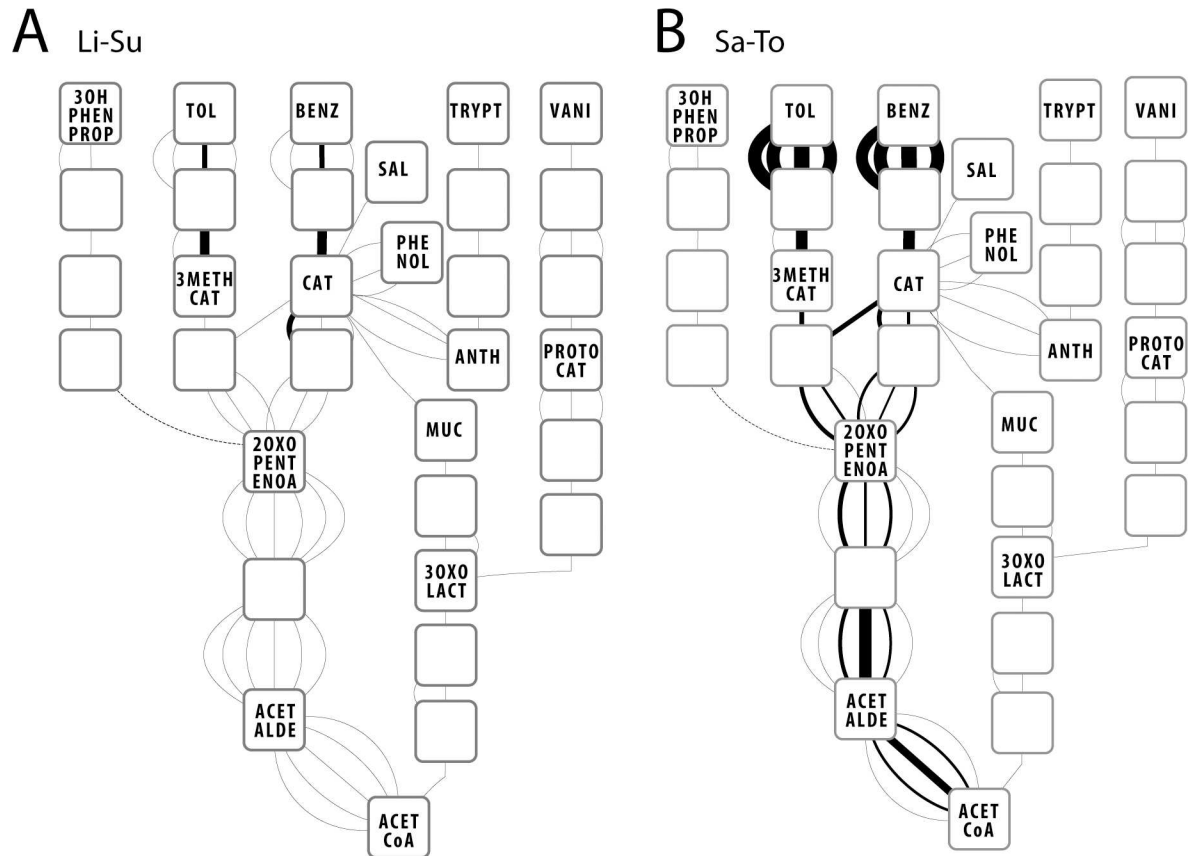


Fig 6. Network analysis of gene transcription for deduced aromatic compound metabolism of *P. veronii* 1YdBTEX2 cultures in (A) liquid medium with succinate (Li-Su) and (B) sand with toluene (Sa-To). Nodes represent substrates and metabolic intermediates. Edges represent the $^2\log$ -transformed average transcription across quadruplates ($^2\log$ CPKM, counts per kilobasepair per million) of the gene coding for the particular enzyme that carries out the reaction between two nodes. Line width linearly decreasing between 100 and 1 for $^2\log$ CPKM between 50 and 5. $^2\log$ CPKM values lower than 5 all have a line width of 1. 2OXOPENTENOA; 2-oxopent-4-enoate, 3OHPHENPROP; 3-(3-hydroxyphenyl)propanoate, 3OXOLACT; 3-oxoadipate enol lactone, 3METHYCAT; 3-methylcatechol, ACETALDE; acetaldehyde, ACETCoA; acetyl-coA, ANTH; anthranilate, BENZ; benzene, CAT; catechol, MUC; cis,cis-muconate, PROTOCAT; protocatechuate, PHENOL; phenol, SAL; salicylate, TOL; toluene, TRYPT; L-tryptophan, VANI; vanillin. See Figure C in S1 File for details of all involved locus numbers.

doi:10.1371/journal.pone.0165850.g006

the interrupted *ipbB* gene within the *ipb* operon [18, 49, 50]. Multiple redundant genes encoding parts of *meta*-cleavage pathways for catechol metabolism are induced upon exposure to toluene. They are located on all three replicons (Fig 6, Figure C and D in S1 File) and one is part of a predicted prophage on the plasmid replicon (pf6, Table A in S2 File). No difference in the types of induced pathways was observed between liquid medium (Li-To) or sand.

By plotting the normalized read counts across the *ipb* gene cluster under conditions of liquid medium with succinate (Li-Su) or with toluene (Li-To), one can easily discern the premature transcriptional termination downstream of the *ipbAd* gene as a result of the transposon insertion (Fig 7A). Hence, whereas transcription of *ipbAa-Ad* is induced around 50-fold by exposure to toluene, there is basically no induction of the downstream genes for the *meta*-cleavage pathway (Fig 7B). In contrast, the genes in the *dmp* cluster were significantly induced by toluene in liquid medium (up to 20 times, Fig 7C), while transcription of those in the *nah* cluster remained unchanged (Fig 7D).

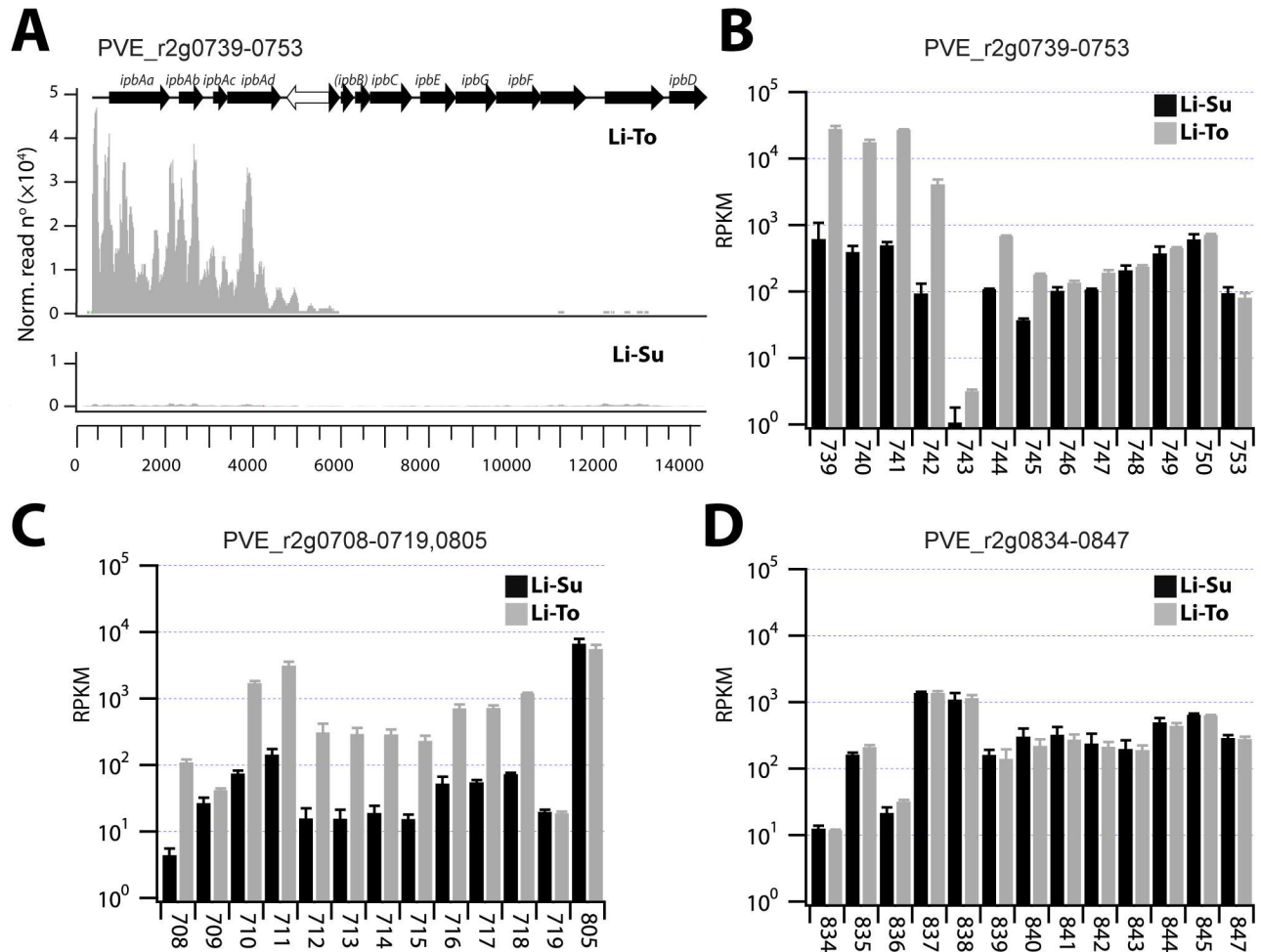


Fig 7. Comparison of catabolic gene transcription involved in toluene or meta-cleavage metabolism by *P. veronii* 1YdBTEX2 in liquid culture with succinate (Li-Su) or toluene (Li-To). (A) Normalized read counts across the *ipb* gene cluster (PVE_r2g0739-0753). Note the decrease as a result of the transposon insertion (white arrow). (B) Expression level (reads per kilobase per million, RPKM, ¹⁰log scale) of the *ipb* cluster genes (numbers refer to PVE_r2g loci). (C) as B, for the *dmp* cluster genes (PVE_r2g0708-0719), and the proposed gene encoding for the dihydrodiol dehydrogenase (PVE_r2g0805). (D) as B, for the *nah* cluster genes (PVE_r2g0834-0847).

doi:10.1371/journal.pone.0165850.g007

Other functions affected by toluene exposure

As expected, GO analysis of genes whose expression is higher upon toluene exposure showed clear enrichment in the category *biological processes* for terms associated with aromatic compound metabolism (Tables I-K in [S2 File](#)). Terms associated with aerobic respiration, leucine biosynthesis and glyoxylate metabolism were also enriched during toluene exposure (Table K in [S2 File](#)). In contrast, for genes with decreased expression upon exposure to toluene, the following terms were enriched: cofactor and amino acid biosynthesis, energy reactions, reduced translation and post-translational reactions, as well as changes in the cell envelope. This suggests major cellular changes upon transition to toluene as growth substrate.

The expression of the gene PVE_r1g2050, which codes for a methyl-accepting chemotaxis protein (McpS), increased both in liquid and sandy microcosms upon induction with toluene (11.3 and 7.5 fold-change, respectively, Table J in [S2 File](#)). This may suggest that *P. veronii* has a similar chemotactic response towards aromatic solvents as *Pseudomonas putida* F1 [51]. Two chromosomally encoded toxin/anti-toxin systems (PVE_r1g1458/1459, PVE_r1g3189/3190)

were also up-regulated upon induction with toluene (8-fold and 3-fold, respectively, Table G in [S2 File](#)). One of the predicted type VI regions (PVE_r1g5937-5956) showed a 2–4 fold lower expression in presence of toluene (Table F in [S2 File](#)).

Analysis of the immediate response of *P. veronii* 1YdBTEX2 to change of environment

Gene expression differences as a function of growth environment (liquid culture versus sand) were far more dramatic than those as a function of carbon substrate. A total of 1348 genes (of which 921 unique) increased and 1251 genes (855 unique) decreased expression in cells exposed to sand compared to liquid medium ([Fig 5B and 5D](#)). GO analysis of enriched pathways suggests that cells in sand increase expression of osmoregulatory systems (e.g., alginate, glutamine, polyamine synthesis), and stress systems (e.g., oxidative stress, mercury defense, protein chaperones). They further redirect signaling and transcriptional regulation, produce reserve material (e.g., glycogen, fatty acid catabolism), and reprogram their central metabolism (Tables L–M in [S2 File](#)). A large variety of other genes are less expressed in cells added to sand compared to liquid, notably including those for RNA precursor synthesis and modification, respiration and cell shape (Table L in [S2 File](#)).

Several groups of genes putatively involved in osmoregulatory adaptation showed increased expression in sand, such as an *osmVWXY* osmoprotectant import system (PVE_r1g0889-0892), or transporters and synthesis of glycine-betaine (PVE_r1g0119,0120; PVE_r1g0416,0417, PVE_r1g5571-5576). Furthermore, two gene clusters associated with putrescine uptake (*potGHI* genes) and putrescine metabolism (*puuABGHF* PVE_r1g3330-3337) were up to 64-fold more expressed in sand (Table M in [S2 File](#)). These genes may code for a putrescine transamination system, leading to the osmoprotectant glutamate. In addition, all the genes coding for the KdpABC high-affinity potassium transport ATPase and its regulation (PVE_rg1744-1746) were up to 56-fold more expressed in sand (Table M in [S2 File](#)). This system is known to become expressed when turgor pressure is low [[52](#), [53](#)]. Finally, also genes putatively encoding trehalose biosynthesis (PVE_r1g2917,4895–4897) were higher expressed in sand (up to 42-fold, Table M in [S2 File](#)).

Other interesting individual genes or gene clusters whose expression changed significantly in sand include several for polysaccharide synthesis (PVE_r1g0180,0181; 1051–1055, 1533,1534; 3441–3446, Table M in [S2 File](#)), or for pili synthesis (PVE_r1g0867,0868; 1320, 4810). A long cluster of genes with encoded transporter systems of unclear specificity exhibited up to 35-fold higher expression in sand (PVE_r1g2142-2157, Table M in [S2 File](#)). Among the genes with the highest increase of expression in sand is a small cluster for hypothetical proteins (PVE_r1g0182-0184), with up to 650-fold higher expression. This suggests that these genes might play an important role in coping with adaptation to the soil environment. Several genes implicated in aromatic compound metabolism were also more expressed in sand (up to 10-fold), independently of toluene addition, suggesting the presence of other available aromatic growth substrates (e.g., PVE_r1g1095-1097; 1438–1446; 3731,3732, Table M in [S2 File](#)). We further identified higher expression levels of a cluster of genes in the plasmid involved in the resistance to copper (PVE_p0049-51, five-fold), of mercury resistance genes on chromosome 2 (PVE_r2g_0813–0817, threefold), and of chromosome 1 genes for a cobalt-zinc-cadmium efflux system (PVE_r1g06092-6099, up to 16-fold, Table H in [S2 File](#)). Part of prophage 2 (*pf2*) genes (PVE_r1g1579-1585) were up to 8-fold more expressed in sand, whereas many genes on *pf5* were less expressed in soil (Table A in [S2 File](#)). Additionally, three genes for an integrase showed increased expression in sand (PVE_r1g2409, twofold; PVE_p0204, p0207, 8-fold), suggesting that this condition can trigger movement of mobile DNA.

Discussion

We determined a complete gapless genome sequence of *P. veronii* strain 1YdBTEX2, a wild-type BTEX degrader isolated from jet-fuel contaminated soil [16]. The genome measured a little over 8 Mb, which is almost 1.3 Mb larger than previously reported [18]. The genome appeared to consist of three replicons, which is unusual for pseudomonads. Chromosome 1 is most closely related to *P. fluorescens*, *Pseudomonas* sp. TKP and *P. trivialis* (Fig 2, Figure A in S1 File), although it contains a large number of unique regions (RGPs 1–27). Among these are four likely intact and one partial prophages, and suspected genomic islands (Fig 1), which underscores the various gene rearrangements by the wild-type strain 1YdBTEX2 with potential adaptive benefit. *P. veronii* 1YdBTEX2 contains a full set of genes essential for denitrification (Fig 4). This may enable the strain to survive and/or grow more easily under microaerophilic conditions in water-saturated soil. The two other replicons of the *P. veronii* genome have very little similarity to other known bacterial replicons (Fig 1B and 1C), and a large proportion of their genes code for hypothetical proteins, conserved or not (Figure B in S1 File). Although not tested here, both chromosome 2 and plasmid replicons may be self-transferable, which we deduce from the presence of gene functions with weak similarity but detectable homology to known type IV secretion systems (Table F in S2 File).

P. veronii 1YdBTEX2 is not particularly versatile in aromatic compound degradation (Fig 6), but the known pathways for BTEX degradation are all encoded on the chromosome 2. The hypothesis that BTEX degradation proceeds via an initial dioxygenation reaction catalyzed by the *ipbAa-Ad* gene products (PVE_r2g0739-0742) is supported by their more than 50-fold induction upon exposure to toluene (Fig 7A and 7B). The gene for dihydrodiol dehydrogenase (*ipbB*) is inactivated by two non-sense mutations and is preceded by a transposon insertion (Fig 4A). The dihydrodiol dehydrogenase function in toluene degradation is most likely complemented by the constitutively expressed gene PVE_r2g0805 (Fig 7C) [49]. Curiously, three *a priori* redundant *meta*-cleavage pathways are encoded on chromosome 2, the first similar to an archetype *dmp* phenol pathway, the second part of the *ipb* pathway and the third similar to a pathway from naphthalene/salicylate degradation (Fig 4). We confirmed the previous observation [49] that induction of the *ipb meta*-cleavage pathway in presence of toluene is abolished as a result of transcription termination inside a transposon inserted upstream of *ipbB* (Fig 7A and 7B). In contrast, the genes for the *meta*-cleavage pathway within the *dmp*-like cluster (PVE_r2g0710-0718) are induced upon toluene exposure, whereas those of the *nah*-like cluster (PVE_r2g0835-847) are not (Fig 7C and 7D). Nonetheless, the genes in all three clusters may contribute to *meta*-cleavage of catechol intermediates, since their expression levels in presence of toluene are quite similar (Fig 7, between 10^2 and 10^3 RPKM).

It is likely that effective growth on toluene of *P. veronii* 1YdBTEX2 also depends on an efflux system, which is encoded on chromosome 2 as well (PVE_r2g0377-0383). This three-component efflux pump and its two-component regulators are homologous to the *ttgGHI* system of *Pseudomonas putida* DOT-T1E [17] and are strongly induced in presence of toluene (Fig 6). It is plausible that *P. veronii* is even able to use toluene as a chemo-attractant, given the 10-fold induction of the *mcpS* gene PVE_r1g2050 on chromosome 1 (Table M in S2 File). This signal might be coupled to the flagellar systems encoded on chromosome 1 (Table E in S2 File). Toluene also induced expression changes in a range of other genes affecting various processes (Table J in S2 File). The induction of two toxin-antitoxin systems by toluene was intriguing (PVE_r1g1458/1459, PVE_r1g3189/3190), and may reflect a more general stress type response. For example, the *higBA* locus in *Mycobacterium tuberculosis* is more expressed under heat shock, nutrient starvation, DNA damage and hypoxia, supporting the hypothesis that it is involved in survival under harmful conditions [54]. *Sphingomonas wittichii* strain RW1, a

degrader of dibenzodioxins, also induces an (entericidin-family) toxin-antitoxin upon exposure to sand [11].

The transition from liquid medium to sand (at 4.8% gravimetric water content) caused a plethora of changes in genome-wide gene expression patterns of *P. veronii* (Fig 5D, Fig 6C). This is surprising given that cells were resuspended in the same medium to inoculate the sand and the liquid. Likely, the small volume of inoculated liquid (with cells) formed small water films around sand particles and filled small pores between sand particles, which may have changed the availability of oxygen to the cells. Indeed, expression of genes involved in aerobic respiration was significantly reduced in sand compared to liquid, which we interpret as a better provisioning of oxygen to the cells in sand. Although counterintuitive, this makes sense given the low overall water content of the sand (4.8%). Diminished gene expression for aerobic respiration in sand under the same conditions was also a hallmark of the initial reaction of *S. wittichii* RW1 upon inoculation to sand contaminated with dibenzofuran or salicylate [11]. Most likely, the inoculated *P. veronii* cells also react immediately to the availability of micronutrients, such as copper and zinc (upregulated expression of transport systems in sand), but also to toxic trace metals like mercury (induction of the mercury defense system in sand). No changes in gene expression of transporters for nitrogen, phosphorous or sulfur were found upon inoculation in sand, but we assume that sufficient nutrients had been provided with the medium.

To better understand the changes observed after inoculation into sand and the cell response to water stress, we compared our results with previous data on differential gene expression in *P. veronii* 1YdBTEX2 induced upon matric (by addition of polyethylene glycol) or solute stress (addition of NaCl) [15]. Of the 274 differentially expressed genes in *P. veronii* 1YdBTEX2 during 1 h contact under either solute or matric stress, 150 were also significantly differentially expressed after 1 h contact in sand (Table O in S2 File). However, this is only 6% of all genes expressed in sand with at least two-fold difference (and $FDR < 1 \cdot 10^{-5}$). Therefore, we conclude that gene expression programs in sand have little in common with reactions to solute and matric stress, and that those are not appropriate proxies for the former as was previously postulated [15]. However, certain reactions were quite similar between solute or matric stress and incubation in sand, notably those involving Table O in S2 File). In contrast to what we expected, genes implicated in flagella biosynthesis and motility that previously had been found as marker genes for response to sand and water stress [11, 15] were not significantly differentially expressed in *P. veronii* 1YdBTEX2 under the conditions used here (Table O in S2 File). The reason for decreasing flagellar gene expression was proposed to be a relay of cellular energy for stress defense rather than for flagellar production and maintenance [55]. Nevertheless, others have shown that the reduction of motility of pseudomonads in response to water stress is not a generalized bacterial response to water-limiting conditions [56].

The accumulation of the compatible solutes glutamate and the uptake of potassium, mediated by the induction of the *kdpABC* operon, seems to be among the strongest “osmoprotectant strategies” employed by *P. veronii* 1YdBTEX2 cells to balance osmotic differences caused by alteration in the solute potential of the extracellular environment upon inoculation in sand (despite cells being inoculated in the same medium). Glutamate accumulation has been reported in *Pseudomonas syringae* strains B728 and DC30000 after osmotic stress, and the *kdp* operons in *Salmonella typhimurium* and *E. coli* are also induced upon K^+ limitation and by osmolarity stress or turgor loss [57–59]. In addition, previous similar inoculation experiments in contaminated sand with the dibenzofuran-degrading bacterium *S. wittichii* also indicated the prime importance of glutamate biosynthesis to compensate for osmolarity differences [15]. Accordingly, we attribute the increase in putrescine metabolism (which produces glutamate) to the same osmotic response of *P. veronii* in sand. Following a fast potassium influx that increases not only the intracellular osmolality, but also the number of positive charges within

the cells, polyamine efflux may occur to compensate for excess ionic charges. This response would be in line with three classical physiological responses of *E. coli* to osmotic stress: polyamine transport and accumulation of glutamate and potassium [58]. Analysis of enriched GO terms also indicated higher expression of *P. veronii* 1YdBTEX2 genes for trehalose and polysaccharide biosynthesis such as alginate upon contact to sand. This is similar to stimulation of alginate production by high osmolality in other *Pseudomonas* species [57], and alginate production has been proposed as a compensatory osmoadaptation mechanism in *Pseudomonas syringae* B728a [59].

Our data support the conclusion that inoculation of cells into sand, even when they are provided with buffered standard medium, provokes a major reorganization in global gene expression patterns (one-third of all genes) involving numerous different cellular pathways (Table J in S2 File). This is similar to previous observations and conclusions using inoculations in sand with *S. wittichii* RW1 [11]. However, the types of cellular pathways seem to bear very little similarity and may thus be quite specific for different bacterial taxa. This can be seen from the actually very few common GO terms enriched in sand incubations with *P. veronii* 1YdBTEX2 and *S. wittichii* RW1, which are phylogenetically distant species (comparison in Table P in S2 File, individual *P. veronii* reaction in Table L in S2 File). However, the reactions of *P. veronii* 1YdBTEX2 common to *S. wittichii* RW1 in sand suggest that during the first contact hour to sand the strain is not actively growing but rather scavenges and readjusts its metabolism to the new environment, while inducing the pathways for toluene metabolism in its presence.

Transcriptomic data are not sufficient to ascertain the capacity of strains to survive inoculation transition periods and successfully colonize a new environment. Nevertheless, our results suggest that *P. veronii* 1YdBTEX2 adapts to a polluted environment (here: sand plus toluene) based on multiple adaptive functions encoded in its genome. Future studies should thus try to confirm that *P. veronii* 1YdBTEX2 can actually grow on BTEX upon inoculation in contaminated material, and characterize global gene expression during growth and survival. It would also be interesting to examine global gene expression in BTEX contaminated materials in which strain 1YdBTEX2 cannot grow, which could help to identify factors critical for its survival and activity in the field. We identified a variety of highly up-regulated genes coding for conserved hypothetical proteins, which would be interesting to investigate further to understand their role in adaptation. The general understanding of factors that contribute or enable better strain survival in non-sterile soils with native communities is still extremely limited, and cannot be suitably predicted from laboratory studies involving standardized growth conditions with chemical stresses, as we demonstrate here. Such knowledge, however, is important for current and future efforts in synthetic engineering of microbial communities.

Supporting Information

S1 File. Figure A. Comparison of the linearized *P. veronii* chromosome 1 replicon (Pv) with its close relatives *Pseudomonas* sp. TKP (Ptkp Acc. No. CP006852.1) and *P. trivialis* IHBB745 (Ptri, CP011507.1). Figure B. Prediction of Cluster of Orthologous Group classification of coding regions in the *P. veronii* 1YdBTEX2 genome distributed across the three replicons. Figure C. Inferred metabolic network for toluene and aromatic compound metabolism by *P. veronii* 1YdBTEX2, displayed using Cytoscape 3.3.0. Figure D. Network analysis of toluene and aromatic compound metabolism by *P. veronii* 1YdBTEX2 exposed in (A) liquid to toluene (Li-To) or in (B) sand to toluene (Sa-To). (DOCX)

S2 File. Table A. Putative intact and incomplete prophages found in the genome of *P. veronii* strain 1YdBTEX2 and their differential gene expression. Table B. Putative regions of genome

plasticity and genomic islands, and their differential gene expression. Table C. Differential expression of genes implicated in aerobic respiration and energy generation. Table D. Differential expression of genes for denitrification. Table E. Differential expression of genes associated with flagella, pilus, or chemotaxis. Table F. Differential expression of genes associated with secretion systems. Table G. Differential expression of genes for putative toxin/anti-toxin systems. Table H. Differential expression of genes for putative heavy metal resistance mechanisms. Table I. Differential expression of genes for putative aromatic hydrocarbon compound metabolism. Table J. Differential expression of genes in the genome of *P. veronii* 1YdBTEX2. Table K. Enriched GO terms among the significantly differentially expressed genes as response to change of carbon source. Table L. Enriched GO terms among the significantly differentially expressed genes as response to change of environment. Table M. Sorted genes significantly higher expressed in sand than in liquid. Table N. Sorted genes significantly lower expressed in sand than in liquid. Table O. Comparison of differentially expressed genes in sand versus under matric or solute stress. Table P. Common GO terms enriched in sand between *P. veronii* 1YdBTEX2 and *Sphingomonas wittichii* RW1. (XLSX)

Acknowledgments

The authors thank Dietmar Pieper for his kind gift of *Pseudomonas veronii* strain 1YdBTEX2.

Author Contributions

Conceptualization: MM VS CB MRR JvdM.

Data curation: MM CB VS JvdM.

Formal analysis: MM VS CB AK NKM AB KH FS FD DF M. Seppely ET PB AET CL DR M. Sulfiotti A. Etienne GRB LP A. Escoriza RM CS EA FBJ OS GR GG OE MD TP MRR JvdM.

Funding acquisition: JvdM.

Investigation: MM VS CB AK NKM AB KH FS FD DF M. Seppely ET PB AET CL DR M. Sulfiotti A. Etienne GRB LP A. Escoriza RM CS EA FBJ OS GR GG OE MD TP MRR JvdM.

Methodology: MM AK VS MD MRR JvdM.

Project administration: MRR JvdM.

Resources: BL AG.

Software: AK CB NKM BL AG TP MRR.

Supervision: MRR JvdM.

Writing – original draft: MM JvdM.

Writing – review & editing: MM CB VS AK FS GRB RM MRR JvdM.

References

1. Haritash AK, Kaushik CP. Biodegradation aspects of Polycyclic Aromatic Hydrocarbons (PAHs): A review. *J Hazard Mat.* 2009; 169(1–3):1–15. doi: <http://dx.doi.org/10.1016/j.jhazmat.2009.03.137>
2. Samanta SK, Singh OV, Jain RK. Polycyclic aromatic hydrocarbons: environmental pollution and bioremediation. *Trends Biotechnol.* 2002; 20(6):243–8. doi: [http://dx.doi.org/10.1016/S0167-7799\(02\)01943-1](http://dx.doi.org/10.1016/S0167-7799(02)01943-1) PMID: 12007492

3. Lan D, Liang B, Bao C, Ma M, Xu Y, Yu C. Marine oil spill risk mapping for accidental pollution and its application in a coastal city. *Marine Poll Bull.* 2015; 96(1–2):220–5. doi: <http://dx.doi.org/10.1016/j.marpolbul.2015.05.023>
4. Camilli R, Reddy CM, Yoerger DR, Van Mooy BAS, Jakuba MV, Kinsey JC, et al. Tracking hydrocarbon plume transport and biodegradation at Deepwater Horizon. *Science.* 2010; 330(6001):201–4. doi: [10.1126/science.1195223](https://doi.org/10.1126/science.1195223) PMID: [20724584](https://pubmed.ncbi.nlm.nih.gov/20724584/)
5. Megharaj M, Ramakrishnan B, Venkateswarlu K, Sethunathan N, Naidu R. Bioremediation approaches for organic pollutants: A critical perspective. *Environ Intern.* 2011; 37(8):1362–75. doi: <http://dx.doi.org/10.1016/j.envint.2011.06.003>
6. Mukherjee S, Juottonen H, Siivonen P, Lloret Quesada C, Tuomi P, Pulkkinen P, et al. Spatial patterns of microbial diversity and activity in an aged creosote-contaminated site. *ISME J.* 2014; 8(10):2131–42. doi: [10.1038/ismej.2014.151](https://doi.org/10.1038/ismej.2014.151) PMID: [25105905](https://pubmed.ncbi.nlm.nih.gov/25105905/)
7. Haghollahi A, Fazaelpoor MH, Schaffie M. The effect of soil type on the bioremediation of petroleum contaminated soils. *J Environ Man.* 2016; 180:197–201. doi: <http://dx.doi.org/10.1016/j.jenvman.2016.05.038>
8. Jeon CO, Madsen EL. In situ microbial metabolism of aromatic-hydrocarbon environmental pollutants. *Curr Opin Biotechnol.* 2013; 24(3):474–81. doi: [10.1016/j.copbio.2012.09.001](https://doi.org/10.1016/j.copbio.2012.09.001) PMID: [22999827](https://pubmed.ncbi.nlm.nih.gov/22999827/)
9. Mrozik A, Piotrowska-Seget Z. Bioaugmentation as a strategy for cleaning up of soils contaminated with aromatic compounds. *Microb Res.* 2010; 165(5):363–75. doi: <http://dx.doi.org/10.1016/j.micres.2009.08.001>
10. Cunliffe M, Kawasaki A, Fellows E, Kertesz MA. Effect of inoculum pretreatment on survival, activity and catabolic gene expression of *Sphingobium yanoikuyae* B1 in an aged polycyclic aromatic hydrocarbon-contaminated soil. *FEMS Microbiol Ecol.* 2006; 58(3):364–72. doi: [10.1111/j.1574-6941.2006.00167.x](https://doi.org/10.1111/j.1574-6941.2006.00167.x) PMID: [17117981](https://pubmed.ncbi.nlm.nih.gov/17117981/)
11. Moreno-Forero SK, van der Meer JR. Genome-wide analysis of *Sphingomonas wittichii* RW1 behaviour during inoculation and growth in contaminated sand. *ISME J.* 2015; 9:150–65. Epub 2014/06/18. doi: [10.1038/ismej.2014.101](https://doi.org/10.1038/ismej.2014.101) PMID: [24936762](https://pubmed.ncbi.nlm.nih.gov/24936762/)
12. Singer AC, van der Gast CJ, Thompson IP. Perspectives and vision for strain selection in bioaugmentation. *Trends Biotechnol.* 2005; 23(2):74–7. doi: [10.1016/j.tibtech.2004.12.012](https://doi.org/10.1016/j.tibtech.2004.12.012) PMID: [15661343](https://pubmed.ncbi.nlm.nih.gov/15661343/)
13. Backman A, Maraha N, Jansson JK. Impact of temperature on the physiological status of a potential bioremediation inoculant, *Arthrobacter chlorophenolicus* A6. *Appl Environ Microbiol.* 2004; 70(5):2952–8. doi: [10.1128/aem.70.5.2952–2958.2004](https://doi.org/10.1128/aem.70.5.2952-2958.2004) PMID: [15128556](https://pubmed.ncbi.nlm.nih.gov/15128556/)
14. Coronado E, Roggo C, Johnson DR, van der Meer JR. Genome-wide analysis of salicylate and dibenzofuran metabolism in *Sphingomonas wittichii* RW1. *Front Microbiol.* 2012; 3:300. Epub 2012/09/01. doi: [10.3389/fmicb.2012.00300](https://doi.org/10.3389/fmicb.2012.00300) PMID: [22936930](https://pubmed.ncbi.nlm.nih.gov/22936930/)
15. Moreno-Forero SK, Rojas E, Beggah S, van der Meer JR. Comparison of differential gene expression to water stress among bacteria with relevant pollutant-degradation properties. *Environ Microbiol Rep.* 2016; 8(1):91–102. doi: [10.1111/1758-2229.12356](https://doi.org/10.1111/1758-2229.12356) PMID: [26616826](https://pubmed.ncbi.nlm.nih.gov/26616826/)
16. Junca H, Pieper DH. Functional gene diversity analysis in BTEX contaminated soils by means of PCR-SSCP DNA fingerprinting: comparative diversity assessment against bacterial isolates and PCR-DNA clone libraries. *Environ Microbiol.* 2004; 6(2):95–110. doi: [10.1046/j.1462-2920.2003.00541.x](https://doi.org/10.1046/j.1462-2920.2003.00541.x) PMID: [14756875](https://pubmed.ncbi.nlm.nih.gov/14756875/)
17. Rojas A, Duque E, Mosqueda G, Golden G, Hurtado A, Ramos JL, et al. Three efflux pumps are required to provide efficient tolerance to toluene in *Pseudomonas putida* DOT-T1E. *J Bacteriol.* 2001; 183(13):3967–73. Epub 2001/06/08. doi: [10.1128/JB.183.13.3967–3973.2001](https://doi.org/10.1128/JB.183.13.3967-3973.2001) PMID: [11395460](https://pubmed.ncbi.nlm.nih.gov/11395460/)
18. de Lima-Morales D, Chaves-Moreno D, Jarek M, Vilchez-Vargas R, Jauregui R, Pieper DH. Draft genome sequence of *Pseudomonas veronii* strain 1YdBTEX2. *Genome announcements.* 2013; 1(3). Epub 2013/05/18. doi: [10.1128/genomeA.00258-13](https://doi.org/10.1128/genomeA.00258-13) PMID: [23682152](https://pubmed.ncbi.nlm.nih.gov/23682152/)
19. Bertelli C, Aeby S, Chassot B, Clulow J, Hilfiker O, Rappo S, et al. Sequencing and characterizing the genome of *Estrella lausannensis* as an undergraduate project: training students and biological insights. *Front Microbiol.* 2015; 6:101. doi: [10.3389/fmicb.2015.00101](https://doi.org/10.3389/fmicb.2015.00101) PMID: [25745418](https://pubmed.ncbi.nlm.nih.gov/25745418/)
20. Gerhardt P, Murray R, Costilow R, Nester E, Wood W, Krieg N, et al. *Manual of methods for general bacteriology.* Washington: American Society for Microbiology; 1981.
21. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Meth.* 2013; 10(6):563–9. doi: [10.1038/nmeth.2474](https://doi.org/10.1038/nmeth.2474) <http://www.nature.com/nmeth/journal/v10/n6/abs/nmeth.2474.html>—supplementary-information. PMID: [23644548](https://pubmed.ncbi.nlm.nih.gov/23644548/)
22. Darling AE, Miklós I, Ragan MA. Dynamics of genome rearrangement in bacterial populations. *PLoS Genet.* 2008; 4(7):e1000128. doi: [10.1371/journal.pgen.1000128](https://doi.org/10.1371/journal.pgen.1000128) PMID: [18650965](https://pubmed.ncbi.nlm.nih.gov/18650965/)

23. Gordon D, Green P. Consed: a graphical editor for next-generation sequencing. *Bioinformatics*. 2013; 29(22):2936–7. doi: [10.1093/bioinformatics/btt515](https://doi.org/10.1093/bioinformatics/btt515) PMID: [23995391](https://pubmed.ncbi.nlm.nih.gov/23995391/)
24. Meyer F, Goesmann A, McHardy AC, Bartels D, Bekel T, Clausen J, et al. GenDB—an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res*. 2003; 31(8):2187–95. Epub 2003/04/12. PMID: [12682369](https://pubmed.ncbi.nlm.nih.gov/12682369/)
25. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform*. 2010; 11:119-. doi: [10.1186/1471-2105-11-119](https://doi.org/10.1186/1471-2105-11-119) PMID: [20211023](https://pubmed.ncbi.nlm.nih.gov/20211023/)
26. Consortium TU. UniProt: a hub for protein information. *Nucl Acids Res*. 2015; 43(D1):D204–D12. doi: [10.1093/nar/gku989](https://doi.org/10.1093/nar/gku989) PMID: [25348405](https://pubmed.ncbi.nlm.nih.gov/25348405/)
27. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990; 215(3):403–10. doi: [10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) PMID: [2231712](https://pubmed.ncbi.nlm.nih.gov/2231712/)
28. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucl Acids Res*. 2016; 44(Database issue):D457–D62. doi: [10.1093/nar/gkv1070](https://doi.org/10.1093/nar/gkv1070) PMID: [26476454](https://pubmed.ncbi.nlm.nih.gov/26476454/)
29. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucl Acids Res*. 1999; 27(1):29–34. PMID: [9847135](https://pubmed.ncbi.nlm.nih.gov/9847135/)
30. Zhou Y, Liang Y, Lynch K, Dennis JJ, Wishart DS. PHAST: A fast phase search tool *Nucl Acids Res*. 2011; 39:347–52.
31. Dhillon BK, Laird MR, Shay JA, Winsor GL, Lo R, Nizam F, et al. IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis. *Nucleic Acids Res*. 2015; 43(W1):W104–8. doi: [10.1093/nar/gkv401](https://doi.org/10.1093/nar/gkv401) PMID: [25916842](https://pubmed.ncbi.nlm.nih.gov/25916842/)
32. Sevin EW, Barloy-Hubler F. RASTA-Bacteria: a web-based tool for identifying toxin-antitoxin loci in prokaryotes. *Genome Biol*. 2007; 8(8):R155–R. doi: [10.1186/gb-2007-8-8-r155](https://doi.org/10.1186/gb-2007-8-8-r155) PMID: [17678530](https://pubmed.ncbi.nlm.nih.gov/17678530/)
33. Shao Y, Harrison EM, Bi D, Tai C, He X, Ou HY, et al. TADB: a web-based resource for Type 2 toxin-antitoxin loci in bacteria and archaea. *Nucl Acids Res*. 2011; 39(Database issue):D606–11. Epub 2010/10/12. doi: [10.1093/nar/gkq908](https://doi.org/10.1093/nar/gkq908) PMID: [20929871](https://pubmed.ncbi.nlm.nih.gov/20929871/)
34. Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Pillay M, et al. IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res*. 2014; 42(Database issue):D560–7. doi: [10.1093/nar/gkt963](https://doi.org/10.1093/nar/gkt963) PMID: [24165883](https://pubmed.ncbi.nlm.nih.gov/24165883/)
35. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Meth*. 2012; 9(4):357–9. doi: [10.1038/nmeth.1923](https://doi.org/10.1038/nmeth.1923) PMID: [22388286](https://pubmed.ncbi.nlm.nih.gov/22388286/)
36. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; 25(16):2078–9. doi: [10.1093/bioinformatics/btp352](https://doi.org/10.1093/bioinformatics/btp352) PMID: [19505943](https://pubmed.ncbi.nlm.nih.gov/19505943/)
37. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015; 31(2):166–9. doi: [10.1093/bioinformatics/btu638](https://doi.org/10.1093/bioinformatics/btu638) PMID: [25260700](https://pubmed.ncbi.nlm.nih.gov/25260700/)
38. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010; 26(1):139–40. doi: [10.1093/bioinformatics/btp616](https://doi.org/10.1093/bioinformatics/btp616) PMID: [19910308](https://pubmed.ncbi.nlm.nih.gov/19910308/)
39. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000; 25(1):25–9. Epub 2000/05/10. doi: [10.1038/75556](https://doi.org/10.1038/75556) PMID: [10802651](https://pubmed.ncbi.nlm.nih.gov/10802651/)
40. Conesa A, Gotz S. Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int J Plant Genomics*. 2008; 2008:619832. doi: [10.1155/2008/619832](https://doi.org/10.1155/2008/619832) PMID: [18483572](https://pubmed.ncbi.nlm.nih.gov/18483572/)
41. Zheng Q, Wang XJ. GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis. *Nucl Acids Res*. 2008; 36(Web Server issue):W358–63. doi: [10.1093/nar/gkn276](https://doi.org/10.1093/nar/gkn276) PMID: [18487275](https://pubmed.ncbi.nlm.nih.gov/18487275/)
42. Alexa A, Rahnenfuhrer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*. 2006; 22(13):1600–7. doi: [10.1093/bioinformatics/btl140](https://doi.org/10.1093/bioinformatics/btl140) PMID: [16606683](https://pubmed.ncbi.nlm.nih.gov/16606683/)
43. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, et al. Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc*. 2007; 2(10):2366–82. doi: [10.1038/nprot.2007.324](https://doi.org/10.1038/nprot.2007.324) PMID: [17947979](https://pubmed.ncbi.nlm.nih.gov/17947979/)
44. Martinez-Garcia E, de Lorenzo V. Engineering multiple genomic deletions in Gram-negative bacteria: analysis of the multi-resistant antibiotic profile of *Pseudomonas putida* KT2440. *Environ Microbiol*. 2011; 13(10):2702–16. Epub 2011/09/03. doi: [10.1111/j.1462-2920.2011.02538.x](https://doi.org/10.1111/j.1462-2920.2011.02538.x) PMID: [21883790](https://pubmed.ncbi.nlm.nih.gov/21883790/)
45. Miyazaki R, Minoia M, Pradervand N, Sulser S, Reinhard F, van der Meer JR. Cellular variability of RpoS expression underlies subpopulation activation of an integrative and conjugative element. *PLoS*

- Genet. 2012; 8(7):e1002818. Epub 2012/07/19. doi: [10.1371/journal.pgen.1002818](https://doi.org/10.1371/journal.pgen.1002818) PGENETICS-D-12-00309 [pii]. PMID: [22807690](https://pubmed.ncbi.nlm.nih.gov/22807690/)
46. Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics*. 2009; 25(1):119–20. doi: [10.1093/bioinformatics/btn578](https://doi.org/10.1093/bioinformatics/btn578) PMID: [18990721](https://pubmed.ncbi.nlm.nih.gov/18990721/)
 47. Guy L, Kultima JR, Andersson SG. genoPlotR: comparative gene and genome visualization in R. *Bioinformatics*. 2010; 26(18):2334–5. doi: [10.1093/bioinformatics/btq413](https://doi.org/10.1093/bioinformatics/btq413) PMID: [20624783](https://pubmed.ncbi.nlm.nih.gov/20624783/)
 48. Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW. Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev*. 2009; 33(2):376–93. doi: [10.1111/j.1574-6976.2008.00136.x](https://doi.org/10.1111/j.1574-6976.2008.00136.x) PMID: [19178566](https://pubmed.ncbi.nlm.nih.gov/19178566/)
 49. de Lima-Morales D, Chaves-Moreno D, Wos-Oxley ML, Jáuregui R, Vilchez-Vargas R, Pieper DH. Degradation of benzene by *Pseudomonas veronii* 1YdBTEX2 and 1YB2 is catalyzed by enzymes encoded in distinct catabolism gene clusters. *Appl Environ Microbiol*. 2016; 82(1):167–73. doi: [10.1128/aem.03026-15](https://doi.org/10.1128/aem.03026-15) PMID: [26475106](https://pubmed.ncbi.nlm.nih.gov/26475106/)
 50. Witzig R, Junca H, Hecht H-J, Pieper DH. Assessment of toluene/biphenyl dioxygenase gene diversity in benzene-polluted soils: links between benzene biodegradation and genes similar to those encoding isopropylbenzene dioxygenases. *Appl Environ Microbiol*. 2006; 72(5):3504–14. doi: [10.1128/aem.72.5.3504-3514.2006](https://doi.org/10.1128/aem.72.5.3504-3514.2006) PMID: [16672497](https://pubmed.ncbi.nlm.nih.gov/16672497/)
 51. Parales RE, Ditty JL, Harwood CS. Toluene-degrading bacteria are chemotactic towards the environmental pollutants benzene, toluene, and trichloroethylene. *Appl Environ Microbiol*. 2000; 66(9):4098–104. PMID: [10966434](https://pubmed.ncbi.nlm.nih.gov/10966434/)
 52. Ballal A, Basu B, Apte SK. The Kdp-ATPase system and its regulation. *J Biosc*. 2007; 32(3):559–68. doi: [10.1007/s12038-007-0055-7](https://doi.org/10.1007/s12038-007-0055-7)
 53. Malli R, Epstein W. Expression of the Kdp ATPase is consistent with regulation by turgor pressure. *J Bacteriol*. 1998; 180(19):5102–8. PMID: [9748442](https://pubmed.ncbi.nlm.nih.gov/9748442/)
 54. Fivian-Hughes AS, Davis EO. Analyzing the regulatory role of the HigA antitoxin within *Mycobacterium tuberculosis*. *J Bacteriol*. 2010; 192(17):4348–56. doi: [10.1128/jb.00454-10](https://doi.org/10.1128/jb.00454-10) PMID: [20585061](https://pubmed.ncbi.nlm.nih.gov/20585061/)
 55. Martinez-Garcia E, Nikel PI, Chavarria M, de Lorenzo V. The metabolic cost of flagellar motion in *Pseudomonas putida* KT2440. *Environ Microbiol*. 2014; 16(1):291–303. doi: [10.1111/1462-2920.12309](https://doi.org/10.1111/1462-2920.12309) PMID: [24148021](https://pubmed.ncbi.nlm.nih.gov/24148021/)
 56. He Z, Zhou A, Baidoo E, He Q, Joachimiak MP, Benke P, et al. Global transcriptional, physiological, and metabolite analyses of the responses of *Desulfovibrio vulgaris* Hildenborough to salt adaptation. *Appl Environ Microbiol*. 2010; 76(5):1574–86. doi: [10.1128/aem.02141-09](https://doi.org/10.1128/aem.02141-09) PMID: [20038696](https://pubmed.ncbi.nlm.nih.gov/20038696/)
 57. Chang W-S, van de Mortel M, Nielsen L, Nino de Guzman G, Li X, Halverson LJ. Alginate production by *Pseudomonas putida* creates a hydrated microenvironment and contributes to biofilm architecture and stress tolerance under water-limiting conditions. *J Bacteriol*. 2007; 189(22):8290–9. doi: [10.1128/jb.00727-07](https://doi.org/10.1128/jb.00727-07) PMID: [17601783](https://pubmed.ncbi.nlm.nih.gov/17601783/)
 58. Csonka LN. Physiological and genetic responses of bacteria to osmotic stress. *Microbiol Rev*. 1989; 53(1):121–47. PMID: [2651863](https://pubmed.ncbi.nlm.nih.gov/2651863/)
 59. Freeman BC, Chen C, Yu X, Nielsen L, Peterson K, Beattie GA. Physiological and transcriptional responses to osmotic stress of two *Pseudomonas syringae* strains that differ in their epiphytic fitness and osmotolerance. *J Bacteriol*. 2013. doi: [10.1128/jb.00787-13](https://doi.org/10.1128/jb.00787-13) PMID: [23955010](https://pubmed.ncbi.nlm.nih.gov/23955010/)

CHAPTER 2

Genome-wide gene expression changes of *Pseudomonas veronii* 1YdBTEX2 during bioaugmentation in polluted soils

Marian Morales, Vladimir Sentchilo, Noushin Hadadi and Jan Roelof van der Meer

Department of Fundamental Microbiology, University of Lausanne, 1015 Lausanne, Switzerland

Previously published in *Environmental Microbiome* 16,8 (2021) <https://doi.org/10.1186/s40793-021-00378-x>

Author Contributions

M. M. and V. S. carried out experimental work. M. M., N. H. and J.M. conducted data analysis. M. M. and J. M. wrote the manuscript. M. M., N. H., V. S. and J. M. commented on the final text. J. M. raised funding. The author(s) read and approved the final manuscript.

RESEARCH ARTICLE

Open Access



Genome-wide gene expression changes of *Pseudomonas veronii* 1YdBTEX2 during bioaugmentation in polluted soils

Marian Morales, Vladimir Sentchilo, Noushin Hadadi and Jan Roelof van der Meer*

Abstract

Background: Bioaugmentation aims to use the capacities of specific bacterial strains inoculated into sites to enhance pollutant biodegradation. Bioaugmentation results have been mixed, which has been attributed to poor inoculant growth and survival in the field, and, consequently, moderate catalytic performance. However, our understanding of biodegradation activity mostly comes from experiments conducted under laboratory conditions, and the processes occurring during adaptation and invasion of inoculants into complex environmental microbiomes remain poorly known. The main aim of this work was thus to study the specific and different cellular reactions of an inoculant for bioaugmentation during adaptation, growth and survival in natural clean and contaminated non-sterile soils, in order to better understand factors limiting bioaugmentation.

Results: As inoculant we focused on the monoaromatic compound-degrading bacterium *Pseudomonas veronii* 1YdBTEX2. The strain proliferated in all but one soil types in presence and in absence of exogenously added toluene. RNAseq and differential genome-wide gene expression analysis illustrated both a range of common soil responses such as increased nutrient scavenging and recycling, expression of defense mechanisms, as well as environment-specific reactions, notably osmoprotection and metal homeostasis. The core metabolism of *P. veronii* remained remarkably constant during exponential growth irrespective of the environment, with slight changes in cofactor regeneration pathways, possibly needed for balancing defense reactions.

Conclusions: *P. veronii* displayed a versatile global program, enabling it to adapt to a variety of soil environments in the presence and even in absence of its target pollutant toluene. Our results thus challenge the widely perceived dogma of poor survival and growth of exogenous inoculants in complex microbial ecosystems such as soil and provide a further basis to developing successful bioaugmentation strategies.

Keywords: RNAseq, Inoculation, Invasion, Survival

Background

The recovery and restoration of soils polluted by organic compounds may be enhanced by introducing specific biodegrader bacteria [1–4]. This process, named bioaugmentation, relies on individual or mixtures of inoculant strains with specific metabolic pathways capable to degrade

particular organic compounds, to invade, survive and propagate in the contaminated environment at the expense of the degraded pollutant [2, 5–8]. Inoculation attempts frequently do not achieve the intended success and the inoculated strains either do not survive and multiply, or do not display their catabolic properties [9–11]. Whereas most studies have addressed very practical aspects of improving bioaugmentation success, e.g., through strain formulations [3, 12] or process management [1], there is a basic lack of understanding of the factors

* Correspondence: janroelof.vandermeer@unil.ch

Department of Fundamental Microbiology, University of Lausanne, 1015 Lausanne, Switzerland



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

determining establishment, growth and survival of exogenous strains inside existing microbial ecosystems. Bioaugmentation in that respect is similar to, for example, application of probiotic bacteria in gut systems [13]. And, in a way, the strategies for controlled species growth may relate to those deployed by many pathogenic bacteria to invade native microbiota [14]. Whereas the questions of controlled invasion are old [15], we believe one can learn more about the process and its limitations from the bacteria themselves, how they perceive the transition from sterile culture medium into non-sterile contaminated sites, and which factors they express specifically during growth and maintenance. Such knowledge may help to define specific process conditions favoring controlled growth within existing communities, and increase future success of bioaugmentation.

Previously, we suggested that more attention should be given to the combination of factors representing the actual expected conditions at polluted sites [16], and the molecular and functional strategies displayed by biodegrader bacteria during inoculation under near-field conditions. For example, Moreno et al., [17] found that one-third of the genes of the dibenzofuran-degrading bacterium *Sphingomonas wittichii* RW1 express differently during transition and growth in (polluted) sand as compared to liquid growth with the same carbon substrate [17]. A complementary study done on the same strain using transposon scanning revealed a wide range of genes with selective effects under “soil-specific” conditions [18]. This suggested the existence and importance of many functions specific for survival in soil conditions, although their general nature has remained unclear [18]. In order to understand whether global adaptive responses to soil environments are conserved, we compared transcriptomic changes in *S. wittichii* with those of *Pseudomonas veronii* strain 1YdBTEX2, isolated from sites contaminated with aromatic compounds [19], during the transition from liquid growth media to sandy soil [20]. Similarly to *S. wittichii* RW1, inoculation into sandy soil provoked a major reorganization in global gene expression of *P. veronii* 1YdBTEX2, implicating one-third of all genes, but with very few pathways and biological processes in common between the two [20].

To complement previous studies on *P. veronii* 1YdBTEX2 gene expression changes during adaptation to a soil environment we focus here on its growth and establishment. The questions we aimed to address here were whether *P. veronii* would be capable to grow and survive in natural non-sterile soils, and whether this would be dependent on the presence of contaminated material or specific added carbon substrate (toluene). We further aimed to understand whether *P. veronii* would react differently in response to soil type and whether its cellular reactions could be revealed from global gene expression

changes. In our experimental design, we inoculated *P. veronii* and measured population growth in three different natural non-sterile soils, in comparison to regular liquid culture and inert silica matrix, in the presence or absence of externally added toluene or in historically contaminated material with (polycyclic) aromatic hydrocarbons. Community RNA was extracted in exponential phase of *P. veronii* growth and in stationary phase, which was reverse-transcribed, ribosomal RNA depleted, and sequenced (RNA-seq). Adaptation and cellular reactions of *P. veronii* under different conditions were interpreted from a combination of global tools, including gene ontology (GO) [21] and cluster of orthologous groups (COG) assignments of biological processes and pathways [22], a previously constructed genome-scale model of *P. veronii* 1YdBTEX2 (iPsvr) [23], as well as from detailed individual gene or operon annotation information. Transcriptome analysis pointed to a variety of common and specific adaptations in soil environments, but a surprisingly conserved core metabolic expression in exponential phase irrespective of the growth environment.

Methods

Media and general culturing conditions of *P. veronii*

P. veronii is a toluene, benzene and *m*- and *p*-xylene degrading bacterium originating from contaminated soil [19]. To more easily count *P. veronii* colony-forming units (CFUs) in soil, we used a derivative with a mini-Tn5 insertion constitutively expressing the green fluorescent protein (GFP) from the P_{circ} promoter of the ICE $_{clc}$ element [24]. We found no effect on toluene growth in liquid medium between wild-type *P. veronii* and the tagged strain (laboratory strain number 3381, Supplementary Figure 1). For all transcriptomic experiments we used the untagged wild-type strain (lab strain number 3371). *P. veronii* was plated freshly for each experiment from a -80°C glycerol stock on nutrient agar (Oxoid), and was grown for 48 h at 30°C . *P. veronii* colonies were restreaked on 21C-type minimal medium (MM) [25] agar plates, which were incubated in a 10-L closed jar for 72 h at room temperature with 0.5 ml pure toluene as sole carbon and energy source dosed through the vapor phase from an open tube placed in the jar. Liquid suspended precultures were prepared in MM with 10 mM succinate, starting from a single toluene-grown colony, which was incubated for 24 h at 30°C with rotatory shaking at 180 rpm. In the case of liquid growth experiments, freshly grown precultures were diluted 1:100 (v/v) in MM with 10 mM succinate and again incubated as before. For soil growth experiments, cells were recovered from exponentially growing cultures on succinate (at culture turbidity $\text{OD}_{600} = 0.4$) by centrifugation ($4000\times g$ for 5 min at room temperature), resuspended in MM without further carbon source and then transferred

to the soil microcosms (see below). Growth of *P. veronii* was determined by CFU counting of appropriate dilutions on MM plates with toluene (gas phase). *P. veronii* colonies were differentiated from any background soil bacterial colonies growing on MM plates with toluene by their green fluorescence, and counted under a digital dual band microscope (Dino-Lite model AM4115T-GRFBY) using the 485 nm blue LED excitation and 510 nm emission filter.

Soil types and contaminated material

Three natural soils were used for microcosm growth studies with *P. veronii*. These consisted of (i) a sandy soil sampled in a lake Geneva beach near in St. Sulpice – named *Sand*, (ii) a silty soil sampled from the bank of the local stream ‘Sorge’ on the university campus – named *Silt*, and (iii) a clay soil sampled in a forest area on campus near to Lake Geneva – named *Clay*. Quantities of ~5 kg were spread on aluminium foil and air-dried on the laboratory bench. Sand and Silt were dried for 7 days (further water losses were not evident with longer drying periods), whereas Clay was dried for 15 days. At the moment of microcosm inoculations, the Sand, Silt, and Clay had gravimetric water content (GWC) of $0.14 \pm 0.01\%$, $0.20 \pm 0.01\%$, and $2.22 \pm 0.06\%$, respectively. pH-H₂O of the materials was 7.14 ± 0.02 (Sand), 8.57 ± 0.2 (Silt) and 7.78 ± 0.02 (Clay). Total organic matter content amounted to 0.028 (Sand), 0.13% (Silt) and 4.0% (Clay). Total cell counts on washed material and stained by SYBR Green I were quantified according to Weinbauer et al. [26]. The effect of drying on viable bacteria was determined with freshly sampled Sand dried for 3 h and for 5 days at ambient air temperature (20 °C, Supplementary Figure 2).

The polluted material (Junction) originated from a former gasification work in Geneva and was obtained through collaboration with Biotech S.A., Geneva (CH). According to previous characterization [27], the material contained primarily alkanes (C₁₀ – C₄₀; 9500 mg kg⁻¹) and polycyclic aromatic hydrocarbons (2300 mg kg⁻¹). In addition, minor concentrations of benzene (3.3 mg kg⁻¹) and methylated monoaromatic compounds were found (22 mg kg⁻¹). The original material consisted mainly of gravel (1–5 cm) covered with a layer of tar-containing mud. In order to obtain more easily handleable material, it was mixed 65/35 (w/w) in small portions with air-dried Silt and sieved through 3-mm diameter to remove the gravel. In the main text we refer to this mixed and sieved material as *Junction*. At the moment of inoculation, Junction had a GWC of $6.55 \pm 0.15\%$.

Growth kinetics estimations

The kinetics of *P. veronii* miniTn5::gfp population growth in soil was assessed in microcosms artificially contaminated with toluene. Four replicate microcosms were

prepared, each consisting of 95 g of dried Sand, Silt or Clay inside 500-ml glass Schott bottles and closed with Teflon-lined screw caps. As control for porosity effects, we used autoclaved quartz at 5% GWC (silica crystals, 50–70 mesh particle size, Sigma-Aldrich ref. 274,739); hereafter referred to as artificial porous matrix or APM.

Toluene was dosed through the vapor phase from a sealed 1-ml micropipette tip, placed inside the glass bottle and containing 0.2 ml of pure toluene. Tubes with toluene were removed before inoculation with 5 ml suspension of washed preculture of *P. veronii* to obtain 2.5×10^4 CFU g⁻¹ material at the start. Microcosm flasks were mixed on a horizontal roller mixer (IKA roller 6 digital) at 80 rpm for 30 min, with manual shaking every 5 min to detach the soil from the walls. After mixing, the tip containing the toluene was placed back inside and the bottles were incubated upright at 24–26 °C in the dark for 60–120 h with regular sampling (see below).

Growth of *P. veronii* miniTn5::gfp in Junction was followed in (triplicate) microcosms of 10 g in 50 ml polypropylene screw-cap tubes (Greiner AG, cat #227261) containing either 100% Junction (see above), 100% Silt, or mixtures of 75% (g g⁻¹), 50, and 25% Junction with Silt. Microcosms were again inoculated with washed *P. veronii* preculture (0.5 ml) to achieve 2.5×10^4 CFU g⁻¹ at the start. Microcosms were homogenized and incubated as above, either without any further addition of carbon, or amended with a sealed 1 ml micropipette tip containing 20 µl pure toluene (as described above).

Growth of *P. veronii* on toluene in liquid suspended culture was measured from four replicate 100-mL screw-cap conical flasks containing 15 ml MM and starting cell concentrations of $\sim 1 \times 10^7$ CFU ml⁻¹. As we found in preliminary experiments that *P. veronii* growth on toluene in liquid culture was less consistent with gas-phase dosing, we deployed an inert oil-toluene mixture instead. For this, we mixed 1:25 (v/v) of toluene:tetradecane ($\geq 99.9\%$; Sigma-Aldrich ref.: 34866; Aldrich ref.: 87140) and added 0.5 ml per 15 ml liquid culture. Flasks were incubated at 30 °C and at 180 rpm on an orbital shaker.

Soil and liquid microcosms were sampled directly after inoculation (1 h) and then four times per day at approximately 6 h intervals (or once, for the Junction series). Samples of 5 g (or 2.5 g for Junction) were retrieved from the microcosms with a spatula and transferred to clean 50 ml polypropylene tubes. 10 ml of sterile saline solution (0.9% NaCl) was added to each tube and cells were extracted by vortexing for 1 min. Larger soil particles were allowed to sediment for a few seconds, after which the supernatant was transferred to a fresh 50 ml tube and serially diluted in sterile saline. For liquid cultures, a 1-mL aliquot was taken directly from the flask and serially diluted. Serial dilutions were drop-plated (10 × 5 µl drops of each dilution) on MM agar, which was

incubated with toluene vapor to quantify the number of *P. veronii* CFU as described above. Identities of *P. veronii* miniTn5::gfp colonies were verified by their GFP fluorescence.

Growth rates (μ_{\max}) of *P. veronii* were calculated from the slope of the mean \log_{10} CFU g^{-1} over time across quadruplicate assays (or triplicates, for Jonction incubations). We further refer to the maximum population size (max pop size) as the highest mean CFU g^{-1} or ml^{-1} observed during the entire experiment for each condition, averaged from four (soils) or three (Jonction) replicates at the same sampling time point. Growth rates were compared among treatments by ANOVA, followed by post-hoc Tukey testing.

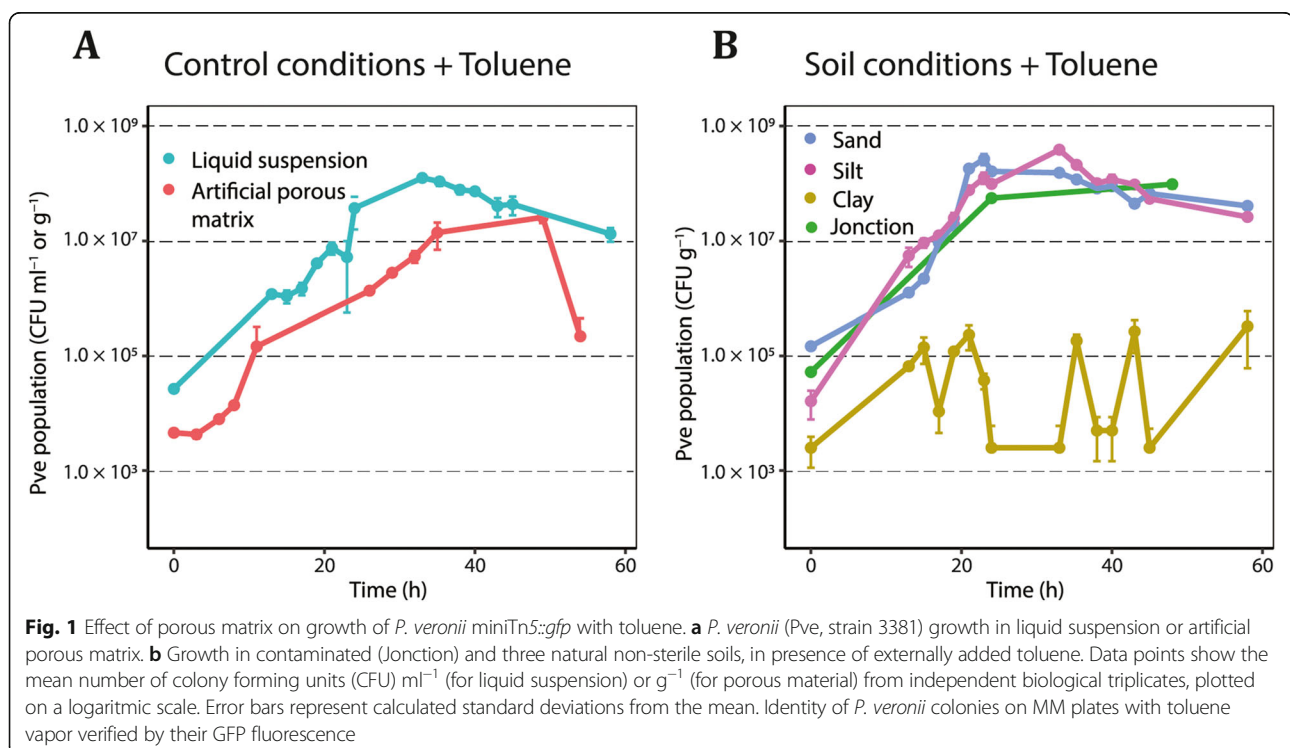
Genome-wide gene expression analysis

For genome-wide expression analysis we used *P. veronii* wild-type, inoculated in the different materials or conditions and sampled after 1 h (transition phase or LAG), during estimated exponential (EXPO) or in stationary (STAT) phase.

For the transition phase we inoculated 5 ml (2.5 ml for Jonction) of a suspension of $\sim 10^7$ ml^{-1} *P. veronii* cells into soil microcosms with 95 g material (47.5 g for Jonction material) in 500 ml capped-glass Schott bottles, with toluene dosage through the gas phase. Cells were prepared from exponentially growing liquid cultures as described above, but resuspended in MM with 0.5 mM succinate to avoid starvation [20]. We used a higher

starting cell density for the transition phase than in the exponential and stationary phases described below, in order to be able to extract sufficient amounts of RNA. As materials we tested here Sand, Silt, Clay, Jonction, APM and liquid suspended growth (LQ). All treatments were started in quadruplicates and incubated at room temperature (24–26 °C) without agitation for 1 h, after which 10 g were sampled from each replicate.

P. veronii during exponential growth and stationary phase was sampled from similar inoculated microcosms (95 g material or 47.5 g for Jonction), but in this case with a low starting cell density of $\sim 10^4$ *P. veronii* CFU g^{-1} or ml^{-1} to achieve sufficient growth in the soils. Cells were precultured as before, but washed and resuspended in MM without succinate. Microcosms were dosed with toluene as before (four replicates each), and sampled (10 g) at approximate exponential and stationary phase for *P. veronii* based on initial growth experiments as in Fig. 1. As materials we tested here Silt, Jonction and APM. For the liquid suspension control, *P. veronii* was inoculated into 100 ml screw-cap conical flasks containing 25 ml of MM amended with 0.5 ml of a 1:19 mixture of toluene:tetradecane to obtain a starting OD₆₀₀ of 0.16 (four replicates). Flasks were incubated at 30 °C and 180 rpm on an orbital shaker. Cells were sampled after 4 h (OD = 0.5, 2 ml) and 24 h (OD = 1.8–1.9, 1 ml), and harvested by centrifugation at 3500×g, for 6 min at 30 °C. Pellets were snap-frozen in liquid nitrogen and stored at –80 °C until RNA extraction.



In order to improve yields for RNA extraction and limit humic acid interference, we first washed cells from solid samples of 10 g by adding 20 ml of 4 °C-cold sterile saline solution (0.9% NaCl) to the tube (50 ml volume) and vortexing for 2 min. Suspensions were allowed to sediment for 1 min, after which the cleared liquid was transferred into a clean 50 ml tube, which was centrifuged for 4 min at 4000×g at 4 °C to collect the cells. After centrifugation, the supernatant was immediately discarded by inversion, the excess of liquid from the walls was quickly removed with a clean paper towel, and the cell pellet was snap-frozen in liquid nitrogen and stored at –80 °C until RNA isolation. Estimated cell recoveries from the washing procedure are reported in Supplementary Figure 3.

RNA isolation, purification and library sequencing

RNA was extracted from frozen cell samples using the RNA PowerSoil Total RNA Isolation Kit (MoBio Laboratories). Cells in thawed samples were disrupted in a bead-beating protocol as recommended by the manufacturer (MoBio Laboratories). RNA was purified following MoBio procedures and the final RNA pellet was resuspended in a volume of 20 µl of RNase-free water. Contaminating DNA was removed by two consecutive treatments with TURBO DNase (Ambion), followed by purification using an RNeasy MinElute Cleanup kit (QIAGEN, Valencia, CA, USA).

RNA quantity (reported in Supplementary Table 1) and quality was verified by reading the 260 nm absorbance and the absorbance ratios at 260/280 nm and 260/230 nm on a NanoDrop spectrophotometer (ThermoFisher Scientific). RNA was migrated on an Agilent 2100 Bioanalyser (Agilent Technologies) to verify the presence of intact 16S and 23S rRNA. Genomic DNA contamination was checked by PCR using specific primers for a unique region in the *P. veronii* genome. RNA samples were then depleted from ribosomal RNAs by using the Ribo-Zero rRNA Removal Kit Bacteria protocol (Epicentre, Madison, WI, USA). Subsequently, the RNA was reverse-transcribed, indexed and amplified by PCR using the ScriptSeq™ v2 Bacteria kit and ScriptSeq™ Index PCR primers set 1 (Epicentre). The resulting directional RNAseq libraries were sequenced using single 100-nt read chemistry on an Illumina HiSeq 2500 platform (Illumina, Inc., San Diego, USA) at the Lausanne Genomic Technologies Facility.

Bioinformatic analysis and statistics

Raw reads were quality-filtered, mapped, sorted and indexed with Bowtie2 [28] and Samtools [29] under default settings, using the finalized gapless *P. veronii* genome sequence as reference (European Nucleotide Archive under bioproject number PRJEB11417). A summary of the total number of mapped reads per condition is listed in Supplementary Table 1. Mapped reads

passing default alignment values were counted with HTSeq [30], then further processed and analysed with edgeR [31]. Only reads counted more than once per million in at least three replicates were kept. Counts were normalised across samples and compared between pair-wise conditions in a modified Fisher's exact test (as implemented in edgeR). ANOVA (also as implemented within edgeR) was used to detect differential gene expression with interpretation groups “natural soils – 1h” (Sand, Silt and Clay) and “controls – 1h” (Liquid and APM) or “polluted soil – 1h” (Jonction) and “controls – 1h”. Comparative data of *P. veronii* transition response (1 h) in liquid suspension and in Sand were taken from previous work [20].

Genes were called significantly differentially expressed between two conditions at a false-discovery rate (FDR) of < 0.05, *p*-value of < 0.01 and a log₂ fold-change > 2, and were subsequently interpreted by using Gene Ontology (GO) analysis [21]. GO terms of *P. veronii* genes were inferred using the program BLAST2GO [32]. The same program was then used to analyze for enrichment of significantly differentially expressed genes in pair-wise comparisons, using the Fisher's exact test and correcting for multiple testing. A simplified network of common and soil-specific enriched GO terms for the biological process category was manually constructed in Cytoscape (version 3.7.2) [33] with nodes representing a biological process and edges the connecting parent-child.

A survey of *P. veronii* 1YdBTEX2 metabolic capacities was extracted from its reconstructed genome-scale metabolic model (iPsvr), which accounts for 1234 metabolic genes [23]. These genes translate potentially into 1812 metabolic reactions. Gene-reaction associations were extracted from the model, using the in-built RAVEN [34] functions, and linked to their potential corresponding metabolic pathways. iPsvr is curated based on KEGG [22] and all the reported reactions, metabolites and pathways follow KEGG nomenclature. Normalized gene expression values for the 1234 metabolic genes under the different conditions were calculated and compared using edgeR [31]. Identified KEGG reactions were visualized on a general metabolic map using iPath3 [35] with line thickness representing the log₂ normalized expression attributed to that reaction. Those reactions whose mean expression differed by two or more standard deviations in EXPO phase among cells in liquid, Silt or Jonction, were highlighted in different color. Maps were exported to Adobe Illustrator (vs 2020). The mean normalized expression of all genes attributed to KEGG reactions under all conditions was further visualized as a heatmap with rows (i.e., genes) clustered in Euclidian distance (MATLAB vs 2016a, *clustergram* function).

Database submission

The raw unmapped RNA-seq reads related to this study have been deposited in the NCBI Short Read Archive under Bioproject accession number PRJNA682712.

Results

Comparative growth of *P. veronii* in different soils with toluene as added carbon substrate

In order to benchmark growth of *P. veronii* in different soils, we inoculated microcosms dosed with toluene in comparison to liquid suspended medium or to artificial porous medium (APM, Fig. 1). Growth rates of *P. veronii* miniTn5::gfp with toluene in polluted (Junction) and two natural soils (Sand and Silt) were slightly higher and statistically different from those measured in liquid or in APM ($p_{\text{adj}} = 3.49 \times 10^{-5}$, ANOVA, followed by post-hoc Tukey test, Table 1). Maximum population sizes were highest in Silt and Sand, then liquid culture and Junction material, followed by APM and Clay (Table 1). In liquid culture and in all microcosms, except perhaps in Junction, the viable *P. veronii* population size decreased 8–10 fold after its maximum (Fig. 1), suggesting cell death. An even more pronounced decline was observed in APM (Fig. 1a).

In contrast to the other microcosms, the *P. veronii* miniTn5::gfp population in Clay developed much poorer, never reaching more than 2.7×10^5 CFU g⁻¹ (Fig. 1b). Furthermore, the mean relative variability between replicates in Clay was higher (65% vs 8–38%), and CFU numbers increased only one order of magnitude in comparison with the measured inoculated *P. veronii* population size (1.28×10^4 CFU g⁻¹). Population size variations in Clay suggested alternating growth and decline, perhaps as a result of cycles of predation (given that the soils were not sterilized). Predation may have been more pronounced in the Clay microcosms because of their slightly higher starting water content (Methods), in which protozoan may have remained alive for longer than in air-dried Sand or Silt [36–38]. Growth of native soil microbiota on added toluene was observed for Clay and Silt but not Sand (Supplementary Figure 4), suggesting that some substrate competition may have occurred with the inoculated strain.

P. veronii growth in soil as a function of degree of contamination

Next, we tested growth in microcosms consisting of Silt-Junction mixtures with different degree of contamination (Methods), in order to determine to what extent *P. veronii* could grow in field-collected polluted soil. Since contaminated material from Junction contains monoaromatic compounds (see above) [27], we hypothesized that *P. veronii* might be able to grow in absence of externally added toluene. Contrary to our expectations, however, *P. veronii* miniTn5::gfp developed poorer in microcosms with a higher degree of Junction material but without added toluene (Fig. 2). Whereas *P. veronii* grew to a population size of 4.1×10^7 CFU g⁻¹ in Silt (without added toluene) within 24 h, in 100% Junction it only reached 2.7×10^5 CFU g⁻¹. Mixing Junction material with Silt overproportionally reduced the final attained *P. veronii* population size (Fig. 2a). For example, with 50% Junction:Silt the final *P. veronii* population size was only one-tenth from that on Silt alone. This indicated that some carbon is available for *P. veronii* growth in (non-sterile) Silt, but that the strain is inhibited by components or factors originating from Junction. The poorer growth in microcosms mixed with Junction material further suggested that *P. veronii* found little available aromatic substrates (Supplementary Figure 4).

In contrast, *P. veronii* grew very rapidly in all Junction microcosms to which external toluene was supplied through the gas phase (Fig. 2b), and reached higher maximum population sizes than in those without (Fig. 2a), indicating that the cells were mainly using toluene as carbon source. Also native microbiota profited from addition of toluene (Supplementary Figure 4). There was less of an effect of increasing Junction proportions, with highest population growth in the microcosms with 25% Junction compared to higher Junction proportions. Toluene may thus have given *P. veronii* some advantage to provide the energy necessary to protect itself from potentially harmful substances present in Junction.

We subsequently used this physiological context to measure the genome-wide expression differences of wild-type *P. veronii* as a function of material and growth

Table 1 *Pseudomonas veronii* miniTn5::gfp growth kinetic parameters on toluene^a in porous media and in liquid suspension

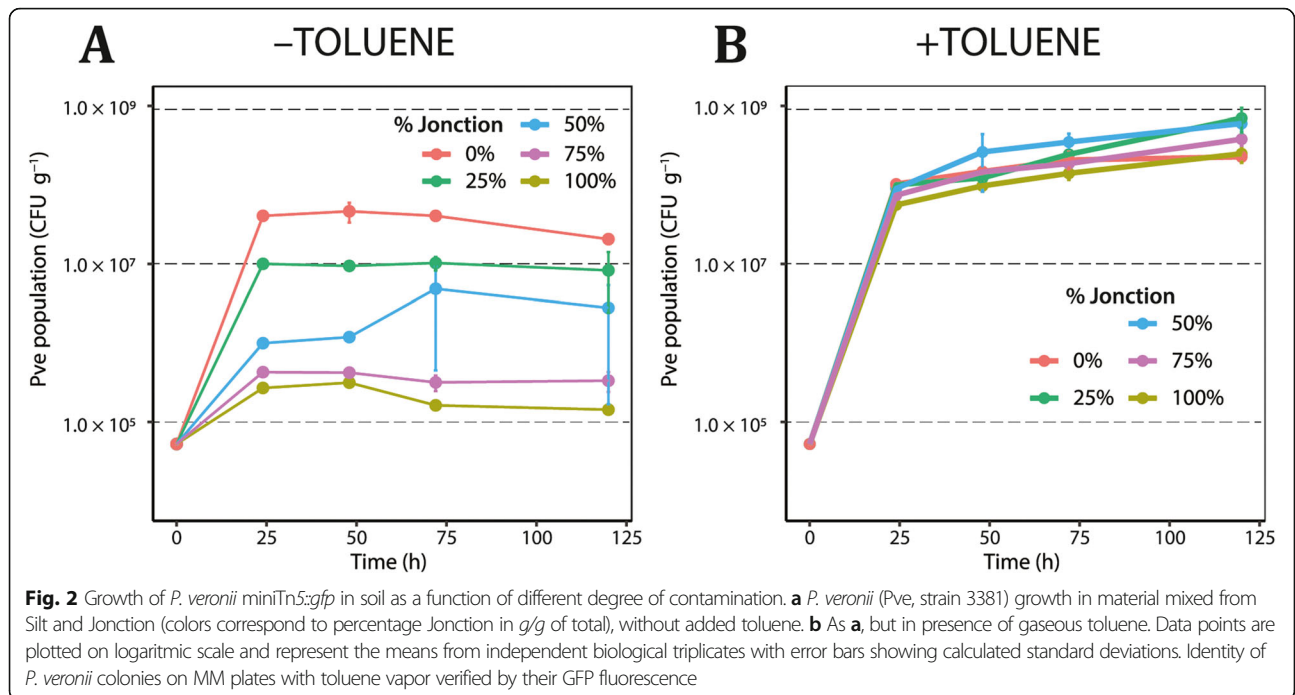
Parameter ^b	Control		Natural soil			Polluted
	Liquid	APM ^c	Sand	Silt	Clay	Junction ^d
max pop size	1.1×10^8	2.6×10^7	2.7×10^8	3.9×10^8	2.7×10^5	1.4×10^8
μ_{max}	0.22 ± 0.01	0.23 ± 0.03	0.27 ± 0.01	0.32 ± 0.01	0.27 ± 0.4	0.29
G	2.43 ± 1.06	2.66 ± 0.03	2.60 ± 0.11	2.43 ± 0.26	2.51	2.39
N	10	10	11	13	5.6	10

^a Toluene dosage in liquid culture was dosed from a secondary oil-phase, whereas in soils and APM it was dosed through the gas phase

^b max pop size = mean maximum population size (CFU g⁻¹ for solid matrix, and CFU ml⁻¹ for liquid suspension); μ_{max} = mean maximum growth rate \pm stdev (h⁻¹, n = 3); G = generation time (h). N = estimated mean number of generations of growth

^c APM Artificial porous matrix

^d Single replicate because of insufficient material. Junction here is 100%



phase. We differentiated and compared three growth phases across the different materials: a transition phase (1 h after inoculation), exponential growth and stationary phase (time points depending on material and estimated from growth of the GFP-tagged variant, as defined in Table 2). Global transcriptomic responses of *P. veronii* to the different materials and conditions were very consistent, given distinct and good replica clustering in multi-dimensional scaling (Fig. 3a).

Effect of porosity on the transcriptomic transition response of *P. veronii*

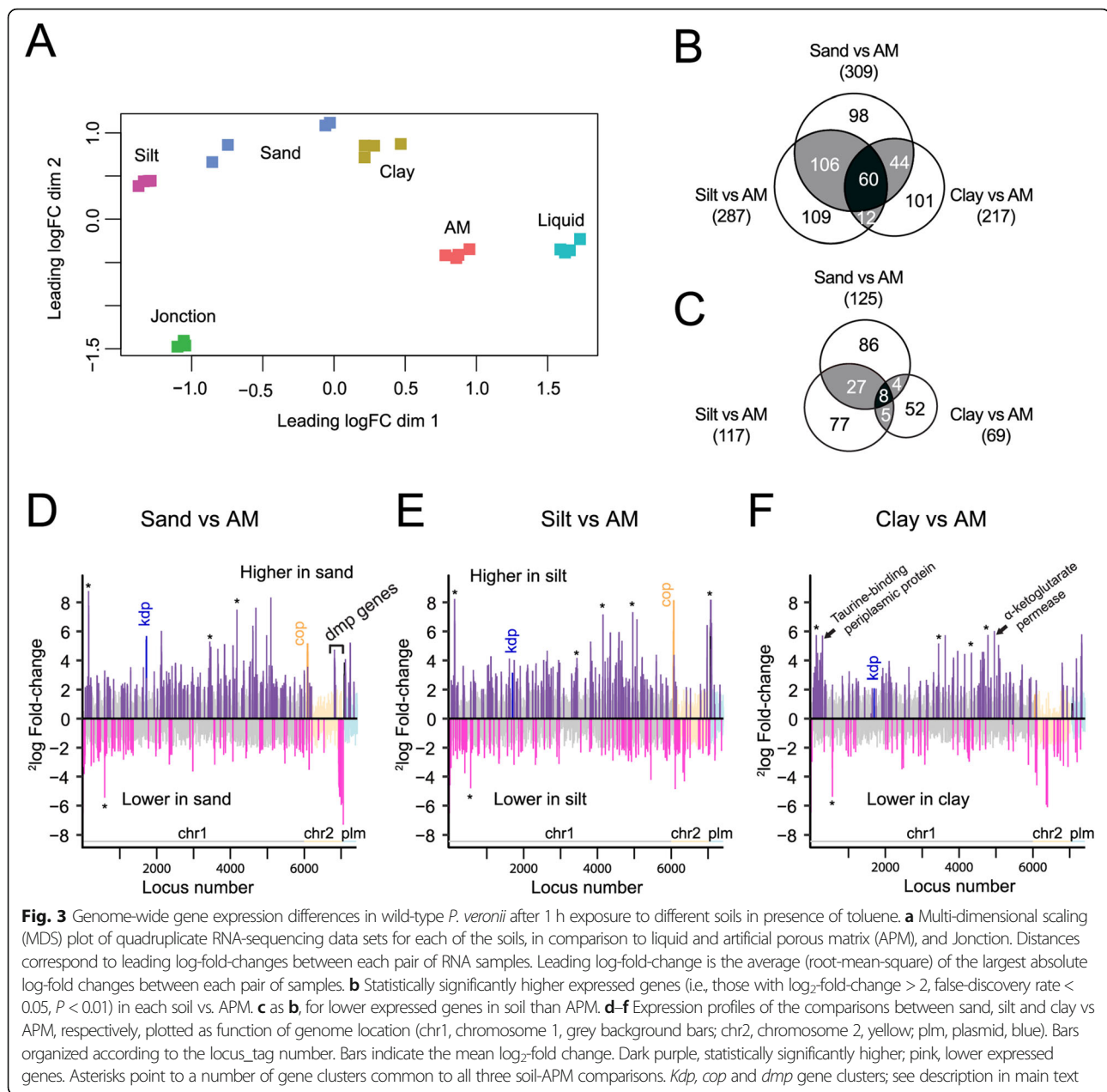
In order to distinguish potential effects originating from matrix porosity, we first compared transcriptomes from inoculated *P. veronii* after 1 h in APM microcosms and in liquid suspension (LQ), both in presence of toluene (Table 2). In comparison to LQ, a total of 319 *P. veronii* genes were differentially expressed in APM (131 higher, 188 lower; Supplementary Data 1). One-third of the

Table 2 Comparison groups and conditions of genome-wide gene expression analysis

Phase	Sample ^a	Material	Sampling Time (h) ^b	Comparison groups
Transition	LQ-1H	Liquid	1	
	APM-1H	Pure silica	1	APM-1H/LQ-1H
	SILT-1H	Silt	1	SILT-1H/APM-1H
	SAND-1H	Sand	1	SAND-1H/APM-1H
	CLAY-1H	Clay	1	CLAY-1H/APM-1H
	JN-1H	Jonction	1	JN-1H/APM-1H JN-1H/soils-1H
Exponential	LQ-EXPO	Liquid	4	LQ-EXPO/LQ-1H
	SILT-EXPO	Silt	29	SILT-EXPO/SILT-1H
	JN-EXPO	Jonction	18	JN-EXPO/JN-1H
Stationary	LQ-STAT	Liquid	24	LQ-STAT/LQ-1H
	APM-STAT	Pure silica	70	APM-STAT/APM-1H
	SILT-STAT	Silt	72	SILT-STAT/SILT-1H

^aLQ, liquid suspension; JN, jonction; APM, artificial matrix; 1H, transient 1-h exposure; EXPO, exponential growth phase; STAT, stationary phase

^bSampling time corresponds to Fig. 1a and b



APM-higher expressed genes coded for conserved or secreted hypothetical proteins. GO analysis showed no particular enrichment of any biological function or pathway (Supplementary Data 2). However, APM-exposed cells showed 4–5 fold increased expression of a gene cluster for the uptake (*potGHI*) and metabolism (*puuACD*, PVE_r1g3330–3336) of putrescine (Supplementary Table 2), which is an important molecule during carbon and nitrogen starvation and in stress defense regulation [39]. Furthermore, a glycine utilization system encoded by the *gcvH2-gcvP2-sdaA-gcvT2* gene cluster was also 4–6 fold higher expressed in APM (Supplementary Table 2).

Among the 188 lower expressed genes in APM compared to LQ, GO analysis indicated enrichment of biological processes involved in growth and energy generation (Supplementary Data 2), many of which were directly implicated in toluene metabolism (e.g., PVE_r2g0711–0717 [*dmpBCDEFGH*], PVE_r2g0739–0742 [*ipbAaAbAcAd*]; Supplementary Table 2). This suggested that *P. veronii* detected less toluene in APM than in liquid. Further genes with lower expression in APM included the *leuBCD* and the *ipdVbkdBbkdA2A1* operons involved in the synthesis of branched-chain amino acids such as leucine, isoleucine, and valine (Supplementary Table 2). GO analysis suggested higher oxygen availability to cells in APM, judged from lowered expression of 7 out of 13 genes from the NADH

dehydrogenase complex I (*nuoHIJKLMN*), genes from the *nar* operon for nitrate respiration (*narGJKLY*), a gene encoding oxygen-independent coproporphyrinogen-III oxidase (*hemN*), and genes coding for the *cbb3* cytochrome C oxidase isoforms, *ccoNI*, *ccoO* and *ccoG2* (Supplementary Table 2).

Collectively, these results indicated that cells in APM perceived carbon and nitrogen limitation, possibly differences in oxygen provision, triggering a response to start recycling nitrogen-rich compounds and amino acids, and decreasing the synthesis of branched-chain amino acids (i.e., leucine, isoleucine and valine). These effects may thus be solely a response to change in porosity from the growth matrix itself.

Effect of soil types on the immediate *P. veronii* response

In comparison to APM, the transient contact of *P. veronii* in different soil types (i.e., 1 h Sand, Silt and Clay) caused a further common core of 68 genes (60 higher and 8 lower) to change expression (Fig. 3b and c, black zones, Table 3). Differences in gene expression were not located to specific genome positions or replicons (Fig. 3d–f). Commonly higher expressed biological processes in soils included, notably, the pathways “nitrogen compound metabolic process” (GO:0006807) and “DNA metabolic process” (GO:0006259), as well as “benzoate catabolic process”, “carbohydrate metabolism and transport”, and “aromatic amino acid catabolic process” (Fig. 4a, Supplementary Datas 3, 4, 5, 6, 7, 8, 9 and 10).

Soil-specific responses in comparison to APM comprised some further 200–300 genes (Fig. 3b and c, white or grey zones, Supplementary Data 3, 4 and 5). Particularly in Sand, these consisted of higher expressed GO-pathways “regulatory and metabolic processes”, “RNA biosynthesis” and “DNA-dependent transcription” (Supplementary Data 6). In Silt, the terms “translation” (GO:0006412), “peptide biosynthetic process” (GO:0043043), and “amide biosynthetic process” (GO:0043604) were underrepresented (Supplementary Data 7, 8). In contrast, for Clay, enriched GO terms associated with decreased “cell motility” (GO:0048870), “localization” (GO:0051179), and “flagellum assembly” (GO:0044780, Fig. 4b, Supplementary Data 9, 10).

In comparison to both LQ and APM controls, there were 293 common genes at least four-fold higher and 22 four-fold lower expressed in soils ($p < 0.05$, ANOVA, Supplementary Data 11). About 29% of those encode either hypothetical proteins, hypothetical secreted proteins, membrane proteins, or conserved hypothetical proteins. The gene coding for ornithine aminotransferase (PVE_r1g5099) was among those whose expression increased the most in soils (165-fold), emphasizing the apparent importance for polyamine biosynthesis [39]. A number of genes with hypothetical function were induced at least by 100-fold (Supplementary Table 3). Other genes with

significantly increased expression in soils included the *pgaABCD* operon, which is involved in the synthesis of extracellular polysaccharide. Also *tauAB*, *kdpFABC*, and at least two *copRSABCD* gene clusters, which are involved in taurine, potassium and copper transport, respectively, were upregulated in soil. Furthermore, gene clusters (PVE_r1g5622–5628) involved in malonate transport and metabolism, and of *bkdA1A2B*, involved in the metabolism of branched-chain amino acids, were higher expressed (Supplementary Table 3). The *dmpBCDEFGH* gene cluster for phenol and catechol *meta*-cleavage degradation together with two additional genes (PVE_p0191–0193) and the *narGIJKY* gene cluster for nitrate respiration were also significantly induced in *P. veronii* in soils after 1 h (Supplementary Table 3).

Collectively, these results thus suggested that, in contrast to APM, *P. veronii* cells in soils do not face nutrient limitation, and possibly gain additional proton motive force from nitrate reduction. Cells were adjusting their metabolism and transport systems for the available resources. The soil environment further triggered defense mechanisms to protect against osmotic stress or favor biofilm formation.

Transition transcriptomic response of *P. veronii* upon inoculation into Junction polluted material in the presence of toluene

In order to understand whether adaptation of *P. veronii* to a field-polluted material is different than to artificially contaminated soils, we compared the transient response in Junction material in the presence of toluene (Tables 3 & 4). Multi-dimensional scaling analysis confirmed that the transcriptomic response to Junction was globally different from the clean soils, despite the common presence of externally added toluene (Fig. 3a). In comparison to APM alone (Fig. 5a), 416 genes were higher, and 246 lower expressed in Junction (Fig. 5b, Supplementary Data 12). Approximately 30% of those correspond to (conserved) hypothetical and hypothetical secreted proteins, and (conserved hypothetical) membrane proteins. A total of 137 of Junction-differentially expressed genes (130 up and 7 down) were shared with those of *P. veronii* in clean soils (Fig. 5c and d, Supplementary Data 13). These included genes involved in L-valine degradation, dissimilatory nitrate reduction, and copper transport (Table 3). Further common to clean and polluted soils was a strong induction of the *hmp* gene, coding for a flavohemoprotein (Fig. 5a and Supplementary Data 13), which has been described to be important in NO detoxification in response to nitrosative stress [40]. This suggests that cells are experiencing stress from noxious nitrogen compounds.

Among the genes induced specifically in Junction and not in clean soils (286 genes, Supplementary Data 14),

Table 3 Common differentially expressed genes of *Pseudomonas veronii* wild-type during transition in soils

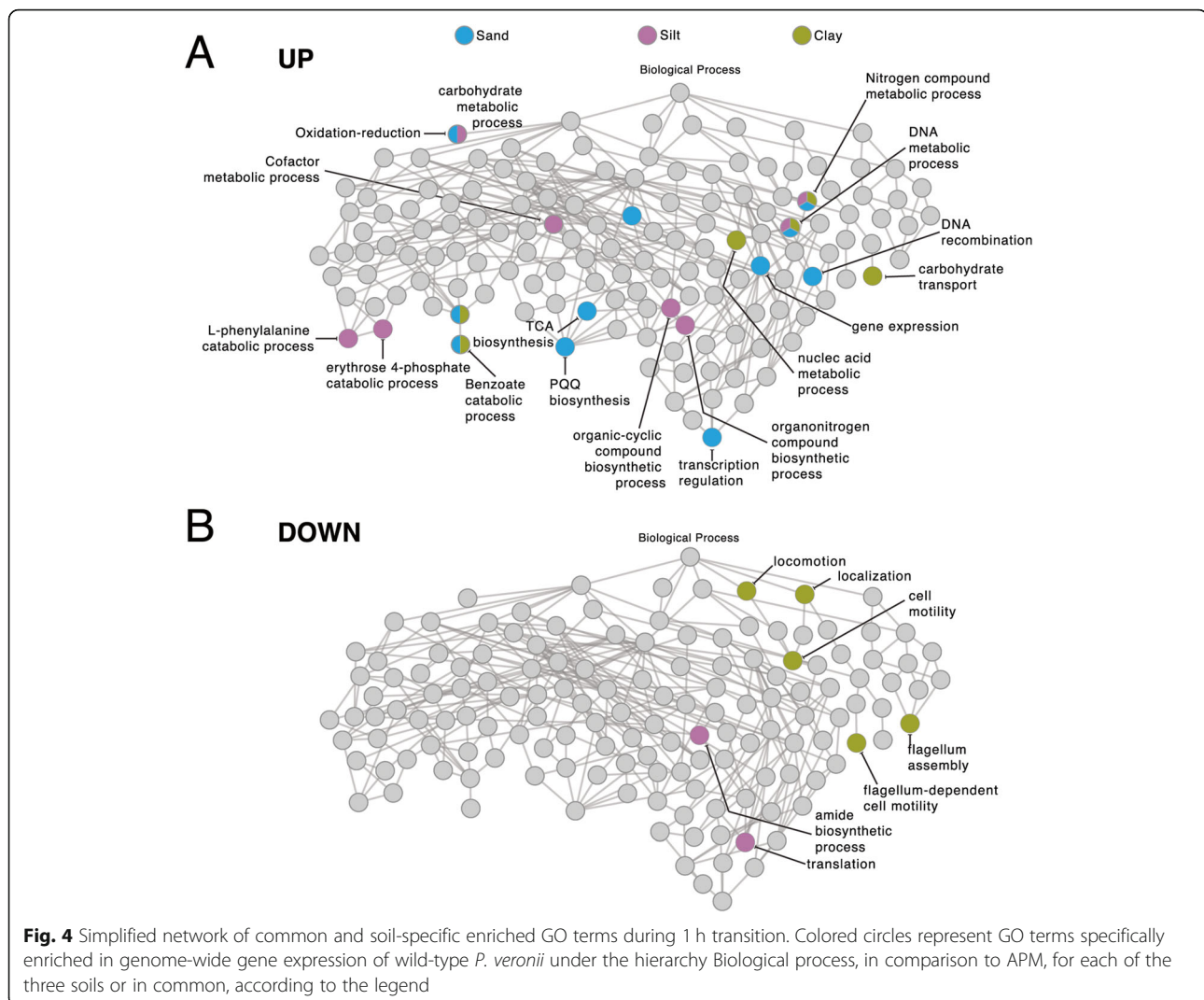
Gene ID	Gene	Gene function	Log ₂ fold-change in comparison ^a					
			Sand vs APM	Silt vs APM	Clay vs APM	Junction vs APM	One-way ANOVA ^b	
					Natural soils vs Controls	Junction vs Controls		
PVE_r1g2130	<i>lpdV</i>	dihydrolipoyl dehydrogenase	2.07	2.42	0.31	2.14	2.00	2.00
PVE_r1g2131	<i>bkdB</i>	dihydrolipoamide branched chain transacylase	3.07	2.39	1.75	3.11	2.57	2.00
PVE_r1g2132	<i>bkdA2</i>	2-oxoisovalerate dehydrogenase subunit beta	3.65	3.16	2.50	3.78	3.25	3.58
PVE_r1g2133	<i>bkdA1</i>	2-oxoisovalerate dehydrogenase subunit alpha	3.44	3.13	1.60	3.84	3.01	3.58
PVE_r1g5053	<i>hibA</i>	probable 3-hydroxyisobutyrate dehydrogenase	2.83	3.40	1.86	3.10	2.91	2.92
PVE_r1g5054	<i>ALDH6B2</i>	methylmalonate-semialdehyde dehydrogenase	3.63	3.54	2.91	3.98	3.45	3.81
PVE_r1g2545	<i>narK</i>	nitrate/nitrite transporter NarK	3.80	1.97	0.00	4.20	2.76	3.87
PVE_r1g2546	<i>narG</i>	respiratory nitrate reductase 1 alpha chain	3.72	2.33	-0.60	3.68	2.77	3.37
PVE_r1g2547	<i>narY</i>	respiratory nitrate reductase 2 beta chain	3.51	2.41	0.03	3.35	2.68	3.04
PVE_r1g2548	<i>narJ</i>	nitrate reductase molybdenum cofactor assembly chaperone	2.84	2.58	0.19	2.85	2.36	2.56
PVE_r1g2549	<i>narI</i>	respiratory nitrate reductase 1 gamma chain	2.60	2.14	0.11	2.56	2.04	2.29
PVE_r1g0681	<i>copA</i>	copper-exporting P-type ATPase A	2.39	3.30	0.38	3.51	2.56	3.22
PVE_r1g6093	<i>copC</i>	copper resistance protein C	4.53	6.00	1.61	3.22	5.04	3.05
PVE_r2g0903	<i>copC</i>	copper resistance protein C	1.27	6.39	1.21	3.35	5.00	3.12
PVE_p0049	<i>copB</i>	copper resistance protein B	3.00	2.86	1.63	2.58	2.68	2.45
PVE_p0050		hypothetical secreted protein	3.19	3.64	2.11	3.19	3.18	3.03
PVE_p0051	<i>copA</i>	copper resistance protein A	3.02	3.50	1.38	2.25	2.97	2.13
PVE_p0052		hypothetical secreted protein	4.04	6.54	0.70	3.53	5.35	3.26
PVE_r1g0242		membrane protein	2.78	2.25	2.60	2.51	2.57	2.51
PVE_r1g0243	<i>pgaB</i>	poly-beta-1,6-N-acetyl-D-glucosamine N-deacetylase	2.57	1.88	2.34	2.57	2.09	2.52
PVE_r1g0244	<i>pgaC</i>	poly-beta-1,6-N-acetyl-D-glucosamine synthase	2.29	1.40	2.03	2.52	2.09	2.46
PVE_r1g0245	<i>pgaD</i>	biofilm PGA synthesis protein PgaD	2.72	1.66	2.11	2.56	2.07	2.46

^a Log₂ fold-change in condition compared with control in soil (FDR < 0.05, *p* < 0.01)

^b One-way ANOVA (analysis of variance) in the condition "soil"; natural or polluted vs controls (LQ and APM)

were those of the *moaA1BA* (PVE_r1g2551–2517) operon, which is responsible for molybdopterin biosynthesis (Table 4), and the *arcDABC* genes, responsible for L-arginine degradation. Furthermore, the *soxGADB-*

glyA3 cluster was induced in Junction, which is responsible for the conversion of sarcosine into serine, and in *Pseudomonas aeruginosa* is involved in adaptation to using alternative carbon, nitrogen and energy sources



for growth [41, 42]. In addition to the *nar* genes, some of the *nir* genes (*nirGLD*-PVE_r1g2514–2516) were also higher expressed in cells in Junction, suggesting further induction of anoxic respiration pathways.

Finally, several other defense systems were higher expressed in Junction (Table 4). These encompassed the *merTPCA*-PVE_r1g0818 operon, which encodes for mercury resistance, genes coding for “Universal stress protein” [43], for chaperones such as *clpB*, *htpG*, *groS*, *groL* and *dnaK*, catalases (e.g., PVE_r1g0164 and PVE_r1g5393) and for the phosphate starvation-inducible protein PsiF. The GO enrichment analysis of transcriptomic response in Junction was coherent with these observations (Supplementary Data 15, 16). This indicated that *P. veronii* perceived a more stressful environment in Junction than in the clean soils, although all contained artificially added toluene.

Genome-wide expression differences during growth

After having uncovered the specific expression differences during the transition of *P. veronii* into clean or contaminated soils compared to liquid and APM, we next focused on measuring its metabolic reprogramming during actual growth in soil. Given that the three soils had shown comparable global transition reactions (Fig. 3) and that *P. veronii* did not grow very well in Clay (Fig. 1), we restricted ourselves in this analysis to Silt. Unfortunately, we could not recover sufficient RNA for analysis from *P. veronii* growing exponentially in APM, nor from stationary phase cells in Junction. Thus, we could finally compare six experimental conditions, all for cells grown in the presence of toluene (Table 2, EXPO and STAT phase).

Growth environments produced clear distinct global signatures with excellent replicate clustering, and EXPO

Table 4 Subset of Junction-specific differentially expressed genes of *Pseudomonas veronii* wild-type during 1 h transition

Locus_tag	log ₂ FC ^a	Gene	Gene function
PVE_r1g0226	2.84		tail protein
PVE_r1g0616	2.11	<i>psiF</i>	Phosphate starvation-inducible protein PsiF
PVE_r1g0858	2.46	<i>clpB</i>	chaperone protein ClpB
PVE_r1g1263	2.17		tail protein
PVE_r1g1264	2.36		phage tail protein
PVE_r1g1271	2.36		pyocin R, lytic enzyme
PVE_r1g1633	2.23		universal stress protein A
PVE_r1g2181	2.10		RND transporter
PVE_r1g2307	3.59		multidrug RND transporter
PVE_r1g2308	3.57		multidrug transporter
PVE_r1g2309	2.29	<i>emrB</i>	multidrug export protein EmrB
PVE_r1g2514	2.34	<i>nirG</i>	Protein NirG
PVE_r1g2515	2.12	<i>nirL</i>	Protein NirL
PVE_r1g2516	2.05	<i>nirD</i>	Protein NirD
PVE_r1g2551	2.75	<i>moaA1</i>	cyclic pyranopterin monophosphate synthase 1
PVE_r1g2614	2.96	<i>moaB</i>	molybdenum cofactor biosynthesis protein B
PVE_r1g2615	2.41	<i>moaA</i>	molybdenum cofactor biosynthesis protein A
PVE_r1g2895	2.65		Transposase for insertion sequence element IS1328
PVE_r1g3164	2.33		universal stress protein
PVE_r1g3188	2.26		universal stress protein
PVE_r1g3381	2.43		RND transporter MFP subunit
PVE_r1g3768	2.26		universal stress protein
PVE_r1g4143	2.30	<i>htpG</i>	chaperone protein HtpG
PVE_r1g4506	2.17		nitrate reductase
PVE_r1g4749	2.36	<i>arcA</i>	arginine deiminase
PVE_r1g4750	2.77	<i>arcB</i>	ornithine carbamoyltransferase, catabolic
PVE_r1g4751	2.28	<i>arcC</i>	carbamate kinase
PVE_r1g5119	2.63	<i>dnaK</i>	chaperone protein DnaK
PVE_r1g5393	2.99	<i>cat</i>	catalase
PVE_r1g5544	2.13	<i>soxG</i>	sarcosine oxidase subunit gamma
PVE_r1g5545	1.92	<i>soxA</i>	Sarcosine oxidase subunit alpha
PVE_r1g5546	2.23	<i>soxD</i>	sarcosine oxidase subunit delta
PVE_r1g5547	2.70	<i>soxB</i>	Sarcosine oxidase subunit beta
PVE_r2g0272	2.51		Insertion element IS2A uncharacterized 48.2 kDa protein
PVE_r2g0273	2.07	<i>insC1</i>	transposase InsC for insertion element IS2
PVE_r2g0743	2.88	<i>insH5</i>	transposase InsH for insertion sequence element IS5Y
PVE_r2g0813	0.22	<i>merR</i>	mercuric resistance operon regulatory protein
PVE_r2g0814	2.30	<i>merT</i>	mercuric transport protein
PVE_r2g0815	2.24	<i>merP</i>	mercury resistance system
PVE_r2g0816	2.39	<i>merC</i>	putative mercury transport protein MerC
PVE_r2g0817	2.25	<i>merA</i>	mercuric reductase
PVE_r2g0818	2.29		conserved hypothetical membrane protein
PVE_r2g0819	2.13	<i>merD</i>	mercuric resistance transcriptional repressor
PVE_r2g0831	3.43	<i>insH5</i>	transposase InsH for insertion sequence element IS5Y

Table 4 Subset of Junction-specific differentially expressed genes of *Pseudomonas veronii* wild-type during 1 h transition (Continued)

Locus_tag	log ₂ FC ^a	Gene	Gene function
PVE_r2g0931	7.23	<i>tnpA1</i>	putative transposase, TnpA1
PVE_p0038	2.16		RND transporter
PVE_p0170	2.00		mobile_element
PVE_p0207	2.61		integrase

^aIn comparison to 1 h in APM

was well-separated from STAT phase responses (Fig. 6a). A total of 175 genes were commonly higher and 246 lower expressed in exponential phase conditions compared to the 1 h transition (Fig. 6b, Supplementary Data 17), and 17 genes were commonly higher and 14 lower expressed in stationary phase compared to the 1 h signatures (Fig. 6c). The number of uniquely

differentially expressed genes in any of the conditions surpassed those of the commonalities (Fig. 6b, c).

As expected, both GO analysis and individual gene annotations confirmed most of the commonly higher expressed genes in exponentially growing cells to be related to growth and energy production (Table 5, Supplementary Data 18, 19, 20, 21, 22 and 23). For example,

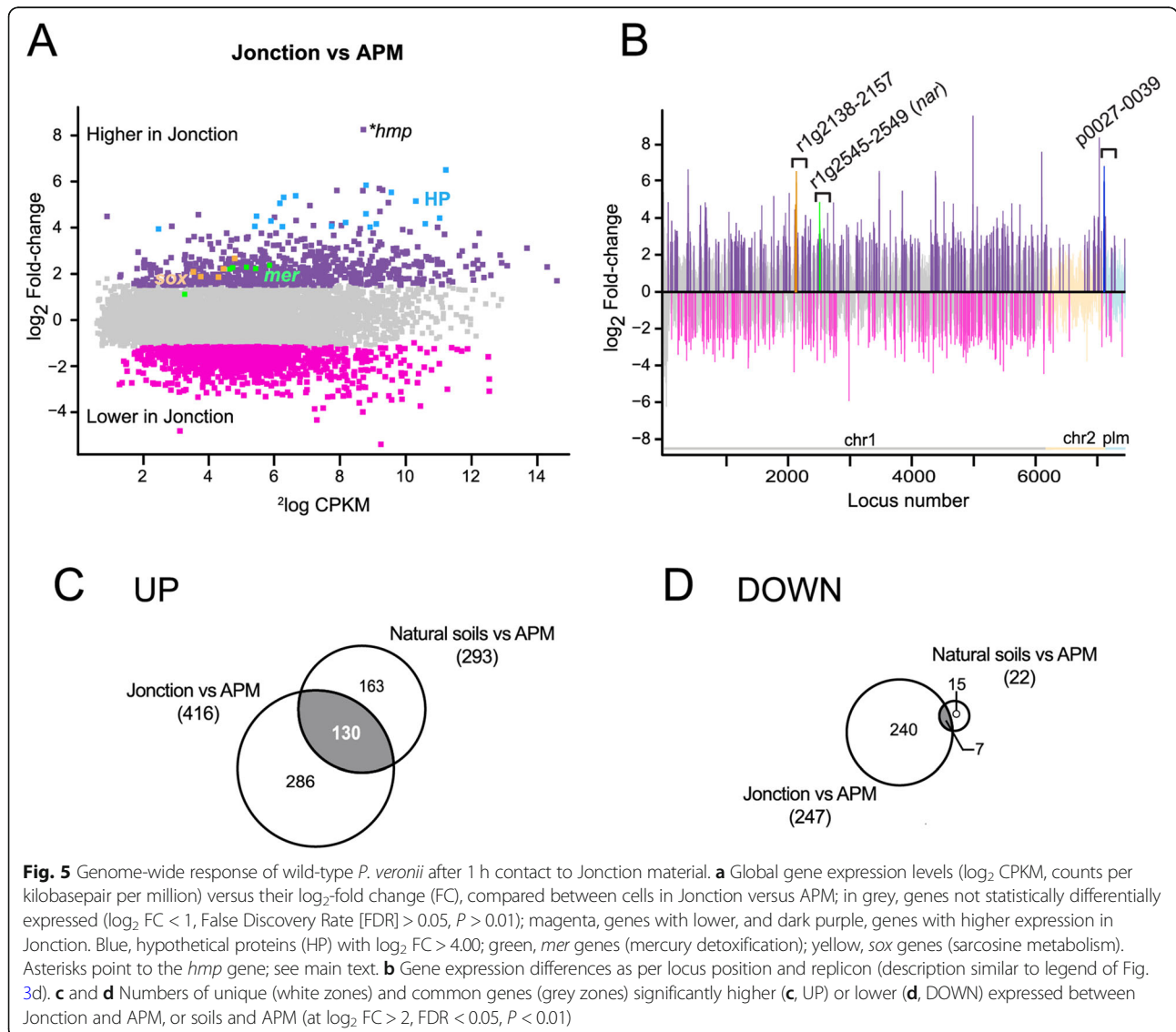
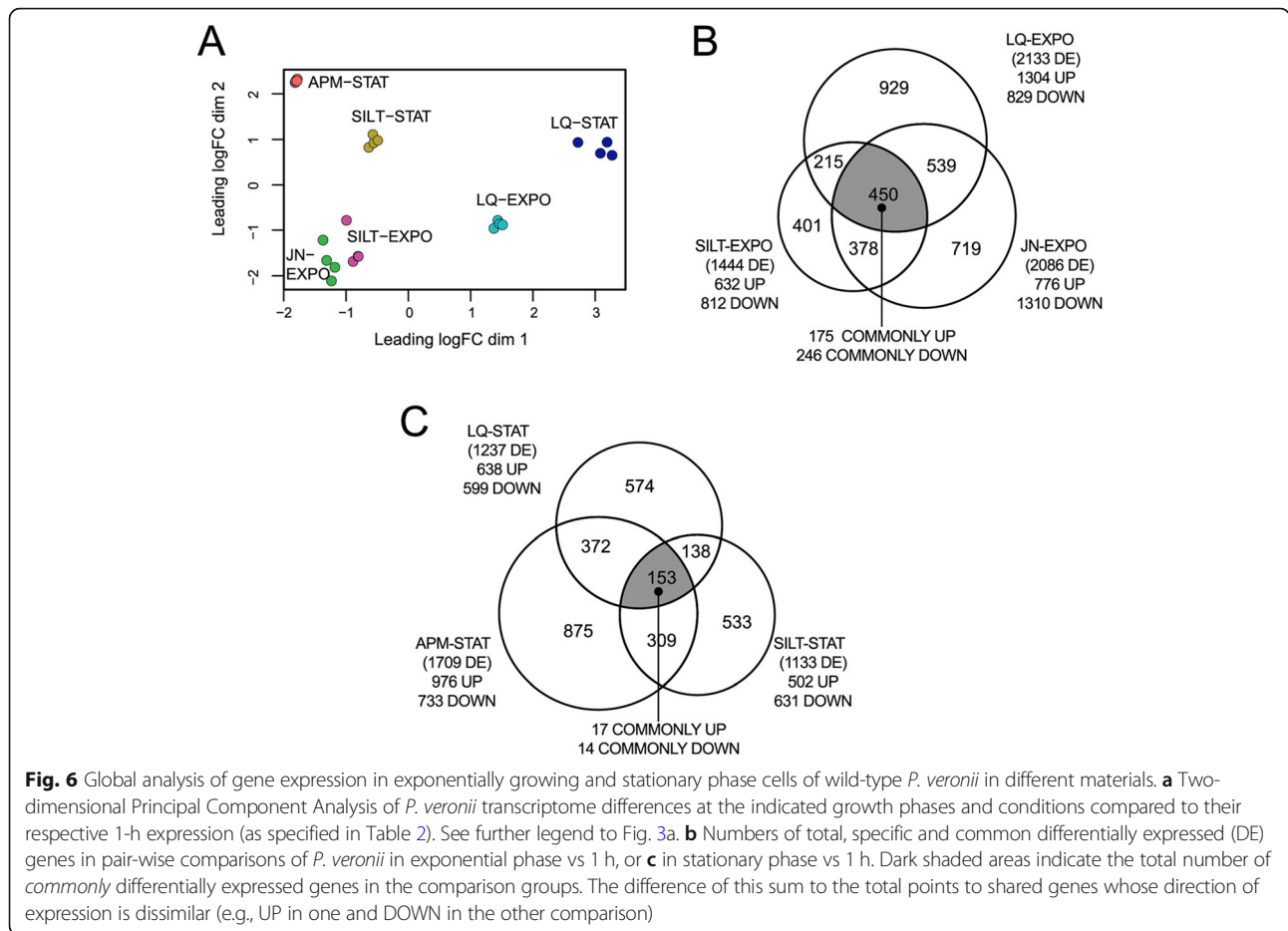


Fig. 5 Genome-wide response of wild-type *P. veronii* after 1 h contact to Junction material. **a** Global gene expression levels (log₂ CPKM, counts per kilobasepair per million) versus their log₂-fold change (FC), compared between cells in Junction versus APM; in grey, genes not statistically differentially expressed (log₂ FC < 1, False Discovery Rate [FDR] > 0.05, P > 0.01); magenta, genes with lower, and dark purple, genes with higher expression in Junction. Blue, hypothetical proteins (HP) with log₂ FC > 4.00; green, *mer* genes (mercury detoxification); yellow, *sox* genes (sarcosine metabolism). Asterisks point to the *hmp* gene; see main text. **b** Gene expression differences as per locus position and replicon (description similar to legend of Fig. 3d). **c** and **d** Numbers of unique (white zones) and common genes (grey zones) significantly higher (**c**, UP) or lower (**d**, DOWN) expressed between Junction and APM, or soils and APM (at log₂ FC > 2, FDR < 0.05, P < 0.01)



higher expressed genes included those for (i) ribosome assembly and protein synthesis (*rpoAB*, *rps* and *rpl*), elongation factors (EF), *fusA* (EF-G), *efp* (EF-P), *tsf* (EF-Ts), and *tuf* (EF-Tu); (ii) DNA replication; (iii) ATP synthesis; (iv) sulfate transport; (v) TCA cycle, and (vi) amino acid biosynthesis.

In contrast, signatures of commonly lower expressed genes were globally less clear. Around 50% of those encoded hypothetical proteins (Table 5). Others included transcriptional factors from the AraC and LuxE, or from unknown families, suggesting fine-tuning of metabolism to the uptake of resources. Genes for Universal Stress Protein A (PVE_r1g0156) and cold shock protein (PVE_r1g3189) were commonly lower expressed (Table 5), possibly associated with a release of stress in dividing cells.

Genes specifically higher expressed in exponential phase in Silt and Jonction included those for proteins involved in phenol and catechol *meta*-cleavage metabolism (*dmp* gene cluster) and urea transport (Table 5). This suggested that cells in soils were increasing flux through the *meta*-cleavage pathway, possibly because of higher perceived toluene availability than in liquid.

Among the genes specifically higher expressed during exponential growth in Silt (Supplementary Data 24) were those for taurine transport, urea degradation, chaperones, pilus biogenesis, flagellum assembly, and chemotaxis (Table 5), suggesting cells to be actively moving around, limited by and scavenging for nutrients. Interestingly, cells in Silt further activated the alternative *nah* pathway for aromatic compound metabolism (Table 5), in addition to *dmp*. This indicated usage of other aromatic carbon sources present in Silt, since the *nah* genes are not induced by toluene [20]. These aromatic compounds may be a common fraction of soil organic matter [44], and this alternate gene expression would explain the observed background growth in Silt without added toluene (Fig. 2a).

Genes induced specifically in exponentially growing cells in Jonction covered multiple unspecified transporters (such as sugar and phosphate ABC transporters, MFS transporters, Supplementary Data 25), uptake of ferripyoverdine, and ribosome biogenesis GTPases. Interestingly, several genes for cell division, cell shape, and peptidoglycan biosynthesis were higher expressed, such as the rod shape-determining proteins MreB and

RodA, and the penicillin binding protein (PVE_r1g5075, Table 5). Defense mechanisms continued to be higher expressed in Junction-growing cells, such as fusaric acid resistance protein (PVE_r1g1039), the multidrug resistance protein MdtC, and a gene coding for a beta-lactamase (PVE_r1g1543, Table 5). Also induction of many transporters may be a sign for defense against toxic compounds.

Analysis of core metabolic reactions

Analysis of the inferred gene sets to be involved in metabolic reactions of *P. veronii* mappable to the KEGG database [22], revealed relatively coherent clustering among gene groups but more loose growth phase and sample clustering (Fig. 7a). This indicated relatively few changes within the core metabolism despite different growth environments. In particular, only 72 metabolic reactions (out of 1234 identified reactions using the iPsvr metabolic scale model [23]) showed outlier behaviour among the three exponential phase datasets (Fig. 7b, Supplementary Table 4). The rest of the metabolic network remained remarkably consistent and similar, given the different growth environments (liquid, Junction, and Silt; Fig. 7b). These 72 reactions occur seemingly arbitrarily within the complete metabolic network, except for a notable pathway expression change in soils connected to the urea cycle (Fig. 7b, '1'). Remarkably, they are two-fold enriched for reactions implicating NADH/NAD⁺ compared to the total metabolic network (Supplementary Table 5), with mostly opposite expression changes between liquid or soil conditions (Junction and Silt, Supplementary Table 4). This might indicate that cells in soils, although they grow at the same rates using essentially the same metabolic pathways as in liquid, replenish NADH/NAD⁺ by activating different side reactions, which they might need for e.g., defense mechanisms.

Condition-specific and commonly expressed genes in stationary phase

Finally, we studied the global responses of *P. veronii* during stationary phase conditions and different growth environments. In comparison to 1 h, cells sampled in stationary phase conditions (Table 2) showed a small number of commonly differentially expressed genes without clear signature (Fig. 6c, Supplementary Table 6). In comparison to LQ, cells in stationary phase conditions in APM or Silt shared 248 higher and 79 lower expressed genes (Fig. 6c, Supplementary Data 26). Among the higher expressed genes were those of the *dmp* cluster and urea catabolism, indicating that cells in Silt or APM were still metabolizing aromatic substrates at that time of sampling (Supplementary Table 7). There was also evidence for increased synthesis of carbon and

energy reserve polymers as known from other pseudomonads [45, 46], such as alginate (*alg*), glycogen (*glg*), and polyhydroxyalkanoates (*phaG*), the latter being induced up to 55-fold in APM and 200-fold in Silt (Supplementary Table 7).

Several transport systems were induced in Silt and APM in stationary phase (Supplementary Table 7), such as for amino acids (*ydh*, *bra*), putrescine (*spuD*), spermidine (*spuE*), nitrate (*nasA*), and sulfate (PVE_r1g3919). Genes associated with resistance mechanisms were higher expressed, for example, organic hydroperoxide resistance protein (*ohr*), and mercury resistance (*mer-ACTP*). Interestingly, several signaling genes such as *cheA*, *cheB2*, *cheW*, and *cheY*, PVE_r1g893, and PVE_r1g2555 were also higher expressed in APM and Silt compared to LQ (Supplementary Table 7). Homologs of those have been implicated in biofilm growth and flagella-mediated twitching [47]. These results indicated that *P. veronii* cells in stationary phase in porous environments remained metabolically active, diverting resources from growth towards synthesizing amino acids and reserve materials to promote survival and attachment.

Discussion

The success of colonization of bacterial inoculants in soil depends on their metabolic and adaptive flexibility encoded in the genome, in broad relation to the environmental conditions of the soil itself, availability of substrates and nutrients, and its prevalent biological factors, such as background microbiota, phages or predators [15]. We show how adaptation, growth, and survival of *P. veronii* under soil conditions requires adjustment of a wide set of metabolic processes, ranging from nutrient requirements and carbon availability, motility and attachment, to respiration, trace metals, and defense against toxicity. By comparing transcriptomic responses in different soils and materials, and at different growth phases, we are confident to have covered a broad range of conditions that help our understanding of the mechanisms necessary for strain adaptation and survival upon inoculation.

Previous studies had indicated the massive difference in gene expression of *S. wittichii* RW1 growing in soil and liquid on dibenzofuran, suggesting that there may be something like a 'soil'-specific transcriptome [48]. Furthermore, we previously studied global responses of three different bacteria taxa (*S. wittichii* RW1, *Arthrobacter chlorophenolicus* A1, and *P. veronii*) in controlled conditions of growth under sublethal solute and matrix stress that mimicked water stress in soils, in order to identify common strategies [17]. By including inert silica particles (APM) here, we could see that a porous matrix by itself causes a global response that is to some extent

Table 5 Common and specific biological processes enriched during exponential growth of wild-type *P. veronii*^a

Regulation	Biological process	Liquid	Silt	Junction
UP	ATP synthesis	<i>nuoEG, atpAFGH</i>	<i>nuoEGKMNL, atpAFGH</i>	<i>nuoEGKLMN, atpABCFGH</i>
	Aerobic respiration	<i>cyoABC</i>	<i>cyoABC</i>	<i>cyoABC</i>
	Toluene catabolic process		<i>dmpBCDFH, nahHIJKLMNOT</i>	<i>dmpBCDGH</i>
	Regulation of transcription		<i>tuf, lepA, efp, arnF, nusA, mfd</i>	<i>nusABG, rho, rpoD, rstA, gntR</i>
	Tricarboxylic acid cycle	<i>sucABC, sdhAB, acnB, icd</i>	<i>sucABC, sdhAB, acnB, icd</i>	<i>sucABC, sdhAB</i>
	Urea catabolism and transport		<i>ureABC, urea ABC transporter</i>	urea ABC transporter
	Translation-ribosome biosynthesis	<i>rpl, rps, prfABC</i>	<i>rpl, rps</i>	<i>rpl, rps, prfABC, GTPase Der, GTPase Era</i>
	Translation-elongation factor	<i>tsf, tuf, fusA, efp</i>	<i>tsf, tuf, fusA, efp</i>	<i>tsf, tuf, fusA, efp</i>
	Glutamate biosynthesis	<i>gltB, gltI</i>	<i>arnF, gltD, gltI</i>	
	Sulfate and taurine transport	<i>cysADW</i>	<i>cysADW, tauABC</i>	<i>cysADW</i>
	Protein folding	<i>surA, cyp18,</i>	<i>htpG, surA, groL, dnaJ, arnF, hscA, cyp18, ppiA</i>	
	Fatty acid biosynthesis	Malonyl-CoA - <i>accABCD, fabDF</i>		Malonyl-CoA - <i>accABCD, fabADFGZ</i>
	DOWN	Water-soluble vitamin metabolic process	Inosine- <i>purADHLMTU</i>	
Bacterial membrane organization		Peptidoglycan/LPS - <i>rfaCG</i> ; cell shape determination - <i>minE, minD</i> ; Lipid A - <i>lpxAB</i> ; Isoprenoids - <i>ispEFG</i> ; cell division <i>zapE</i>	Flagellum assembly - <i>fliDKM, flgBCDE, flaG</i>	Peptidoglycan/LPS - <i>uppSP, lptD, rfaPCG</i> ; Cell shape determination - <i>minE, mrdB, mreB, minD</i> ; Lipid A - <i>lpxABDH</i> ; Isoprenoids - <i>ispEFG</i> ; cell division <i>zapE</i> ; ferripyoverdine receptor - <i>fpvA</i>
Transcriptional regulators:		AraC; ArsR	AraC; ArsR; MerR; LysR; TetR; LuxR; <i>ttgVW</i>	AraC; ArsR, <i>ttgRVW</i> ; <i>qseB</i> ; <i>norR</i> ; <i>copG</i> ;
Organic substance metabolism		branched-chain amino acid - <i>bkdA1A2B</i>	branched-chain amino acid - <i>bkdA1A2B</i> ; arginine - <i>arcABCD</i> ; malonate - <i>mdcBCG</i> ; biopolymer - <i>phbCB</i> ; phenylalanine - <i>hmgA, maiA</i> ; molybdenum - <i>modAA1B</i>	branched-chain amino acid - <i>bkdA1A2B</i> ; <i>lpdV</i> ; vanillate - <i>vanAB</i> ; biopolymer - <i>phhAB</i> ; arginine - <i>dauAB</i> ; molybdenum - <i>modAA1B</i> ; putrescine - <i>putA, puuABC</i>
Nitrate respiration		<i>narJ, nirQS, nosZ</i>	<i>narGJLY, nirQS, norRR2</i>	<i>narGJKL, nirQS, nosZ</i>
Transposition			integrase, transposases, base plate protein	transposases, insertion element
Aerobic respiration		cytochrome B559-B561-CBB3; <i>cydB</i>	cytochrome B559-B561-CBB3; <i>cydB</i>	cytochrome B559-B561-CBB3; <i>cydAB, ctaD</i> ; hem biosynthesis - <i>cpo, hemN, hemH</i> ; ferredoxin, cytochrome c oxidase
Response to stress		flavohemoprotein - <i>hmp</i> ; heat shock protein, universal stress protein A	flavohemoprotein - <i>hmp</i> ; heat shock protein; Universal stress protein A; cold shock protein	flavohemoprotein - <i>hmp</i> ; osmotic stress - <i>cysG</i> ; heat shock protein; virulence sensor <i>bvgAS</i> ; universal stress protein A; Catalase - <i>katE</i> ; cold shock protein

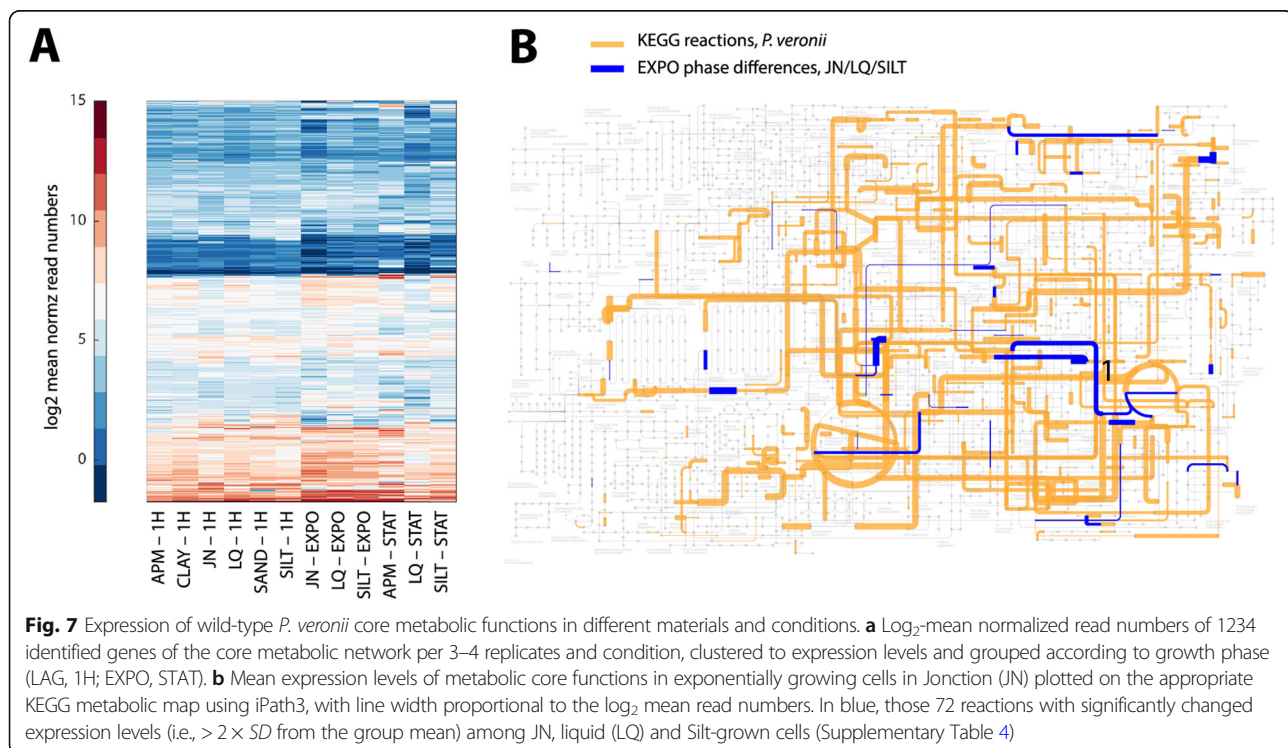
Table 5 Common and specific biological processes enriched during exponential growth of wild-type *P. veronii*^a (Continued)

Regulation	Biological process	Liquid	Silt	Jonction
	Transport			putrescine – <i>potAGH</i> ; glycine betaine – <i>opuAA-AB</i> ; magnesium – <i>mgtAB</i> ; mercury – <i>merATP</i> ; RND transporter; sulfate transporter; citrate transporter, MFS transporter

^aEnrichment defined in comparison to the 1 h transition phase transcriptome signatures

similar to water stress imposed by the addition of solutes or inert swelling agents (matric stress) [17, 49, 50]. Notably, incubation in APM similarly triggered cellular osmoregulation, amino acid recycling and oxidative stress, together with reduced growth, which therefore seem to originate from perceived water stress. It is further worth mentioning that half of the genes that were specifically induced upon entry into porous medium encode hypothetical functions. Even though their function is unknown, they might correspond to a set of genes with important roles in general adaptation to and growth in porous media [17]. On the other hand, the *P. veronii* transcriptomic response observed here in APM and, previously, under imposed matric and solute stresses [17], was different from those in the soils. There is thus a response to soil different from that of porosity alone. Unexpectedly, growth of *P. veronii* with toluene in APM was worse than in liquid or natural soils, suggesting that porous conditions caused differences in the regulation of toluene metabolism in *P. veronii*.

One of the aims of this work was to study the effect of soil type on physiological responses and growth of *P. veronii*, and to understand if there are soil factors that would determine its adaptation and survival. The soils we worked with here have different microbiota background levels, and some showed growth on toluene, which may have invoked substrate competition on the introduced *P. veronii* cells. Unfortunately, we do not understand the nature of signs for competition in the *P. veronii* transcriptome. Soils further varied in pH, organic carbon content, structures and textures, leading to expected differences in e.g., substrate and water availability, matric stress, oxygen fluxes [51], or protozoan activity [38]. Indeed, upon soil inoculation, we observed the onset of nitrate respiration by *P. veronii* and induction of cytochrome c oxidases *cbb₃*, which have been observed as low oxygen environmental responses of *P. aeruginosa* [52], *Pseudomonas stutzeri* A15, and *Pseudomonas fluorescens* YT01 [53, 54]. Nitrate respiration may confer a fitness advantage to *P. veronii* to adapt to soil



microniches with low oxygen concentrations, as previous studies in controlled porous matrices have shown [55]. Possibly, nitrate respiration under low oxygen could be beneficial for toluene metabolism, to divert oxygen to the necessary dioxygenases for toluene breakdown and maintain respiratory energy generation from nitrate as electron acceptor. Adaptation in each soil and condition was accompanied by the selective induction of a variety of transporter systems (such as permeases, porins, multi-drug transporters, amino acid ABC transporters). This illustrates the metabolic versatility of *P. veronii* to adapt to living in soil.

Regardless of the soil type and even in contaminated material such as Junction, transition caused *P. veronii* cells to induce an osmoprotective response (putrescine and potassium uptake), to induce genes for exopolysaccharide production, and to regulate genes for copper homeostasis, indicating the importance of these processes for the initial soil adaptation. Of particular interest is the strong upregulation of systems involved in putrescine uptake and metabolism. The roles of putrescine (and other polyamines) are manifold, having been described both as signals for and during carbon or nitrogen starvation, regulation and oxidative stress defense [39], and recycling and metabolism of arginine [41]. The accumulation of compatible solutes such as putrescine (from which glutamate can be produced) and the uptake of potassium have been reported as the strongest “osmoprotectant strategies” employed by environmental bacteria to balance osmotic differences caused by the alteration in the solute potential of the extracellular environment [56]. Our data showed that inoculation into soil causes osmotic imbalance for *P. veronii*, whose reaction is in line with the three classical physiological responses to osmotic stress: polyamine transport and accumulation of glutamate and potassium [56]. Particularly in Clay, we observed the induction of alginate synthesis, which also has been proposed as an osmo-adaptation mechanism of pseudomonads in environments with high osmolality [57, 58].

Importantly, we found that the *P. veronii* population developed in all non-sterile environments, more extensively when toluene was added as specific growth substrate, but also in its absence. This is in contrast to the much cited incapacity of inoculants to grow and be metabolically active in natural non-sterile materials, which may lead to bioaugmentation failures [9–11]. However, other reports have indicated adaptation, growth and survival of inoculants purposed for bioremediation at least in realistic non-sterile microcosms [48, 59–61]. We acknowledge that microcosms are not the same as field experiments, but at least this shows that many bacteria with potentially interesting properties for bioremediation are capable of establishing in

complex microbial ecosystems. Those studies and including our own results indicated strain level adaptation to the specific environment (of the microcosm) and the carbon source [48, 59–61]. The variety of different responses, unfortunately, precludes some sort of global ‘one-for-all’ interpretation of adaptive characters that inoculants would need to have in order to be ‘fit’ for their intended deployment, except a recurrent signature of differentially expressed chemotactic and flagellar biosynthesis genes. We aimed to find additional conditions under which *P. veronii* would not be able to adapt and proliferate, but unfortunately, we could not really exploit this at transcriptomic level. Growth in Junction material without added toluene was clearly not favorable for *P. veronii*, but the transcriptomic response from 1 h exposure in Junction with toluene (in which it could grow) did not particularly show signs of poor adaptation. Except for increased induction of stress defense systems and changes in expression of genes for respiratory activity, there were no particular signs of physiological breakdown in cells in Junction. The only other environment, which did not lead to strong population development of *P. veronii* on toluene was Clay. We think that the most likely causes for population loss here were predation and, possibly, competition of native soil microbiota for toluene. Unfortunately, we did not manage to isolate sufficient RNA for transcriptome analysis from *P. veronii* during exponential growth in Clay. However, during the 1-h transition in Clay we observed a decrease of expression of genes for flagellar assembly and motility, which was not present in Silt or Sand. Flagella allow bacteria to explore their environment, to search for nutrients and to escape from predators or adverse conditions, but activation of flagellar genes have also been implicated in the solvent stress tolerance response, where the flagellar export apparatus is used to export other proteins unrelated to flagellar assembly [62]. Overall, the role of flagella synthesis during soil adaptation remains unclear. Growing *S. wittichii* RW1 in sand [48], or clay with dibenzofuran [60] diminished expression of flagella synthesis genes. Both *S. wittichii* RW1, *Artrobacter chlorophenolicus* A6 and *P. veronii* reduce expression of flagella synthesis under solute and matric stress [17], and similar behaviour has been detected in other bacteria under water stress [58, 63]. In contrast, *P. putida* KT2440 exposed to water stress on ceramic plates did not show significant difference in flagellar gene expression [49], whereas *P. veronii* during growth in Silt with toluene up-regulated flagellar gene synthesis (Table 5). This contrasting behaviour may indicate strain-specific finetuning of flagellar expression in relation to available energy levels. Perhaps, therefore, the diminished expression of *P. veronii* flagellar genes during the 1-h adaptation to Clay was a result of redirected energy resources. The

resulting reduced cell motility could then have favored grazing by protozoa.

The analysis of gene expression during growth of *P. veronii* allowed us to identify further differences in soils versus liquid. It has been reported that in bulk soils carbon rather than nutrients such as N, S, P, and Fe is the limiting factor for bacteria growth and fitness [64, 65]. In contrast, in hydrocarbon- or PAH-polluted soils the addition of macro- (N, P, S, K) and micro-elements (Fe) is usually practiced to enhance bacterial activity, because of the excess carbon posed by the pollution [66–69]. Expression of functions related to nutrient scavenging by *P. veronii* cells in soils suggests that they quickly perceive nutrient limitation and can adapt to some extent. For example, the observed higher expression of genes involved in sarcosine recycling may have been a response to limiting nitrogen. In addition, perceived nitrogen limitation may have led to increased turnover of branched-chain amino acids and asparagine, plus leading to different usage of the urea pathways. Growth in soils may also have led to sulfur limitation and to induction of increased assimilation of sulfite from sulfate, and catabolism of cysteine and taurine. These examples indicate that availability of nutrients is important for establishment of inoculants in polluted soils.

Conclusion

In conclusion, our study clearly showed *P. veronii* adaptation, growth, and survival in different non-sterile soils and contaminated material. Although we externally added toluene as a specific carbon and energy source for the proliferation of *P. veronii*, its previously observed abundance and activity in contaminated sites [19, 70] indicate its capacity to survive under field conditions. These are important observations because they contradict the regular notion of poor growth and survival for exogenous strains in complex microbial ecosystems [71], and thus, provide basis to better select inoculants for applications in bioaugmentation. Under the tested conditions, we did not find a single “soil-transcriptomic” program but instead identified and highlighted both a core of commonly as well as specifically induced functions (many of which consist of uncharacterized proteins) in soils that contribute to the strain’s adaptation under different conditions. The strain expressed a remarkably robust metabolic program during growth, which was maintained irrespectively of its environment. We did not identify critical factors associated with the failure of the strain upon inoculation beyond potential predation, possible substrate competition and signs of nutrient limitations in later growth phases.

At this point, our comprehension (and that of many others) is necessarily a ‘narrative’ of understanding how cells adapt and grow, concluded from the conglomerate of

global (e.g., GO and COG) analysis as well as that of individual gene annotations. However, specific cellular reactions are clearly different between typical liquid cultures and in soil. Therefore, it seems crucial to us to study strain behaviour under the conditions of the expected complex environments (i.e., soil, gut, skin), and not in standard liquid culture. In the future, this knowledge may help to better predict the success of inoculants.

Abbreviations

EXPO: Exponential growth phase; STAT: Stationary phase; RNA-seq: Reverse-transcribed ribosomal RNA depleted library sequencing; GO: Gene ontology; COG: Cluster of orthologous groups; CFUs: Colony-forming units; GWC: Gravimetric water content; APM: Artificial porous matrix; OD: Optical density (culture turbidity); FDR: False discovery rate

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40793-021-00378-x>.

Additional file 1: Supplementary Figure 1. Comparison of growth rates on toluene of *P. veronii* wild-type and the *P. veronii* gfp-tagged variant. **Supplementary Figure 2.** Background soil microbiota counts. **Supplementary Figure 3.** Cell washing recoveries of wild-type *P. veronii* cells inoculated to soils. **Supplementary Figure 4.** Background growth of *P. veronii* miniTn5:gfp on soil organic carbon or soil microbiota on toluene.

Additional file 2: Table S1. Summary of total mapped reads of inoculated microcosms exposed to toluene. **Table S2.** Subset of genes whose expression levels responded to inoculation into glass beads (APM) versus liquid (FDR < 0.05). **Table S3.** Subset differentially expressed genes of *P. veronii* 1YdBTEX2 in all soils vs controls (Liquid&APM). **Table S4.** Relevant KEGG reactions with outlier expression in exponential phase conditions. **Table S5.** Biochemical term enrichment in the exponential phase outlier reactions compared to all metabolic reactions. **Table S6.** Commonly differentially expressed genes in Liquid, glass beads (APM) and Silt in stationary phase. **Table S7.** Subset of commonly differentially expressed genes in glass beads (APM) and Silt in stationary phase.

Additional file 3: DATA 1. Complete list of differentially expressed genes in APM-1H versus liquid-1H (FDR < 0.05). **DATA 2.** Enriched GO terms among the significantly differentially downregulated genes in cells growing on glass beads (APM) versus liquid media after 1 h inoculation. **DATA 3.** Complete list of differentially expressed genes in Sand-1H versus APM-1H (FDR < 0.05). **DATA 4.** Complete list of differentially expressed genes in Silt -1H versus APM-1H (FDR < 0.05). **DATA 5.** Complete list of differentially expressed genes in Clay-1H versus APM-1H (FDR < 0.05). **DATA 6.** Enriched GO terms among the significantly differentially upregulated genes in Sand versus APM after 1 h contact. **DATA 7.** Enriched GO terms among the significantly differentially expressed upregulated genes in Silt versus glass beads (APM) after 1 h contact. **DATA 8.** Enriched GO terms among the significantly differentially expressed downregulated genes in Silt versus glass beads (APM) after 1 h contact. **DATA 9.** Enriched GO terms among the significantly differentially expressed upregulated genes in Clay versus glass beads (APM) after 1 h contact. **DATA 9.** Enriched GO terms among the significantly differentially expressed downregulated genes in Clay versus glass beads (APM) after 1 h contact. **DATA 11.** Differentially expressed genes in all natural soils-1H vs controls-1 h (ANOVA, FDR < 0.05). **DATA 12.** Complete list of differentially expressed genes in Junction-1H vs APM-1H (FDR < 0.05). **DATA 13.** Commonly differentially expressed genes in soils and Junction soils after 1 h contact versus control conditions -1H (ANOVA, FDR < 0.05). **DATA 14.** Unique differentially expressed genes in Junction-1H vs APM-1H. **DATA 15.** Enriched GO terms among the significantly differentially expressed upregulated genes in Junction versus glass beads (APM) after 1 h contact. **DATA 16.** Enriched GO terms among the significantly differentially expressed downregulated genes in Junction versus glass beads

(APM) after 1 h contact. **DATA 17.** Commonly differentially expressed genes in *Pseudomonas veronii* cells growing exponentially in Liquid, Sand and Silt vs same material -1H (FDR < 0.05). **DATA 18.** Enriched GO terms among the significantly differentially expressed upregulated genes in LIQUID-EXPO vs LIQUID-1 h. **DATA 19.** Enriched GO terms among the significantly differentially expressed downregulated genes in LIQUID-EXPO vs LIQUID-1 h. **DATA 20.** Enriched GO terms among the significantly differentially expressed upregulated genes in SILT-EXPO vs SILT-1 h. **DATA 21.** Enriched GO terms among the significantly differentially expressed downregulated genes in SILT-EXPO vs SILT-1 h. **DATA 22.** Enriched GO terms among the significantly differentially expressed upregulated genes in JUNCTION-EXPO vs JUNCTION-1 h. **DATA 23.** Enriched GO terms among the significantly differentially expressed downregulated genes in JUNCTION-EXPO vs JUNCTION-1 h. **DATA 24.** Complete list of differentially expressed genes in SILT-EXPO vs SILT-1H (FDR < 0.05). **DATA 25.** Complete list of differentially expressed genes in Junction-EXPO vs JN-1H (FDR < 0.05). **DATA 26.** Commonly differentially expressed genes in APM and Silt in stationary phase (FDR < 0.05). **DATA 27.** Unique differentially expressed genes Silt in stationary phase (FDR < 0.05). **DATA 28.** Unique differentially expressed genes in APM-STAT vs APM-1H (FDR < 0.05). **DATA 29.** Enriched GO terms among the significantly differentially expressed upregulated genes in SILT-STAT vs SILT-1 h. **DATA 30.** Enriched GO terms among the significantly differentially expressed downregulated genes in SILT-STAT vs SILT-1 h. **DATA 31.** Enriched GO terms among the significantly differentially expressed upregulated genes in AS-STAT vs AS-1 h. **DATA 32.** Enriched GO terms among the significantly differentially expressed downregulated genes in AS-STAT vs AS-1 h. **DATA.** Enriched GO terms among the significantly differentially expressed upregulated genes in LIQUID-STAT vs LIQUID-1 h. **DATA.** Enriched GO terms among the significantly differentially expressed downregulated genes in LIQUID-STAT vs LIQUID-1 h.

Acknowledgements

The authors thank Loïc Sauvan, Renyi Li and Senka Čaušević for their help in the initial phases of this project.

Authors' contributions

M. M. and V. S. carried out experimental work. M. M., N. H. and J.M. conducted data analysis. M. M. and J. M. wrote the manuscript. M. M., N. H., V. S. and J. M. commented on the final text. J. M. raised funding. The author(s) read and approved the final manuscript.

Funding

This work was supported by the grant *MicroScapesX* from the Swiss Initiative in Systems Biology *SystemsX.ch*.

Availability of data and materials

The raw sequence reads of the RNA-seq data sets are available from the Short Read Archive under Bioproject number PRJNA682712.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors have read and approved the final version for submission.

Competing interests

The authors declare no competing interests.

Received: 13 December 2020 Accepted: 10 April 2021

Published online: 29 April 2021

References

- Ahn YB, Liu F, Fennell DE, Haggblom MM. Biostimulation and bioaugmentation to enhance dechlorination of polychlorinated dibenzo-*p*-dioxins in contaminated sediments. *FEMS Microbiol Ecol.* 2008;66(2):271–81. <https://doi.org/10.1111/j.1574-6941.2008.00557.x>.

- Cycon M, Mrozik A, Piotrowska-Seget Z. Bioaugmentation as a strategy for the remediation of pesticide-polluted soil: a review. *Chemosphere.* 2017;172:52–71. <https://doi.org/10.1016/j.chemosphere.2016.12.129>.
- Garg N, Lata P, Jit S, Sangwan N, Singh AK, Dwivedi V, et al. Laboratory and field scale bioremediation of hexachlorocyclohexane (HCH) contaminated soils by means of bioaugmentation and biostimulation. *Biodegradation.* 2016;27(2-3):179–93. <https://doi.org/10.1007/s10532-016-9765-6>.
- Megharaj M, Wittich R-M, Blasco R, Pieper DH, Timmis KN. Superior survival and degradation of dibenzo-*p*-dioxin and dibenzofuran in soil by soil-adapted *Sphingomonas* sp. strain RW1. *Appl Microbiol Biotechnol.* 1997;48(1):109–14. <https://doi.org/10.1007/s002530051024>.
- Atashgahi S, Sanchez-Andrea I, Heipieper HJ, van der Meer JR, Stams AJM, Smidt H. Prospects for harnessing biocide resistance for bioremediation and detoxification. *Science.* 2018;360(6390):743–6. <https://doi.org/10.1126/science.aar3778>.
- El Fantroussi S, Agathos SN. Is bioaugmentation a feasible strategy for pollutant removal and site remediation? *Curr Opin Microbiol.* 2005;8(3):268–75. <https://doi.org/10.1016/j.mib.2005.04.011>.
- Mrozik A, Piotrowska-Seget Z. Bioaugmentation as a strategy for cleaning up of soils contaminated with aromatic compounds. *Microbiol Res.* 2010;165(5):363–75. <https://doi.org/10.1016/j.micres.2009.08.001>.
- Tyagi M, da Fonseca MM, de Carvalho CC. Bioaugmentation and biostimulation strategies to improve the effectiveness of bioremediation processes. *Biodegradation.* 2011;22(2):231–41. <https://doi.org/10.1007/s10532-010-9394-4>.
- McKew BA, Coulon F, Yakimov MM, Denaro R, Genovese M, Smith CJ, et al. Efficacy of intervention strategies for bioremediation of crude oil in marine systems and effects on indigenous hydrocarbonoclastic bacteria. *Environ Microbiol.* 2007;9(6):1562–71. <https://doi.org/10.1111/j.1462-2920.2007.01277.x>.
- Tchelet R, Meckenstock R, Steidle P, van der Meer JR. Population dynamics of an introduced bacterium degrading chlorinated benzenes in a soil column and in sewage sludge. *Biodegradation.* 1999;10(2):113–25. <https://doi.org/10.1023/A:1008368006917>.
- de Lorenzo V, Pieper D, Ramos JL. From the test tube to the environment - and back. *Environ Microbiol.* 2013;15(1):6–11. <https://doi.org/10.1111/j.1462-2920.2012.02896.x>.
- Johri AK, Dua M, Tuteja D, Saxena R, Saxena DM, Lal R. Degradation of α -, β -, γ - and δ -hexachlorocyclohexane by *Sphingomonas paucimobilis*. *Biotechnol Lett.* 1998;20(9):885–7. <https://doi.org/10.1023/A:1005323811769>.
- Lawley TD, Clare S, Walker AW, Stares MD, Connor TR, Raisen C, et al. Targeted restoration of the intestinal microbiota with a simple, defined bacteriotherapy resolves relapsing *Clostridium difficile* disease in mice. *PLoS Pathog.* 2012;8(10):e1002995. <https://doi.org/10.1371/journal.ppat.1002995>.
- Endt K, Stecher B, Chaffron S, Slack E, Tchitchek N, Benecke A, et al. The microbiota mediates pathogen clearance from the gut lumen after non-typhoidal *Salmonella* diarrhea. *PLoS Pathog.* 2010;6(9):e1001097. <https://doi.org/10.1371/journal.ppat.1001097>.
- van Veen JA, van Overbeek LS, van Elsas JD. Fate and activity of microorganisms introduced into soil. *Microbiol Mol Biol Rev.* 1997;61(2):121–35. <https://doi.org/10.1128/61.2.121-135.1997>.
- Coronado E, Valtat A, van der Meer JR. *Sphingomonas wittichii* RW1 gene reporters interrogating the dibenzofuran metabolic network highlight conditions for early successful development in contaminated microcosms. *Environ Microbiol Rep.* 2015;7(3):480–8. <https://doi.org/10.1111/1758-2229.12276>.
- Moreno-Forero SK, Rojas E, Beggah S, van der Meer JR. Comparison of differential gene expression to water stress among bacteria with relevant pollutant-degradation properties. *Environ Microbiol Rep.* 2016;8(1):91–102. <https://doi.org/10.1111/1758-2229.12356>.
- Roggo C, Coronado E, Moreno-Forero SK, Harshman K, Weber J, van der Meer JR. Genome-wide transposon insertion scanning of environmental survival functions in the polycyclic aromatic hydrocarbon degrading bacterium *Sphingomonas wittichii* RW1. *Environ Microbiol.* 2013;15(10):2681–95. <https://doi.org/10.1111/1462-2920.12125>.
- Junca H, Pieper DH. Functional gene diversity analysis in BTEX contaminated soils by means of PCR-SSCP DNA fingerprinting: comparative diversity assessment against bacterial isolates and PCR-DNA clone libraries. *Environ Microbiol.* 2004;6(2):95–110. <https://doi.org/10.1046/j.1462-2920.2003.00541.x>.
- Morales M, Sentchilo V, Bertelli C, Komljenovic A, Kryuchkova-Mostacci N, Bourdilloud A, et al. The genome of the toluene-degrading *Pseudomonas*

- veronii* strain 1YdBTEX2 and its differential gene expression in contaminated sand. *PLoS One*. 2016;11(11):e0165850. <https://doi.org/10.1371/journal.pone.0165850>.
21. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25(1):25–9. <https://doi.org/10.1038/75556>.
 22. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res*. 2016;44(D1):D457–62. <https://doi.org/10.1093/nar/gkv1070>.
 23. Hadadi N, Pandey V, Chiappino-Pepe A, Morales M, Gallart-Ayala H, Mehl F, et al. Mechanistic insights into bacterial metabolic reprogramming from omics-integrated genome-scale models. *NPJ Syst Biol Appl*. 2020;6(1):1. <https://doi.org/10.1038/s41540-019-0121-4>.
 24. Sentchilo V, Czechowska K, Pradervand N, Minoia M, Miyazaki R, van der Meer JR. Intracellular excision and reintegration dynamics of the ICE_{cl} genomic island of *Pseudomonas knackmussii* sp. strain B13. *Mol Microbiol*. 2009;72(5):1293–306. <https://doi.org/10.1111/j.1365-2958.2009.06726.x>.
 25. Gerhardt P, Murray RGE, Costilow RN, Nester EW, Wood WA, Krieg NR, Briggs Phillips G (Eds): manual of methods for general bacteriology. Washington, D.C.: American Society for Microbiology; 1981.
 26. Weinbauer MG, Beckmann C, Hofle MG. Utility of green fluorescent nucleic acid dyes and aluminum oxide membrane filters for rapid epifluorescence enumeration of soil and sediment bacteria. *Appl Environ Microbiol*. 1998;64(12):5000–3. <https://doi.org/10.1128/AEM.64.12.5000-5003.1998>.
 27. Lima-Morales D, Jauregui R, Camarinha-Silva A, Geffers R, Pieper DH, Vilchez-Vargas R. Linking microbial community and catabolic gene structures during the adaptation of three contaminated soils under continuous long-term pollutant stress. *Appl Environ Microbiol*. 2016;82(7):2227–37. <https://doi.org/10.1128/AEM.03482-15>.
 28. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods*. 2012;9(4):357–9. <https://doi.org/10.1038/nmeth.1923>.
 29. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. Genome project data processing S: the sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9. <https://doi.org/10.1093/bioinformatics/btp352>.
 30. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166–9. <https://doi.org/10.1093/bioinformatics/btu638>.
 31. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139–40. <https://doi.org/10.1093/bioinformatics/btp616>.
 32. Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, et al. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res*. 2008;36(10):3420–35. <https://doi.org/10.1093/nar/gkn176>.
 33. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, et al. Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc*. 2007;2(10):2366–82. <https://doi.org/10.1038/nprot.2007.324>.
 34. Agren R, Liu L, Shoaie S, Vongsangnak W, Nookaew I, Nielsen J. The RAVEN toolbox and its use for generating a genome-scale metabolic model for *Penicillium chrysogenum*. *PLoS Comput Biol*. 2013;9(3):e1002980. <https://doi.org/10.1371/journal.pcbi.1002980>.
 35. Darzi Y, Letunic I, Bork P, Yamada T. iPath3.0: interactive pathways explorer v3. *Nucleic Acids Res*. 2018;46(W1):W510–3. <https://doi.org/10.1093/nar/gky299>.
 36. Karakoc C, Singer A, Johst K, Harms H, Chatzinotas A. Transient recovery dynamics of a predator-prey system under press and pulse disturbances. *BMC Ecol*. 2017;17(1):13. <https://doi.org/10.1186/s12898-017-0123-2>.
 37. Rønn R, McCaig AE, Griffiths BS, Prosser JL. Impact of protozoan grazing on bacterial community structure in soil microcosms. *Appl Environ Microbiol*. 2002;68(12):6094–105. <https://doi.org/10.1128/AEM.68.12.6094-6105.2002>.
 38. Wanjugi P, Harwood VJ. The influence of predation and competition on the survival of commensal and pathogenic fecal bacteria in aquatic habitats. *Environ Microbiol*. 2013;15(2):517–26. <https://doi.org/10.1111/j.1462-2920.2012.02877.x>.
 39. Shah P, Swiatlo E. A multifaceted role for polyamines in bacterial pathogens. *Mol Microbiol*. 2008;68(1):4–16. <https://doi.org/10.1111/j.1365-2958.2008.06126.x>.
 40. Rodionov DA, Dubchak IL, Arkin AP, Alm EJ, Gelfand MS. Dissimilatory metabolism of nitrogen oxides in bacteria: comparative reconstruction of transcriptional networks. *PLoS Comput Biol*. 2005;1(5):e55. <https://doi.org/10.1371/journal.pcbi.0010055>.
 41. Yang Z, Lu CD. Functional genomics enables identification of genes of the arginine transaminase pathway in *Pseudomonas aeruginosa*. *J Bacteriol*. 2007;189(11):3945–53. <https://doi.org/10.1128/JB.00261-07>.
 42. Willsey GG, Wargo MJ. Sarcosine catabolism in *Pseudomonas aeruginosa* is transcriptionally regulated by SouR. *J Bacteriol*. 2016;198(2):301–10. <https://doi.org/10.1128/JB.00739-15>.
 43. Nachin L, Nannmark U, Nystrom T. Differential roles of the universal stress proteins of *Escherichia coli* in oxidative stress resistance, adhesion, and motility. *J Bacteriol*. 2005;187(18):6265–72. <https://doi.org/10.1128/JB.187.18.6265-6272.2005>.
 44. Schmidt-Rohr K, Mao JD, Olk DC. Nitrogen-bonded aromatics in soil organic matter and their implications for a yield decline in intensive rice cropping. *Proc Natl Acad Sci U S A*. 2004;101(17):6351–4. <https://doi.org/10.1073/pnas.0401349101>.
 45. Mozejko-Ciesielska J, Pokoj T, Ciesielski S. Transcriptome remodeling of *Pseudomonas putida* KT2440 during mcl-PHAs synthesis: effect of different carbon sources and response to nitrogen stress. *J Ind Microbiol Biotechnol*. 2018;45(6):433–46. <https://doi.org/10.1007/s10295-018-2042-4>.
 46. Kim HY, Schlichtman D, Shankar S, Xie Z, Chakrabarty AM, Kornberg A. Alginate, inorganic polyphosphate, GTP and ppGpp synthesis co-regulated in *Pseudomonas aeruginosa*: implications for stationary phase survival and synthesis of RNA/DNA precursors. *Mol Microbiol*. 1998;27(4):717–25. <https://doi.org/10.1046/j.1365-2958.1998.00702.x>.
 47. Güvener ZT, Tifrea DF, Harwood CS. Two different *Pseudomonas aeruginosa* chemosensory signal transduction complexes localize to cell poles and form and remould in stationary phase. *Mol Microbiol*. 2006;61(1):106–18. <https://doi.org/10.1111/j.1365-2958.2006.05218.x>.
 48. Moreno-Forero SK, van der Meer JR. Genome-wide analysis of *Shingomonas wittichii* RW1 behaviour during inoculation and growth in contaminated sand. *ISME J*. 2015;9(1):150–65. <https://doi.org/10.1038/ismej.2014.101>.
 49. Gülez G, Dechesne A, Workman CT, Smets BF. Transcriptome dynamics of *Pseudomonas putida* KT2440 under water stress. *Appl Environ Microbiol*. 2012;78(3):676–83. <https://doi.org/10.1128/AEM.06150-11>.
 50. Johnson DR, Coronado E, Moreno-Forero SK, Heipieper HJ, van der Meer JR. Transcriptome and membrane fatty acid analyses reveal different strategies for responding to permeating and non-permeating solutes in the bacterium *Shingomonas wittichii*. *BMC Microbiol*. 2011;11(1):250. <https://doi.org/10.1186/1471-2180-11-250>.
 51. Miedema R. Applications of micromorphology of relevance to agronomy. *Adv Agron*. 1997;59:119–69. [https://doi.org/10.1016/S0065-2113\(08\)60054-1](https://doi.org/10.1016/S0065-2113(08)60054-1).
 52. Alvarez-Ortega C, Harwood CS. Responses of *Pseudomonas aeruginosa* to low oxygen indicate that growth in the cystic fibrosis lung is by aerobic respiration. *Mol Microbiol*. 2007;65(1):153–65. <https://doi.org/10.1111/j.1365-2958.2007.05772.x>.
 53. Ghiglione JF, Gourbiere F, Potier P, Philippot L, Lensi R. Role of respiratory nitrate reductase in ability of *Pseudomonas fluorescens* YT101 to colonize the rhizosphere of maize. *Appl Environ Microbiol*. 2000;66(9):4012–6. <https://doi.org/10.1128/AEM.66.9.4012-4016.2000>.
 54. Rediers H, Vanderleyden J, De Mot R. Nitrate respiration in *Pseudomonas stutzeri* A15 and its involvement in rice and wheat root colonization. *Microbiol Res*. 2009;164(4):461–8. <https://doi.org/10.1016/j.micres.2007.03.003>.
 55. Borer B, Tecon R, Or D. Spatial organization of bacterial populations in response to oxygen and carbon counter-gradients in pore networks. *Nat Commun*. 2018;9(1):769. <https://doi.org/10.1038/s41467-018-03187-y>.
 56. Csonka LN. Physiological and genetic responses of bacteria to osmotic stress. *Microbiol Rev*. 1989;53(1):121–47. <https://doi.org/10.1128/MR.53.1.121-147.1989>.
 57. Chang WS, van de Mortel M, Nielsen L, Nino de Guzman G, Li X, Halverson LJ. Alginate production by *Pseudomonas putida* creates a hydrated microenvironment and contributes to biofilm architecture and stress tolerance under water-limiting conditions. *J Bacteriol*. 2007;189(22):8290–9. <https://doi.org/10.1128/JB.00727-07>.
 58. Freeman BC, Chen C, Yu X, Nielsen L, Peterson K, Beattie GA. Physiological and transcriptional responses to osmotic stress of two *Pseudomonas syringae* strains that differ in epiphytic fitness and osmotolerance. *J Bacteriol*. 2013;195(20):4742–52. <https://doi.org/10.1128/JB.00787-13>.
 59. Chai B, Tsoi T, Sallach JB, Liu C, Landgraf J, Bezdek M, et al. Bioavailability of clay-adsorbed dioxin to *Shingomonas wittichii* RW1 and its associated genome-wide shifts in gene expression. *Sci Total Environ*. 2020;712:135525. <https://doi.org/10.1016/j.scitotenv.2019.135525>.

60. Chai B, Tsoi TV, Iwai S, Liu C, Fish JA, Gu C, et al. *Sphingomonas wittichii* strain RW1 genome-wide gene expression shifts in response to dioxins and clay. *PLoS One*. 2016;11(6):e0157008. <https://doi.org/10.1371/journal.pone.0157008>.
61. Jin HM, Jeong HI, Kim KH, Hahn Y, Madsen EL, Jeon CO. Genome-wide transcriptional responses of *Alteromonas naphthalenivorans* SN2 to contaminated seawater and marine tidal flat sediment. *Sci Rep*. 2016;6(1):21796. <https://doi.org/10.1038/srep21796>.
62. Segura A, Hurtado A, Duque E, Ramos JL. Transcriptional phase variation at the flhB gene of *Pseudomonas putida* DOT-T1E is involved in response to environmental changes and suggests the participation of the flagellar export system in solvent tolerance. *J Bacteriol*. 2004;186(6):1905–9. <https://doi.org/10.1128/JB.186.6.1905-1909.2004>.
63. Fida TT, Breugelmanns P, Lavigne R, Coronado E, Johnson DR, van der Meer JR, et al. Exposure to solute stress affects genome-wide expression but not the polycyclic aromatic hydrocarbon-degrading activity of *Sphingomonas* sp. strain LH128 in biofilms. *Appl Environ Microbiol*. 2012;78(23):8311–20. <https://doi.org/10.1128/AEM.02516-12>.
64. Hobbie JE, Hobbie EA. Microbes in nature are limited by carbon and energy: the starving-survival lifestyle in soil and consequences for estimating microbial rates. *Front Microbiol*. 2013;4:324.
65. Mirleau P, Wogelius R, Smith A, Kertesz MA. Importance of organosulfur utilization for survival of *Pseudomonas putida* in soil and rhizosphere. *Appl Environ Microbiol*. 2005;71(11):6571–7. <https://doi.org/10.1128/AEM.71.11.6571-6577.2005>.
66. Mills MA, Bonner JS, Page CA, Autenrieth RL. Evaluation of bioremediation strategies of a controlled oil release in a wetland. *Mar Pollut Bull*. 2004;49(5-6):425–35. <https://doi.org/10.1016/j.marpolbul.2004.02.027>.
67. da Silva AC, de Oliveira FJ, Bernardes DS, de Franca FP. Bioremediation of marine sediments impacted by petroleum. *Appl Biochem Biotechnol*. 2009;153(1-3):58–66. <https://doi.org/10.1007/s12010-008-8457-z>.
68. Yu KS, Wong AH, Yau KW, Wong YS, Tam NF. Natural attenuation, biostimulation and bioaugmentation on biodegradation of polycyclic aromatic hydrocarbons (PAHs) in mangrove sediments. *Mar Pollut Bull*. 2005;51(8-12):1071–7. <https://doi.org/10.1016/j.marpolbul.2005.06.006>.
69. Kalantary RR, Mohseni-Bandpi A, Esrafilii A, Nasserii S, Ashmagh FR, Jorfi S, et al. Effectiveness of biostimulation through nutrient content on the bioremediation of phenanthrene contaminated soil. *J Environ Health Sci Eng*. 2014;12(1):143. <https://doi.org/10.1186/s40201-014-0143-1>.
70. Junca H, Plumeier I, Hecht HJ, Pieper DH. Difference in kinetic behaviour of catechol 2,3-dioxygenase variants from a polluted environment. *Microbiology*. 2004;150(12):4181–7. <https://doi.org/10.1099/mic.0.27451-0>.
71. Jones ML, Rivett DW, Pascual-Garcia A, Bell T. Productive bacterial communities exclude invaders. *bioRxiv*. 2019. <https://doi.org/10.1101/2019.12.18.881102>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)



CHAPTER 3

Identification of genes involved in soil adaptation in *Pseudomonas veronii* 1YdBTEX2

Marian Morales, Vladimir Sentchilo, Noushin Hadadi, Dominique Vuarambon and Jan Roelof van der Meer

Department of Fundamental Microbiology, University of Lausanne, 1015 Lausanne, Switzerland

Abstract

Pseudomonas veronii 1YdBTEX2 is a promising candidate bacterium to be used in bioaugmentation of polluted soil by oil spills given its capacity to use monoaromatic compounds as a sole carbon and energy source. In order to describe and understand the capacity of the microorganism to survive in soils we used a genome-wide transposon scanning approach. Two independent mutant libraries were generated with a total of 2.7×10^5 unique insertions. Libraries were grown in multiple growth cycles for an average of 50 generations either in batch liquid medium or soil microcosms (sand or silt) in presence of toluene. From the positions and proportions of the inserted transposed kanamycin resistance gene in each sample compared to an *in silico* randomly distributed library we identified potential generally essential genes. Soil growth and survival functions were further assessed from loss or gain of the relative abundance of insertion mutations over generation times. Our data showed relatively sparse overlap between soils, but pointed to crucial functions of urea and short-chain fatty acid metabolism, oxidative stress defense and nutrient/element transport for soil growth. Interestingly, mutants inactivating flagella biosynthesis and motility had a strong fitness gain in soil incubations, but seemed more sensitive to predation and extinction in the long run. Comparison to transcriptomic data strengthened the hypothesis that *P. veronii* can maintain central metabolism and growth, irrespective of and without major reprogramming to the soil environment.

Introduction

Bioaugmentation to accelerate or extend contaminant degradation rates is a potentially useful technology, but relies on ecological properties of inoculant bacteria (and resident microbiota) that are generally ill-understood. Previous studies conducted under standardized growth conditions that mimic environmental stresses in natural soils highlighted several mechanisms that bacteria have evolved to survive and proliferate in polluted soil. The most frequent mechanisms described in the literature include diverse osmoprotectant strategies to cope with changing water potential (Bojanovič et al., 2017; Moreno-Forero and van der Meer, 2015; Polonenko et al., 1986; Reva et al., 2006), the expression of multiple efflux systems capable of extruding heavy metals or toxic solvent molecules (Armalyté et al., 2019; Blanco et al., 2016; Martinez et al., 2009), the presence of denitrification pathways to survive under microaerophilic or anaerobic conditions and the capacity of producing iron-chelating molecules to thrive under low iron conditions (Mendonca et al., 2020). Previously, by global transcriptomics, others and we showed the complex multifactorial physiological and metabolic responses to cope with contaminants in environmental compartments, which represent a toxic threat and a carbon source at the same time (Domínguez-Cuevas et al., 2006; Krell et al., 2012;

Molina-Santiago et al., 2017; Moreno-Forero and van der Meer, 2015). Solvent degradation requires fine-tuning the cell's membrane fluidity, altering the composition of the cytoplasmic and outer membranes and the cell surface (Eberlein et al., 2018; Johnson et al., 2011; Murínová and Dercová, 2014; Pieper and Reineke, 2000; Sardesai and Bhosle, 2002), and activating multiple stress-response systems to counteract the reactive oxygen species generated during the catabolism of toxic compounds (Akkaya et al., 2018; Chavarría et al., 2013; Nikel et al., 2021). Soil survival and growth may require specific adaptation programs that seemingly have little overlap between different bacteria (Moreno-Forero and van der Meer, 2016), making it difficult to discern *a priori* the potential success for strain colonization of a non-sterile polluted environment based on its genome information.

Inferring the global importance of gene functions to environmental survival can be facilitated by transposon sequencing. Transposon sequencing is based on the activity of transposases to insert DNA fragments with recognizable ends (Goryshin et al., 1998; Green et al., 2012). Whereas natural transposon activity is subject to extensive regulatory control by the host cell, synthetic hyperactive transposon constructs allow the production of bacteria libraries, in which each member is a mutant generated by the insertion of a marker gene delivered by that transposon. This so-called *transposon mutagenesis* can be coupled with next-generation sequencing technologies to identify and quantify all insertions *en masse* in a mutant library by high-throughput sequencing (hence, also known as Tn-seq) (Liu et al., 2016; van Opijnen and Camilli, 2013; Sternon et al., 2018; Wong et al., 2016). Tn-seq is a powerful tool that has been utilized to study specific metabolic pathways (Nguyen et al., 2016), gene essentiality (Gallagher et al., 2020), mechanisms of pathogenesis (Harrison et al., 2020; Jacobs et al., 2003; Smith and Enquist, 1999), or genes essential for survival and fitness in soil or water stress (Coronado et al., 2014; Roggo et al., 2013). The basic concept in Tn-seq is that if a marker insertion has inactivated a gene, and the gene function is essential or has a fitness effect, the proportion of the mutant in the overall library will change over time with respect to the original library, or even disappear from the meta-population of all mutants. Problematic for many of the Tn-seq applications is that the library preparation itself requires selection for the marker insertion and thus some growth, resulting in insertions in essential genes being counter-selected and absent at the start (Gallagher et al., 2020). Furthermore, judging fitness loss from proportional insertion abundances may be complicated by polar effects and biased insertion specificities.

The major objectives of the underlying study were to identify genes by Tn-seq implicated in environmental survival of the bacterium *Pseudomonas veronii* 1YdBTEX2 and better understand its growth strategies in soils under pollution stress. *P. veronii* has been mentioned as a candidate strain to use for bioaugmentation of contaminated soils because of its natural presence in contaminated sites (Junca and Pieper, 2004) and for its metabolic capacity

to utilize monoaromatic compounds as a sole carbon and energy source (de Lima-Morales et al., 2013, 2016). Using wild-type strains isolated from polluted soils to help restoring environments with similar pollution conditions may have the advantage that efficient degradation capacity is coupled to a selected inherent adaptability to stressful conditions. Unfortunately, this concept functions with moderate success as a result of frequent poor inoculant growth and survival in soils upon re-inoculation (Cunliffe et al., 2006; van Overbeek et al., 1995). There is thus a need to better understand the functions and characteristics of strains to optimally respond and adapt to their new growth environment, and to discern whether different bacteria have different strategies in this respect.

By using conjugation from a hyperactive transposon *Escherichia coli* donor and subsequent counterselection and -treatment, we produced two high-density independent genome-wide insertion libraries of *P. veronii* 1YdBTEX2. Mutant libraries were inoculated and cultured in up to 15 growth-dilution cycles in two different non-sterile soils supplemented with toluene and each run in quadruplicates, which was contrasted to toluene growth in liquid suspension. Cell densities of the *P. veronii* mutant meta-population in soil microcosms were determined after each cycle, and total DNA was extracted at three cycles of increasing generations of meta-population growth. Positions and abundances of marker insertion sites were determined by Illumina high-throughput sequencing, compared over time and as a function of treatment, and to randomized *in silico* insertions in order to estimate their fitness effects. Groups of fitness-affected genes were further examined based on their annotation and by means of pathway or clustered orthologies, and further compared to previous transcriptome data on the same organism (Morales et al., 2021). In contrast to our expectations from previous analyses (Roggo et al., 2013), common gene sets with measurable soil fitness for *P. veronii* were relatively scarce, suggesting streamlined metabolic adaptation irrespective of growth environment.

Results

Genome-wide random insertion libraries of P. veronii

Transposon libraries in *P. veronii* 1YdBTEX2 were produced by conjugating a hyperactive mini-transposable *aph* marker (Km-resistance) from an *E. coli* donor and subsequent depletion of the donor by phage treatment and recipient by antibiotic selection pressure. After cleaning and removing empty reads, between 119,289 and 178,769 unique insertion positions were detected in two independent *P. veronii* transposon libraries under all incubation conditions and sampling time points. By combining all the libraries, we detected 267,979 unique insertions (Figure 1A), corresponding to an average of 156 to 200, but a median of 2⁴ unique insertions per 10 kb (Figure 1B). Both libraries had few (30,081) overlaps, indicating the mostly random targets of the marker insertions on the genome. Log₂ per 10 kb insertion density distributions were similar among both libraries, but the spread still suggested a considerable difference in transposed marker target selectivity (Figure 1B). Insertion densities in Chromosome 1 (Chr1) were significantly lower than the other two *P. veronii* replicons (Chromosome 2 -Chr2 - and Plasmid - Plm, Figure 1C& D); probably because more genes with essential functions are located on Chr1. A slight 'smiley' effect of insertion densities in Chromosome 1 was visible, which is likely the result of chromosome replication effects (i.e., more target DNA being available close to the origin of replication).

Table 1. Libraries sequenced in this study

Series	Library	Starting library	Description	N ^b	Total raw reads	Mapped reads	% genes with insertions
Pv	Lib1	-	Starting library 1	-	2'788'655	1'924'227	83.8
	Lib2	-	Starting library 2	-	3'448'875	2'001'522	82.1
LIQ	T2	Lib1	1st cycle ^a	6	3'167'777	1'951'562	57.7
	T3		4th cycle	24	4'481'652	2'738'039	52.6
	T4		8th cycle	48	5'074'464	3'191'586	51.9
SAND	T2	Lib1	1st cycle	9	3'172'673	1'544'349	81.8
	T3		3th cycle	15	2'561'061	2'108'135	75.4
	T4		15th cycle	51	1'983'358	960'881	27.8
SILT	T2	Lib2	1st cycle	9	3'092'741	1'624'322	66.4
	T3		2nd cycle	12	2'875'037	1'779'345	65.6
	T4		3th cycle	15	1'580'122	1'073'113	54.5

a) Growth cycles in presence of toluene as carbon and energy substrate.

b) N, approximate number of generations of growth of *P. veronii* after that cycle.

The unique insertions in the starting libraries mapped to between 83.76% and 82.15% of the predicted *P. veronii* genes (Table 1, Supplementary Figure 1), with at least a single insertion. For the library samples transferred in liquid culture (LIQ), between 3 and 5 million clean reads were recovered, of which between 51-57% were mapped to the *P. veronii* genome (Table 1). The soil-derived transposon libraries (SAND and SILT) produced approximately 1.7-

times less mapped reads (Table 1). Libraries sequenced at later time points on average mapped to fewer genes, indicative for fitness effects of the growth conditions.

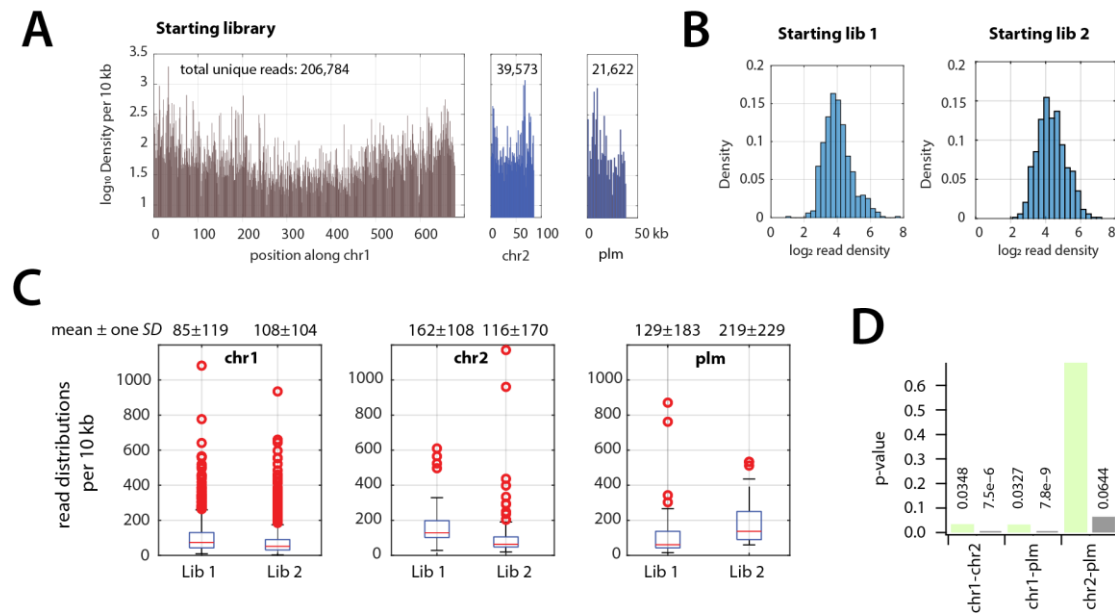


Figure 1. Gene insertion distributions in *P. veronii* transposon libraries. **A)** Log₁₀ transposon insertion density per 10 kb in one of the starting libraries plotted as a function of genomic location (chr 1, chromosome 1; chr 2, second replicon; plm, third replicon). Numbers in the upper part of each panel correspond to the total number of unique mapped reads across all libraries. **B)** Log₂ insertion density distributions per 10 kb in chr 1 in both starting libraries. **C)** Mean insertion densities are lower in chr 1 than in the second and third replicons (chr 2 and plm). Boxplots show read distributions per 10 kb for both libraries (Lib1, Lib 2) per replicon. The values reported at the top of each box correspond to the mean density ± one SD. **D)** P-values correspond to t-test comparisons (two-sided) of all grouped 10-kb densities per sample between Chr1: Chr2, Chr1: Plm, and Chr2: Plm, for both initial libraries.

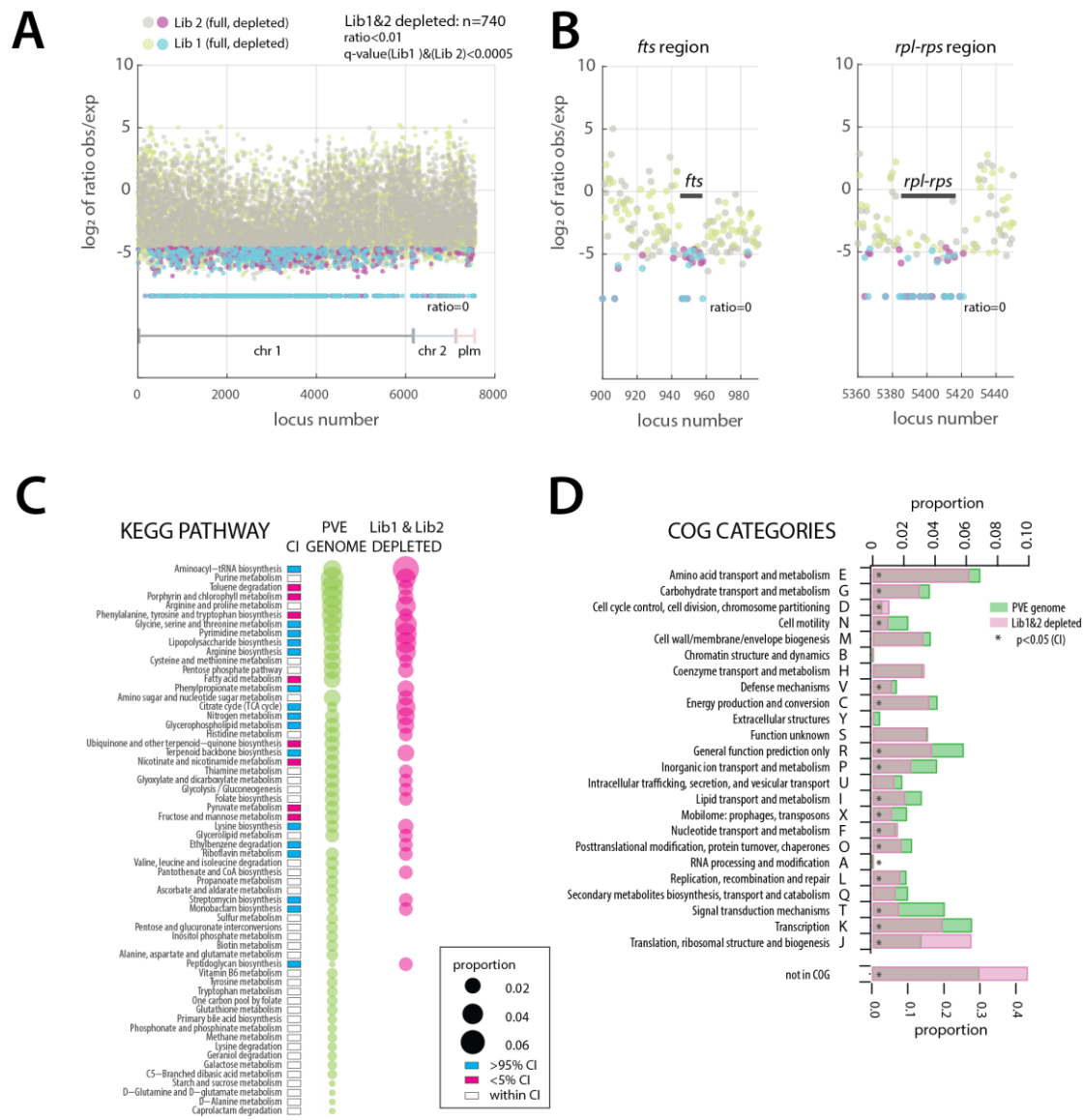


Figure 2. Global analysis of gene depletion in *P. veronii* starting transposon libraries. A) Scatter plot of the log₂ ratio of observed/expected reads per 10 kb in the starting libraries, summed per gene region (chr1, chr2, and plm; plotted according to locus_tag number). Each dot represents a gene; Lib 1, green; Lib 2, grey. Lib1&2-depleted: Genes with ratios in both starting libraries <0.01 and q-values <0.0005 for Lib 1 (cyan) and Lib 2 (magenta). **B)** Zoomed-in panels show gene regions surround the mostly essential cell division (*fts*) genes and ribosomal protein coding genes (*rpl-rps*), respectively. **C)** KEGG pathway attributed genes of *P. veronii*. Circle sizes show the proportion of genes attributed to the corresponding KEGG pathway for the full *P. veronii* genome (PVE, as per total of KEGG-attributed genes) compared to that of the list attributed from genes depleted in Lib1 and Lib2 (as in A, but with q<0.01; n = 1293). Colored boxes indicate whether the proportion in the Lib1&2-depleted series is higher (cyan) or lower than (magenta), or within (white) the 95% confidence interval (CI) calculated on the mean of the *P. veronii* genome KEGG attributions, randomly subsampled 10 times independently for the same sampling size (n = 1293). **D)** Same as (C) but for cluster of orthologous group (COG) attribution in the full genome

and the set of depleted genes in Lib1&2 at start. Asterisks indicate significance outside the calculated confidence interval as above (CI, $p < 0.05$).

Essential genes in P. veronii starting mutant libraries

As both starting libraries had been enriched to deprive them of wild-type *P. veronii* and *E. coli* donor, we expected that the procedure in itself had resulted in selective conditions and loss of certain gene insertions. In order to identify gene insertions that possibly impaired mutant fitness already in the starting libraries, we scored those genes for which the number of insertions was statistically significantly lower than what would be expected from a library with completely random insertions (taking into account gene size and the observed $1.2\text{--}1.7 \times 10^5$ unique insertions per genome in the starting libraries, summed per 10-kb regions). At a q-value of < 0.0005 (p-value < 0.01) and a ratio < 0.01 the initial libraries overlapped in 740 genes with lower than expected insertions (Figure 2A, Supplementary table S1). 440 of those overlapped with the essential gene set ($n = 565$) identified from *Acinetobacter baylyi* ADP1 for growth on minimal medium (Gallagher et al., 2020)(Supplementary table S2). Several obvious gene regions into which transposon insertions were expected to cause fitness loss (e.g., *fts*, *rpl*, *rps*, Figure 2B; *sdh* genes for succinate dehydrogenase, succinate being used for library selection, Supplementary table S2) were indeed strongly depleted for insertion densities. Other genes, e.g., encoding amino acid or nucleotide biosynthetic pathways, were also depleted for transposon insertions compared to a random expected insertion library, but were not completely devoid of any detectable insertions (Figure 2C, e.g., arginine, pyrimidine). Other biosynthetic pathways were not significantly affected (Figure 2C, e.g., tryptophane, alanine). This suggests such mutants survive to a low extent in the ‘meta’-population of transposon mutants (which was grown as a pool) perhaps because of metabolite sharing or of pathway redundancies. At the level of COG-attributed classes, the genes depleted in the starting libraries were significantly enriched in COG-D (Figure 2B, Cell cycle control, cell division and chromosome partitioning, $p < 0.05$ to randomly drawn genes, $n = 10$) and COG-J (Translation). Genes within these classes were thus associated with negative fitness effects, as expected from Figure 2A. However, most other COG classes were underrepresented, suggesting overall positive fitness effects of insertions (e.g., COG-N motility, COG-T signal transduction, Figure 2B). The proportion of genes without any COG attribution was also significantly enriched in the starting libraries compared to what was expected from the full *P. veronii* genome (Figure 2B). This may indicate a high number of poorly characterized genes with negative fitness effects on growth upon inactivation.

Most of the genes without any insertion in any of the libraries (around 200 genes), consisted of very small genes (< 150 bp), which simply by chance or by insertional preference may not have been targeted by the transposon. For these reasons, we consider that criteria such

as ‘zero reads’ in any of the libraries was not useful for scoring gene essentiality. For the remainder, therefore, we used time trends in relative insertion abundances and ratio thresholds as the more important criteria for assessing the fitness effects of gene insertions to treatment conditions.

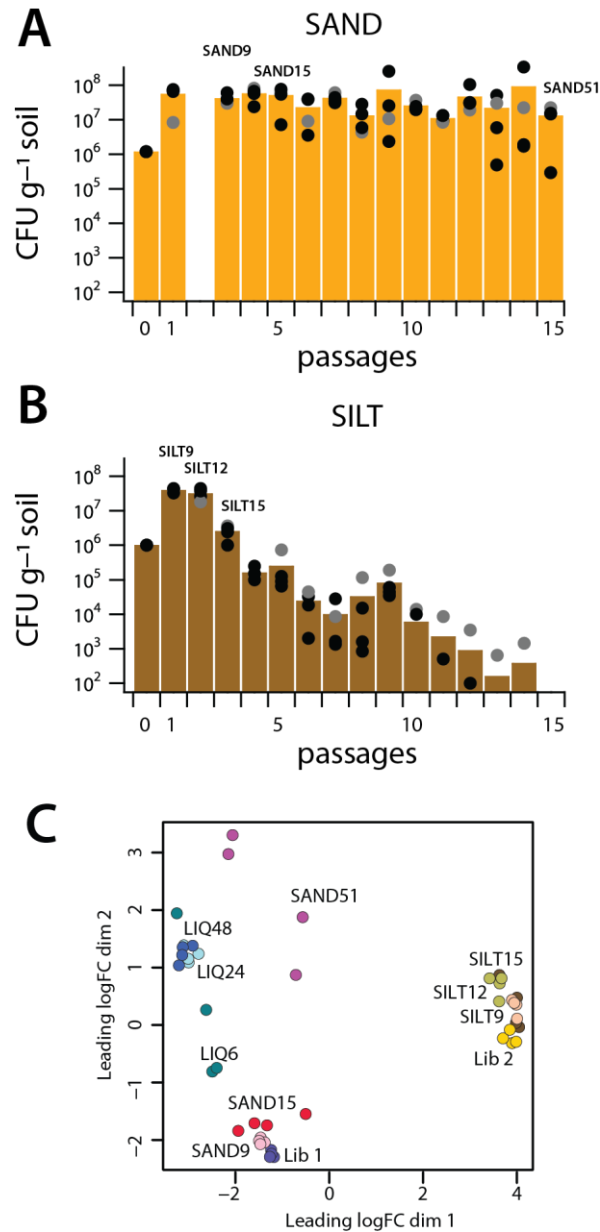


Figure 3. Soil transposon libraries of *P. veronii*. **A)** *P. veronii* meta-population sizes after each passage in Sand microcosms under toluene vapor. Soil growth cycles consisted of a 1:10 w/w dilution into fresh soil. CFU, colony-forming units on plates with toluene vapor. SAND9, etc., estimated number of generations of growth of the *P. veronii* meta-population. **B)** As (A) but for Silt. **C)** Multidimensional clustering of initial library replicates (Lib1 and Lib2, corresponding to the starting libraries), treatments and sampling times (in generations). SAND, passaged libraries in Sand microcosms; SILT, Silt microcosms; LIQ, liquid suspended culture. All further libraries grown in presence of toluene. SAND and LIQ derive from Lib 1; SILT derives from Lib 2.

Gene insertion fitness effects as a function of population dynamics and conditions

In order to identify genes with fitness effects, we followed and contrasted the growth of *P. veronii* mutant populations under three different culturing conditions, and through multiple culturing cycles to produce an estimated ~50 generations of ‘meta’-population growth. Our hypothesis here was that the specific culturing regimes would create selective effects, leading to a gradual change in the relative proportions of individual *P. veronii* mutants in the starting libraries, which would be reflected in changes in relative proportions of per gene insertion read densities in the sequenced libraries. From Lib1, we cultured cells either in liquid suspension (LIQ) or in Sand microcosms (SAND) with added toluene as carbon and energy substrate, and passaged such cultures multiple times. Mutant populations were recovered after an estimated 6, 24 and, 48 (LIQ), or 9, 15, and 51 generations (SAND, Table 1). The meta-population of *P. veronii* mutants maintained itself quite well during repeated passages in Sand (Figure 3A).

In contrast, mutant populations from Lib2 cultured and passaged in Silt microcosms maintained population densities for the first three passages, after which it decreased (Figure 3B). Consequently, we restricted our sampling to the first three cycles with estimated 9, 12 and, 15 generations of growth (SILT, Table 1). The reason for this difference may be the 100-fold higher background soil microbiota in Silt than Sand (Morales et al., 2021) and, therefore, increasing competition or predation from native microbiota. Replicate Tn-seq samples grouped close together in multi-dimensional scaling and further clustered by treatment and time, with samples taken after more generations of mutant library growth tending to drift ‘further’ from the starting libraries (Figure 3C). Replicate variability on average resulted in a 10% variation of the mean insertion density per gene, with notable exceptions at the highest generations (SAND-51) and in liquid suspension (LIQ, 35-100% variation, Supplementary figure S2).

To better discern consistent effects of marker insertions in genes on the overall fitness of the mutants for adaptation, growth, and survival in soil, we took advantage of the three-time sampling points and scored time trends in the relative abundances of insertion densities per gene in comparison to that at the start. First, we calculated the regression coefficients on the mean of the normalized insertion densities per gene ($n = 4$ replicates) across all time points per treatment (Figure 4A). We differentiated cases where the means decreased over time, meaning the relative abundance of insertions in those genes diminished as the libraries passed through more generations of growth (*down*, Figure 4B). Similarly, the *up* category signified increasing relative abundances, whereas *flat* regression lines pointed to no change in relative abundance (Figure 4B). Among the *flat* category, we further differentiated to those genes, which at the first time point had at least two reads (*neutral*) and those that had less (*zero*). Proportionally speaking, genes with *zero* or *down* behaviour comprised the majority of cases, whereas *up* and *neutral* each encompassed some 600–700 genes (Figure 4C). Genes in the categories *up* and

down may point to those where insertions led to fitness gain or loss, respectively, as a function of the treatment. Those in the category *neutral* may point to genes with insertions that do not further influence fitness under the used growth conditions. Given the overall gene insertion variability per replicate (Supplementary figure S2) we concluded there would not be sufficient statistical power to infer any further meaning to the class of *flat-zero*. This category comprises many genes that on the basis of a comparison of the expected versus observed insertion densities (Figure 2, Supplementary table S1) had already been associated with fitness effects in the starting libraries.

Table 2. Underrepresented functions in *P. veronii* mutant libraries shared among soil and suspended liquid culturing or only in both soil conditions.

Subsystem	Function	Depleted all	Locus_tag Depleted soils
Extracellular or cell wall polymers	Alginate biosynthesis	r1g0181, r1g5800	r1g1054
	Glycosyltransferase	r1g1060, r1g0485-0488, r1g0526	
	Polysaccharide biosynthesis	r1g0400-0402, r1g1724-1737	
	Biofilm formation	r1g1304-1307	
	Biopolymer transport	r1g2401, r1g5927-5930	
	Pilus assembly and biogenesis	r2g0191-0195, r2g0893	
	Membrane protein modification		r1g0191-0192, r1g0389, r1g0611
Global regulation	Hfq	r1g0563-0568	
	<i>recA/recX</i>	r1g1274-1275	
	<i>lexA</i>	r1g1643	
	Sigma factors	r1g3422, r1g3602, r1g3573, r1g4475	r1g1384
	Polyphosphate kinase	r1g5787	
	Chaperones		r1g4921, r1g5118
	DNA methyltransferase		r1g4069
Central metabolism	Tryptophane synthesis	r1g0361	r1g5451-5454
	Short chain fatty acid metabolism		r1g4493, r1g4527, r1g4583
	Entner Doudoroff metabolism	r1g0144-0146, r1g4705, r1g4817, r1g4825	
Nutrient transport	Phosphate	r1g0616, r1g5974-5975, r1g5969	
	Iron	r1g0833, r1g1231, r1g5743-5744, p0131-0135	r1g5754
	Magnesium	r1g3472, r1g3844, r1g4627, r1g4808	
	Ammonium	r1g1817	
	Urea	r1g3555	r1g0606, r1g1504-1506, r1g2167
	Copper	r1g6093, 6098	
	Porins	r1g4474, r1g537	
	Amino acids	r1g0629	r1g0387, r1g0596
	Sulfate		r1g0274-0275, r1g0601, r1g5694
Transporter system	Phospholipids reflux	r1g0964-0971	
	Multidrug efflux	r2g0379-0381	
	Potassium and sodium	r1g498, r1g4078	r1g1337
Solute, osmoprotection	Spermidine and putrescine	r1g1793, r1g3317, r1g5459-5460, r1g5993	
	Glycine-betaine transport		r1g5572
Respiration	NADH-quinone oxidoreductase	r1g2343-2355	r1g2343-2355
	Alternative cytochromes		r1g150-153, r1g4815
Oxidative stress	Peroxioredoxin	r1g2391, r1g4247, r1g5448	
	SoxR	r1g4016	

Rhizopine	Rhizopine binding and catabolism	r1g2694-2700	
Resistance	Mercury	r2g0813-819	
	Arsenic		r1g1869
Hypotheticals		(118 members)	(40 members)
Transcription regulators		(32 members)	(7 members)

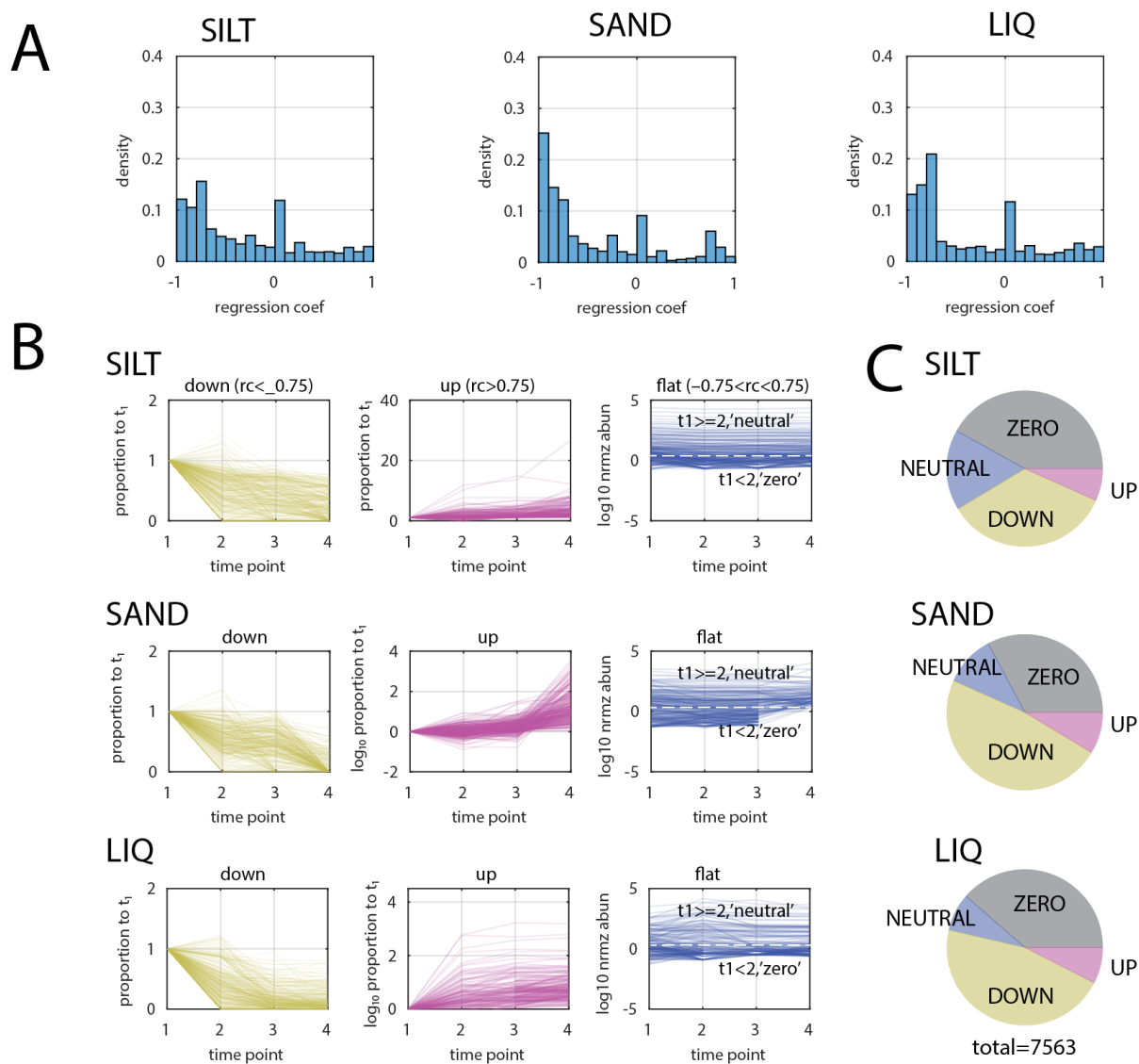


Figure 4. Global insertion mutant behaviour over time. A) Regression coefficient distributions for each of the treatment series: Silt, Sand and Liquid. B) Relative abundances of individual gene insertion numbers over sampling time points, categorized as consistent downwards (i.e., negative correlation coefficient <-0.75), consistent upwards (i.e., positive correlation coefficient >0.75), or flat (correlation coefficient in between -0.75 and 0.75 , further differentiated to those genes which at the first sample point had at least 2 or more counts, and those that had less). Note that time points refer to sampling points as in table 1. Further note that *down* and *up* are plotted as proportion compared to time point 1; *flat* is plotted as absolute read abundance. C) Proportional mean attribution of per gene insertions to each of the four categories defined in (B, 'flat'-neutral, and 'flat'-zero), for each of the treatments.

Positive fitness effects during soil incubation

In order to tease out genes with insertions causing a fitness gain under soil conditions, we compared the *up* categories defined above among the Silt- and Sand-incubations with that of liquid suspended cells grown with toluene. In addition to the slope criteria defined above, we

further included the criterium that a gene at the last time point should have at least two reads and should have more than five times the number of reads at start., There were 27 genes in this category among both soil incubations, 23 of which did not occur in liquid (Supplementary table S3). Surprisingly, the large majority of those fell in genes related to flagellar biosynthesis and motility, in two different genome locations (Figure 5A; Table 2, *fla*, *flh*, *fli*, *flg* genes). This was the strongest signal we observed and suggests a strong fitness gain from the inactivation of flagella in unsaturated soils for *P. veronii*. Insertions in a few other genes also caused strong fitness gains, such as coding for an AAA-ATPase, a chaperonin, a permease and, a universal stress protein (Supplementary table S3).

Of the same starting library (Lib1), there were 23 *up* cases passing the filtering thresholds in common between liquid and Sand incubations (Supplementary table S4). These did not carry a single common signature, but four of those caused massive fitness gain for cells in liquid. These comprised an insertion in a hypothetical gene PVE_r1g5203, two regulatory genes on replicon 2 (PVE_r2g0382 and PVE_r20399), coding for the toluene efflux pump HTH-type transcriptional regulator TtgV and an RNA polymerase sigma factor RpoH, respectively (Supplementary table S4). Finally, insertion in a plasmid-located gene PVE_p0076 coding for an 8-amino-7-oxononanoate synthase also resulted in a significant fitness gain (Supplementary table S4).

Negative fitness effect during soil incubation

Since all experimental conditions included the presence of toluene as growth substrate, we expected to see fitness differences for insertions in genes for the known aromatic metabolic pathways of *P. veronii*. In particular, insertions in the *ipbAa-Ad* genes encoding the first step of toluene degradation were disproportionately low, even in the starting libraries, but their densities diminished further to complete absence in all conditions (Supplementary table S5). Insertions in the *dmp* and *ibpC-D* genes for the *meta*-cleavage branch were less drastically affected than *ipb*, possibly because they are functionally redundant; but they also decreased at later time points (Supplementary table S5). In contrast, insertions in the *nah*-gene encoded *meta*-cleavage pathway were without measurable fitness cost, suggesting they are either not expressed or important for the metabolism of other aromatic substrates for which our experiments did not select (Supplementary table S5).

In order to see whether genes attributed to *down*, *up*, or *flat* classes affected particular subsystems, we grouped them into COG-classes (Figure 5B). COG-attributed genes of the *flat* class among all conditions but with reads at start < 2 (i.e., *flat-zero*) were similar to that of the subset of genes depleted at the start in both libraries (Figure 5B, ALL(ZERO) and intersect ZERO&T0-DEplete). This suggested that insertions in those genes affected fitness in the

starting libraries and did not re-appear in any of the subsequent incubations, or were gene regions with very poor target insertion selectivity (e.g., Figure 1B). Therefore, we did not consider this category further. In contrast, among all three culturing conditions there was a core set of 67 genes (at the applied thresholds), without any measurable fitness defect upon insertion (Figure 5B, ALL NEUTRAL). These cover a variety of COG categories but without any particular coherent genomic context or pathway structure (Supplementary table S6).

We focused then on the *down* categories, indicative of insertions causing fitness loss under the soil or liquid toluene culture conditions. Consistently observed among the *down* gene subsets in the soil incubations (Sand and Silt, either individual or in overlap) were enrichments in the COG-categories H, P, and T (Figure 5B, $p < 0.05$ on the calculated confidence interval of randomly subsampled gene subsets). This suggests an overproportional assignment of insertions causing fitness loss during soil growth and survival to genes implicated in *Coenzyme transport and metabolism*, *Inorganic ion transport*, and *Signal transduction mechanisms*. Table 2 lists a subset with the most obvious functions affected by the transposed marker insertions under toluene cultivation in both soil conditions, or shared with culturing in suspended liquid medium (full set reported in Supplementary table S7).

In contrast to our expectations, very few functions were exclusive for causing fitness defects in both soils but not in liquid medium, whereas most were shared among all growth conditions (Table 2). A smaller proportion of gene functions associated with decreased fitness appeared exclusively among cells growing in Sand, perhaps because of the higher number of growth cycles achieved in Sand than in Silt (Figure 3A, Supplementary table S7). Most functional defects that appeared during prolonged culturing related to extracellular or cell wall modifications, global regulators, nutrient transport and, oxidative stress (Table 2). About one-third of these have detectable overlap to the gene set previously defined for fitness importance in Sand of *Sphingomonas wittichii* RW1 (Moreno-Forero and van der Meer, 2015) (193 out of 748, at BLASTp e-value of $< 1e-10$ and query overlap length of > 0.75 ; Supplementary table S8). In addition, a large number of uncharacterized transcription factors and genes coding for hypothetical proteins diminished *P. veronii* fitness upon marker insertion (Table 2). Of particular mention were genes implicated in the Entner-Doudoroff pathway, that notably for pseudomonads provide the capacity to regenerate NADPH (Nikel et al., 2013, 2015). Unexpectedly, insertions in the main NADH-quinone oxidoreductase system (a single operon on PVE_r1g2343-2355) were not counterselected in cells at start, but decreased fitness over time in cells cultured on toluene, and with stronger effects in soil incubations than in liquid (Table 2, Supplementary table S7). A rather weak fitness effect was detected for subsystems predicted to be implicated in solute and osmoprotection, perhaps because *P. veronii* encodes several redundant systems for this. One of the few signatures that we picked up specific for cells

cultured in soil were insertions in genes for short-chain fatty acid metabolism (Table 2), and genes for urea metabolism and sulfate uptake/metabolism (Table 2), which possibly give an advantage for growing in soil. There were a few genes with specific fitness defects either in Sand or in Silt (Supplementary table S7), but these are difficult to attribute to coherent pathways or subsystems. This suggests that at least for *P. veronii* there is not a large common set of genes important for survival in 'soil', although the overlap between Silt and Sand may have been more difficult to interpret as a result of the shorter overall number of growth cycles (and thus generations) in Silt than in Sand.

Finally, we tested whether genes associated with fitness decrease in Sand or Silt were more likely to be higher expressed among all *P. veronii* genes, which would attest to their importance for growth under these conditions. For this, we took advantage of the measured transcriptome of *P. veronii* during exponential growth in Silt (Morales et al., 2021), and scored whether genes in the *down* categories were more likely to have expression above the median \log_2 CPM expression value of all differentially expressed genes between Silt and liquid (Figure 5C). In comparison to repeated random samplings ($n = 10$) of the same-sized gene sets, indeed, there was a statistically significant tendency that *down* genes are among the higher expressed genes (Figure 5C, $p < 0.05$ confidence interval). For comparison, also the flagellar genes highlighted in Figure 5A, are among the higher expressed genes of *P. veronii* in Silt, although in this particular case marker insertion resulted in fitness gain in soil.

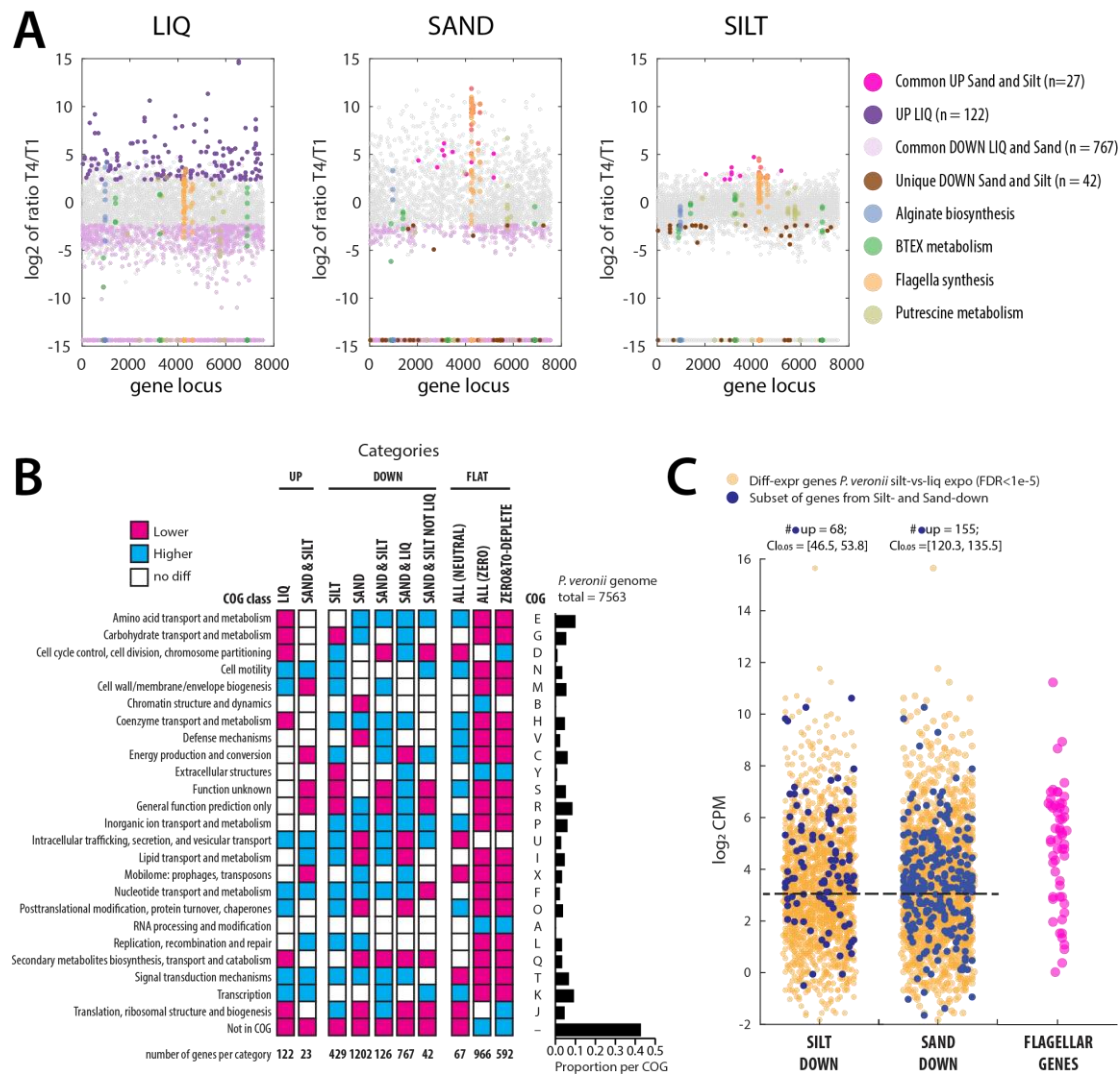


Figure 5. Common and condition-specific fitness-affected genes of *P. veronii* during growth with toluene. **A)** Diagrams plot the log₂-ratio of the mean normalized per gene insertion density (organized according to gene locus number) at the last sampling point (T4) and the first (T1, light grey circles) for the three growth conditions; LIQ, liquid suspension; SAND and SILT, Sand and Silt microcosms. Magenta, purple, light magenta and brown; fitness-affected genes (UP and DOWN, and their numbers in the sets) as defined from time series analysis as in Figure 4. Cyan, green, ochre and light green; *P. veronii* genes involved in the indicated metabolic or biosynthetic pathways. Note the high enrichment of genes for flagellar synthesis and motility with fitness benefit under Sand or Silt growth. **B)** COG attribution of enriched or depleted gene insertion read density sets. Colored boxes show observed higher (cyan) or lower than (magenta) expected COG-class attributions ($p < 0.05$ confidence interval of the mean of 10 randomly subsampled distributions with the same size as the observed number of genes per category). White boxes indicate within the confidence interval. Categories contain gene sets as defined by time regressions and threshold filtering. Barplot in black shows COG-attributions predicted for the full *P. veronii* genome. **C)** *P. veronii* gene expression (as log₂ normalized counts per million, CPM) in Silt in exponential phase, overlaid with that of the subsets of genes depleted during growth in Silt (SILT-DOWN) or Sand (SAND-DOWN), and in comparison of that of the flagellar genes (same subset as in A). Numbers of

genes in the subsets $> \log\text{CPM}=3$ are listed on top, in comparison to the confidence interval at $p=0.05$ ($\text{CI}_{0.05}$) calculated from randomly sampled gene subsets of the same size ($n=10$).

Discussion

Bioaugmentation relies on the capacity of exogenous strains with xenometabolic potential to adapt, grow and survive after inoculation in a polluted environment. However, it is still rather unclear which inherent characteristics of strains contribute to the intended growth and survival. The major aim of the underlying study was thus to identify genes in the candidate strain for monoaromatic compound bioaugmentation *P. veronii* 1YdBTEX2 that are crucial for fitness maintenance in soil under pollution stress. We produced two independent high density genome-wide transposed random insertion libraries and contrasted the relative abundances of gene insertions over multiple generations in either cycles of liquid suspended or soil growth artificially contaminated with toluene. In contrast to previous data with the dibenzofuran-degrading bacterium *S. wittichii* RW1 (Roggo et al., 2013), relatively few gene insertions caused *P. veronii* fitness defects exclusively in contaminated soils. These included urea metabolism, sulfate uptake/metabolism, short-chain fatty acid metabolism, and genes involved in oxidative stress defense. This sparsity is in agreement with recent transcriptomic data, which suggested that *P. veronii* maintains a remarkably constant global core expression profile independent of growth in liquid or soil (Morales et al., 2021). Both transcriptomics and transposed insertion library sequencing thus seem to indicate that *P. veronii* adapts very easily and similarly, irrespective of the environment it is in, provided the absence of too severe obvious water or toxic stresses (Moreno and van der Meer, 2016). The reason for this different physiological 'wiring' than, for example, *S. wittichii* (Roggo et al., 2013), is not clear, but may be dependent on differences in global regulatory mechanisms and gene redundancies.

Tn-seq is extremely powerful to identify fitness-essential genes and has been widely applied to understand gene essentiality under different growth conditions in pathogenic bacteria (Hardy et al., 2021; Harrison et al., 2020; Jacobs et al., 2003; Liberati et al., 2006; Murray et al., 2009; Qin et al., 2004; Subashchandrabose et al., 2013) but less frequently for environmental bacteria with bioaugmentation potential (Coronado et al., 2014; Mirleau et al., 2005; Roggo et al., 2013). The technique also poses a number of challenges and pitfalls that complicate the quantitative analysis of inserted marker frequencies in analyzed libraries. Most frequently, generated mutant libraries are incubated *en masse* under the test conditions with the assumption that if a marker insertion inactivates a gene with a fitness effect, the cell bearing that mutation will divide at a different rate in the meta-population of mutants. Consequently, the relative proportion of that gene insertion will change over time within the library as a whole. However, depending on the host strain and the characteristics of the (hyperactive)

transposon, marker insertion positions may not be unbiased, as we showed here as well. Furthermore, marker insertions can have polar effects on downstream-located genes, the difficulty of which can be somewhat alleviated by co-interpreting effects of insertions in transcriptional units rather than single genes only.

On the other hand, depending on the position within the gene, not all insertions may produce the same fitness effect, although for quantitative comparison, insertion densities are grouped per gene (to have sufficient statistical power). Finally, complete absence of any selection during the production of the transposed insertion libraries is almost impossible to achieve, and, therefore, even the first sequenced overview of insertion positions and abundances is already biased for the absence of highly essential genes (e.g., see Gallagher et al. 2020). In case of *P. veronii*, we had to rely on conjugation to deliver the transposon from the *E. coli* donors. Although our procedure was designed to diminish as much as possible further cell division before the first screening, some stages of cultivation were essential to establish the inserted marker in the meta-population and reduce the wild-type donor and recipient cell abundances (by more than 4–5 orders of magnitude). Essential genes that might have been lost already at the stage of the first libraries were thus assessed by comparing insertion position frequencies to that of a completely *in silico* randomized insertion library, under further consideration of observed per gene insertion distributions. This essentiality assessment corresponded reasonably well with a recent list of true essential genes identified in *A. baylyi* ADP1 by transposon transformation without selection bias (Gallagher et al., 2020), and with obvious cell division and translation markers. Furthermore, many other genes involved in basic cellular functioning such as respiration and energy generation (ATP synthesis, cytochrome bb3 oxidase and NADH dehydrogenase), peptidoglycan biosynthesis and DNA replication, as well as growth on succinate (used for library enrichment), were underrepresented in the *P. veronii* starting libraries. Genes for several biosynthetic pathways, however, were not depleted as much as expected upon marker insertion. This may be the effect of co-culturing meta-populations of mutants that to some extent share essential metabolites or amino acids.

More important for our study than an ‘absolute’ gene essentiality analysis was the potential identification and interpretation of genes affecting conditional *P. veronii* fitness in soils under pollutant stress. For this, we applied a conservative data analysis method that took time trends into consideration, rather than a single incubation time point comparison to the starting libraries. Furthermore, we produced two independent libraries that were screened in two different soils artificially contaminated with toluene to increase detection sensitivity. Growth cycling in one of the soils (i.e., Silt) could not be completed as intended, because of a dramatically reduced overall *P. veronii* population size after the third passage (estimated 15 generations). This suggested much poorer survival of the *P. veronii* mutant meta-population in

Silt than Sand with gaseous toluene, which may be due to competitive loss against the higher number of resident microbes in Silt (Morales et al., 2021), or to increased effects of grazing by protozoa on the altered mutant meta-population (see below). Grazing has frequently been proposed as a reason for poor inoculant survival (Tyagi et al., 2011).

Despite these differences in the *P. veronii* mutant meta-population behaviour in the two soils, time regression analysis identified a number of genes with shared fitness effects in the two independent libraries. In terms of aromatic compound metabolism, these consisted of (expected) fitness defects from insertions in the *ipb*-genes coding for the first steps of toluene transformation, but less so in genes for the *meta*-cleavage pathways of (methyl)catechols (i.e., *ipbC-H* and *dmp*). Both *ipb* and *dmp meta*-cleavage gene clusters are induced in the presence of toluene (Morales et al., 2016), but are seemingly redundant from their observed minor fitness loss. In contrast, insertions into the *nah* genes for the third *meta*-cleavage pathway in *P. veronii* were without fitness cost under conditions with toluene in either soils or liquid. As the *nah* genes are not specifically induced by toluene (Morales et al., 2016), we conclude that they might have a role in the metabolism of other aromatic compounds. Mutations in genes coding for a toluene efflux pump (*ttgGHI*) also had a fitness cost on *P. veronii*, similar as shown for other toluene-degrading *Pseudomonas* (Molina-Santiago et al., 2014; Rodríguez-Herva et al., 2007). Interestingly, insertion mutants in the *ttgV* transcriptional regulator of the *ttgGHI* operon were actually positively selected during liquid suspended growth with toluene. This is similar as observed for *Pseudomonas putida* DOT-T1E mutants with a *ttgV* deletion that caused a fourfold higher expression of the *ttgGHI* operon, making it tolerant to higher toluene concentrations (Fillet et al., 2009; Rojas et al., 2003). *P. veronii ttgV* insertion mutants may thus have become selected because of the higher toluene availability in liquid (and toxicity), in contrast to soils (Huertas et al., 1998).

The list of genes where marker insertion decreased fitness exclusively in soils was narrowed by using a combination of pathway analysis (KEGG and COG), individual gene annotation, and potential transcriptional units. This list did not include major specific metabolic pathways, except parts of metabolism of urea, uptake and metabolism of sulfate and other trace elements or growth factors, steps in short-chain fatty acid metabolism, and oxidative stress response. Although some of these overlapped with genes causing fitness defects of *S. wittichii* RW1 in dibenzofuran-polluted Sand (Roggo et al., 2013), the program as a whole was much less distinguishable. This suggests, therefore, a very different strategy of *P. veronii* than *S. wittichii* that consists of maintaining central carbon metabolism no matter the environment. This conclusion is further supported by the distinct correlation observed here between gene fitness effects and their expression in soils (Morales et al., 2021). One possible biochemical mechanism underlying this maintenance is the Entner-Doudoroff (ED) pathway and the main NADH-

quinone oxidoreductase system of pseudomonads. The ED pathway and the activity of the NADH-quinone oxidoreductase are critical for replenishing the NAD(P)H pools upon oxidative stress (exerted either by endogenous or external perturbations)(Chavarría et al., 2013; Nikel et al., 2021). NADPH serves as the reductant for several reactive-oxygen-species-detoxifying enzymes and mechanisms, such as glutathione regeneration (Cabiscol et al., 2000; Chavarría et al., 2013; Nikel et al., 2021; Pérez-Pantoja et al., 2012). Insertions in *P. veronii* genes for both ED and NADH pathways had a stronger negative effect for growth in soils than in liquid, in agreement with the concept that an abundant and constant NADPH supply is a key feature to ensure survival and maintenance in hostile environments such as polluted soils.

Finally, as observed before for Tn-seq generated mutants of *S. wittichii* in soil (Roggo et al., 2013), some insertions caused massive *P. veronii* fitness gain. This was particularly evident from the fitness increase in soil of insertions in almost all of the genes for flagellar biosynthesis and assembly (two genome locations independently). Previous transcriptomic data had indicated that flagellar genes are higher expressed during growth in soils (Morales et al., 2021), suggesting motility as an important adaptive feature for soil establishment. Indeed, other studies have reported that flagella provide a fitness advantage to cells in soils through improved dispersal and access to nutrient resources (Pion, 2013). This is supported by studies on *S. wittichii* RW1 showing that interruption in flagellar genes resulted in strong fitness loss in soil (Roggo et al., 2013; Wang and Or, 2010), and on *Pseudomonas fluorescens* demonstrating significant better survival of motile than non-motile strains 21 days after inoculation in non-sterile sandy loam (Turnbull et al., 2001). In contrast, flagella biosynthesis and motility itself are costly, and might be counterselected in soils with constant humidity and plenty of carbon availability (as in our experimental designs). Non-motile *P. veronii* mutants, therefore, may have favorably redirected energy to other critical functions such as toluene metabolism or cell integrity maintenance, providing faster growth. Possibly, however, the poor long-term survival of the *P. veronii* mutant meta-population that became largely dominated by non-motile strains became at risk to easier predation in Silt with its higher native microbial background, and went almost extinct. Also, previous growth studies had suggested high sensitivity of *P. veronii* to predation under non-motile conditions in soils such as clay (Morales et al., 2021). Flagella biosynthesis and motility can thus be a two-edged sword for cells: benefitting from chemotaxis for nutrient availability (Ni et al., 2020), particularly during short and infrequent wetting events in unsaturated soils (Dechesne, 2009), and escape from predators, but energy costly. More systematic studies to evaluate the actual impact of motility on predation during strain inoculation into soil might be interesting to be carried out (Kota et al., 1999).

In conclusion, the high-throughput scanning of insertion mutant behaviour under environmentally relevant conditions provided us with a global picture of the strategies and

capabilities of *P. veronii* to grow and survive in contaminated soils. Their fast adaptability and large redundant genomes enable pseudomonads to maintain their main central metabolic routes and grow fast in spite of changing environmental conditions, which seems different from the strategies of other bacteria (e.g., *Sphingomonas*) to mount a specific 'soil-type' response and possibly, grow more slowly. In addition to selecting strains for bioaugmentation based on their catabolic properties, future soil microbiome engineering efforts should thus take other ecological properties into account as well.

Material and Methods

Culture conditions

P. veronii 1YdBTEX2 and *E. coli* strains were routinely grown in lysogeny broth (LB, 10 g l⁻¹ tryptone, 5 g l⁻¹ yeast extract, 10 g l⁻¹ sodium chloride) on a rotatory shaker at 180 rpm at 30°C or 37°C, respectively. As defined minimal medium (MM) we used type 21C mineral medium (Gerhardt et al., 1981). Whenever necessary, medium was solidified with 1.5% agar. As carbon sources for *P. veronii* we used 20 mM sodium succinate or toluene. Kanamycin was supplemented at 50 µg ml⁻¹ (Km50) to select for transposon insertions in *P. veronii* or for maintenance of plasmid pRL27 (see below). Toluene as sole carbon source was supplied to cells growing on MM-agar plates through the vapor phase in a gas-tight reservoir with an open 10 ml glass vial containing 1 ml of liquid toluene.

To obtain bacteriophage lysate for the transposon mutant library enrichment, one ml of Sextaphag® polyvalent pyobacteriophage preparation (Microgen, Russia) was added to 50 ml of *E. coli* BW20767 culture grown in LB with Km50 at a culture turbidity at 600 nm of 0.8, and incubated for 3 h at 37°C and 180 rpm. After that, the lysate was amended with 100 µl of 20 mg ml⁻¹ DNase I and incubated for 15 min as above to reduce viscosity. Finally the lysate (>10⁹ phage per ml) was cleared by centrifugation for 10 min at 17400 × *g* in a FA-45-6-30 rotor (Eppendorf), filtered through 0.45-µm cellulose acetate and stored at 4°C until needed for enrichment (see below).

Transposon mutant library construction

Transposon mutant libraries of *P. veronii* were prepared through conjugation of a suicide plasmid from the *E. coli* BW20767 (pRL27) donor, which expresses a hyperactive transposase leading to insertion of a Km-resistance gene flanked by transposon ends. *P. veronii* and *E. coli* cultures were grown overnight in 80 ml of LB and LB-Km50, respectively; harvested by centrifugation at 3,220 × *g* for 7 min in an Eppendorf swing-out A-4-62 rotor at 25 °C and resuspended each in 40 ml of fresh LB for 1 h at 30 °C without shaking to resume growth and remove traces of Km. Two culture aliquots, each with 1.4×10¹⁰ donor and 1.8×10¹⁰ recipient

cells were then combined in a 15 ml centrifuge tube and centrifuged as described above. After removing the supernatant, the cell slurries were gently aspirated with a pipette and each transferred to a 0.2- μm pore size 25 mm \varnothing cellulose acetate filter disc (Sartorius). Disks were placed on LB-agar and incubated for 48 h at 30°C. The same volumes of donor and recipient cells were separately treated in the same way to serve as controls for the efficiency of the selection process. Following the incubation, cells from each filter were resuspended in 2 ml phosphate-buffered saline (PBS) by vortexing, after which the two duplicate mating mixtures were pooled. Aliquots of 0.1 ml from all cell suspensions were serially diluted in MM and plated on selective media as described below. The remaining mating mixture volume (1.9 ml) was supplemented with glycerol to a final concentration of 15% (v/v), aliquoted, flash frozen in liquid nitrogen and stored at -80°C until library enrichments.

E. coli BW20767 donor cells were counted on LB-Km50 agar plates incubated at 37 °C, which is a non-permissive temperature for *P. veronii*. Wild type *P. veronii* (recipient) was counted on MM agar plates incubated at 30°C in toluene vapour. A Hundred colonies were replica plated on LB-Km50 agar to verify any spontaneous appearance of mutants with Km resistance. Potential *P. veronii* transposon insertion mutants were counted on MM-Km50 agar plates incubated with toluene vapor as sole carbon source. This procedure excludes *P. veronii* toluene-negative mutants, thus the colony counts are used as an approximation.

P. veronii Tn5-mutant library enrichment

In order to reduce the incidence of wild type *P. veronii* and *E. coli* cells in the raw transposon libraries, we thawed an aliquot of 0.6 ml of the frozen mating mix (with an estimated 3.6×10^5 Km-resistant *P. veronii* mutants) on ice, and inoculated this into 50 ml of MM with 20 mM succinate and 50 $\mu\text{g ml}^{-1}$ Km in a 250-ml-sized conical flask. The resulting liquid culture was amended with 2 ml of the *E. coli* BW20767-specific bacteriophage preparation and incubated for 48 h at 30°C and 180 rpm. Following the enrichment, we counted the number of Km-resistant and Km-sensitive *P. veronii*, and the remaining *E. coli*, as described above. A few dozen individual colonies of Km-resistant *P. veronii* were analyzed by PCR-amplification of a fragment of the aminoglycoside phosphotransferase (*aph*) gene and a *P. veronii* genome marker, respectively, using primer pairs *aph_rev/aph_fw* and *Pv_chr2_fw/Pv_chr2_rev* (Supplementary table S9). This whole procedure (starting from the matings) was repeated once more and two enriched libraries, named Lib1 and Lib2, were aliquoted and stored at -80°C with 15% glycerol. At the point of freezing, the stocks had approximately 3.6×10^8 Km-resistant *P. veronii* cells (estimated from CFU on MM-toluene-Km plates), and less than 1.9×10^4 *E. coli* (estimated from CFU on LB-Km at 37°C).

Selection conditions in liquid medium

To potentially identify *P. veronii* mutants with insertions in genes causing fitness loss, we imposed the following selective conditions: (i) growth in defined MM with toluene as sole carbon source, (ii) growth in Sand, or (iii) Silt with toluene as sole added carbon source. For conditions (i) and (ii) we used the Lib1 library; for condition (iii) we used Lib2.

To prepare the libraries for inoculation in the selective conditions, 3.6 ml of the library glycerol stock was revived in 30 ml of LB with Km50 in a 100 ml conical flask, incubated at 30 °C and 180 rpm for 3 h until a culture turbidity of 0.5 (ca 5×10^8 cells ml⁻¹). The 10-ml culture was centrifuged at $3,220 \times g$ for 8 min in an Eppendorf swing-out A-4-62 rotor at 25 °C, after which the cell pellet was resuspended in 12 ml of MM without added carbon substrate. This cell suspension was further diluted 100-fold in the same medium (50 ml) to produce the inoculum for the selection experiments (ca 5×10^6 cells ml⁻¹). Four aliquots of 2 ml (of OD~0.5) were kept separately; cells were harvested by centrifugation as before and stored at -80°C until DNA extraction (t=0 samples of Lib1 and Lib2).

For growth in liquid culture with toluene, we transferred 5 ml of the inoculum directly to 50 ml-sized polypropylene tubes (Greiner AG, cat #227261) in four replicates. A 1-ml pipette tip sealed on one end containing 0.2 ml pure toluene was placed into the tubes, after which they were closed with a screw cap and incubated upright at 25 °C for 72 h. After incubation, the toluene containing tip was removed from the tube and the cell suspension was vigorously vortexed, after which a 0.1 ml subsample was taken, serially diluted and plated on MM with Km50 incubated in the presence of toluene to count the number of Km-resistant *P. veronii*. A further subsample of 0.1 ml was diluted in 5 ml MM into a new 50-ml Greiner tube, and again incubated with toluene for 3 d. This passaging was repeated 8 times in total, each corresponding approximately 6 generations of growth. The remaining culture volumes (4.8 ml) after cycles 1, 4 and 8 were centrifuged as above to pellet the cells, which were frozen at -80 °C until DNA isolation.

Selection conditions in soil microcosms

For selection in Sand and Silt with added toluene, we prepared microcosms in 250 ml Schott flasks each containing 50 g of non-sterile air-dried (7–10 days at room temperature) of Sandy soil or Silt (for sampling locations, see (Morales et al., 2021)). Four replicate microcosms were started simultaneously, which were inoculated with 5 ml of the reconstituted and diluted library suspension of above (to start at 2.5×10^6 cells g⁻¹), and mixed manually with a sterilized spatula and then placed on a roller board for 30 min. Flasks were then placed upright and a 1-ml pipette tip with one end sealed was placed inside with 0.4 ml pure toluene. The soil microcosms were closed with a teflon-lined screw cap and incubated without shaking at 25 °C for 3 days.

After incubation, the tip was removed from the flask and the content of the microcosm was homogenated by rotation on a roller board (IKA roller 6 digital) during 30 min. Two aliquots of ~5 g were then retrieved, one of which was used to count Km-resistant *P. veronii* (see below), the other to inoculate a next microcosm with 50 g material. The new microcosm was amended with 4.5 ml MM to maintain the same gravimetric water content. This passaging was repeated 15 times, each cycle was considered to allow ~3 generations of growth in soil. The remaining material was then used for DNA isolation (see below).

Selective growth with toluene in Silt (condition-iii) was followed using the Lib2 library. In this case, we used microcosms consisting of 90 g of air-dried Silt inside 500 ml glass Schott bottles with Teflon-line septa and screw cap, into which a one-end sealed 5 ml pipette tip with 0.2 ml of pure toluene was placed. Four replicate microcosms were prepared, each inoculated with 10 ml of the washed revived Lib2 suspension (achieving the same starting cell density in the soil). Passaging consisted of sampling 11 g of soil using a sterile spatula, which was transferred to a fresh 90 g microcosm, further amended with 9 ml of MM to maintain the GWC.

Counting P. veronii in soil microcosms

Aliquots of 5 or 11 g (for Sand and Silt microcosms, respectively) were transferred into 50-ml Greiner tubes. 9 ml of MM was added to the soil and the tube was vortexed at maximum setting for 1 min. Larger soil particles were allowed to sediment for a few seconds, after which the supernatant was transferred to a fresh 50 ml tube. An aliquot of 50 μ l was removed and serially diluted in MM, which was plated on MM-Km50 and incubated with toluene vapor to count the number of colony-forming units of Km-resistant toluene-degrading *P. veronii*. The remaining volume was centrifuged for 8 min at 4000 $\times g$ to pellet the cells, which were frozen with liquid nitrogen and stored at -80°C for DNA extraction and library preparation (see below).

Soil DNA extraction and P. veronii DNA quantification

Total DNA was extracted from frozen cell pellets using the DNeasy PowerSoil Kit (Qiagen, Cat No: 12888-100) according to manufacturer's recommendations. DNA quantity and quality in the purified solutions were verified by reading the 260 nm absorbance and the absorbance ratios at 260/280 nm and 260/230 nm, by using a NanoDrop spectrophotometer (ThermoFisher Scientific). In addition, DNA was examined for purity and integrity on 1% agarose-Tris-acetate-EDTA gel electrophoresis.

P. veronii DNA was quantified by qPCR to normalize the input of DNA from *P. veronii* across library samples, targeting the PVE_r1g_3276 gene using forward (5' tcaccggtactactgaaacgc) and reverse primers (5' attccgcaattgaccaggc, 165-bp amplicon). Each qPCR reaction was performed in triplicate in a total volume of 10 μ l using SYBR® Select Master 846 Mix (Applied Biosystems)

with 1 µl of isolated DNA solution. Each plate contained a positive control with *P. veronii* wild-type DNA and a water control. All qPCR reactions were carried out in a 96-well plate on a StepOnePlus instrument (Applied Biosystems) with the thermal cycling conditions as follows: denaturation stage at 50°C for 2 min followed by 95°C for 2 min, 40 amplification cycles at 95°C for 15 s, and 60°C for 1 min.

DNA library preparation and Illumina sequencing

Aliquots of 0.5 µg DNA (normalized by qPCR results above) was used for library preparation in a customized protocol. First, the DNA sample was enzymatically fragmented (300bp – 700bp) using the NEBNext Ultra II FS Enzyme Mix according to the supplier's instructions (New England Biolabs). Fragmented DNAs were made blunt-ended and phosphorylated, after which an A-nucleotide was added to the 3' ends. Subsequently, NEBNext loop adaptors with a single T-overhang were ligated (all according to instructions of the suppliers). Excess adapters and small (< 100 bp) fragments were removed using SPRI magnetic bead purification (Agencourt Solid Phase Reversible Immobilization magnetic beads, Beckman-Coulter). Purified adapter-ligated fragments containing *aph*-transposed insertions were subsequently enriched by PCR amplification for 21 cycles, using a customized primer complementary to the left-end transposon end and the indexed P5 Primer/P7 primer for Illumina (NEBNext Multiplex Oligos for Illumina; index primers Set 1 and Set 2 – New England Biolabs. Supplementary table S9). Two subsequent purification steps with SPRI-beads were followed to remove excess of primers, PCR reagents and to select DNA amplicons between 300 – 700 bp (i.e., by using 0.5- and 0.2-fold SPRI magnetic bead solution to total reaction volume).

In a final nested PCR with 6 cycles, fragments that contained the Tn5 transposon were linked to the NEBNext adaptor sequences for Illumina sequencing (Supplementary table S9), after which the DNAs were purified twice using 0.8-fold volumes of SPRI beads. DNA concentrations in the prepared libraries were measured using the Qubit dsDNA high-sensitivity assay (Thermo Fisher Scientific). DNA was further migrated on an Agilent 2100 Bioanalyser (Agilent Technologies) to confirm quality and distribution of DNA sizes between 300 and 700 bp.

For sequencing, 24 libraries were pooled together in equimolar concentrations (40 ng) and the pool was then diluted to 1 nM. The sequencing was conducted on a MiniSeq sequencer (Illumina, Inc., USA) using the MiniSeq System Mid-Output Kit 150 cycle (Illumina, Inc., USA) with custom primers (TruSeq Read1 and TruSeq_IndexRead1, Supplementary table S9), producing single-end reads of 151 bp.

Processing of sequencing data

Depending on the sample, we obtained between 337'734 and 1'367'243 raw sequence reads (151 bp). Reads without the 19-bp transposon end sequence were removed by using Cutadapt (version 2.5)(Martin, 2011). Next, we removed the Illumina adaptors using Trimmomatic (version 0.36, default parameters)(Bolger et al., 2014). With a custom-made in-house perl script we removed all the remaining sequences that contained only transposon sequence and no connecting genome insert position sequence. The filtered, trimmed and cleaned fastq files were then filtered for quality (fastQC version 0.11.5)(Andrews et al., 2010), and the filtered remaining reads were aligned and mapped onto the *P. veronii* genome using bowtie2 under standard stringency (version 2.3.4.1)(Langmead and Salzberg, 2012). Finally, the mapped reads were counted by an in-house perl script and analyzed in MatLab (R2016a).

Statistical analysis

To estimate the deviation between the observed per gene distributions of *aph* inserts in the two starting libraries (Lib1 and Lib2) and expected random insertion distributions without selective effects, we four times independently randomly subsampled 1.2×10^5 unique insertions (close to the observed number) without replacement from the *P. veronii* combined genome (chromosome 1, replicons 2 and 3). These were then grouped per gene to have four replicates of expected random distributions, normalized by total, and compared to the sum-normalized Lib1 and Lib2 per gene distributions using a *mattest* and *mafdr* in MATLAB (vs. 2016a) to estimate significant differences. Gene insertions were considered to cause immediate fitness effects if their ratio of mean observed to mean expected distributions was < 0.01 with a corrected q-value of < 0.01 in both starting libraries. The KEGG-pathway and COG attributions of the resulting gene list were compared to that of the full genome by calculating a confidence interval (at $p = 0.05$) on a ten times independently subsampled and attributed random gene list of the same size.

To identify genes whose insertion influence *P. veronii* fitness in Sand or Silt, we normalized all sequence libraries using the 'Pseudoreference Sample' model, taking into account the median read density across samples. We then created three categories of time (generation) trends among treatments: *up*, *down* and *flat*, from a regression analysis of the observed mean grouped normalized read densities per gene (function *regression* in MATLAB v. 2016a). Regression coefficients (rc) > 0.75 pointed to genes with insertions becoming more abundant over time and were categorized as *up*. Similarly, those with rc < -0.75 were categorized as *down*, and those with rc in between -0.75 and 0.75 as *flat* (Fig. 4, cutoffs chosen to encompass groups detected in rc distributions). In the *flat* category we further filtered those genes with starting read number of less than 2 to a category *zero* (most of which encompass

genes that were already identified above as essential for fitness at start). Resulting gene lists per treatment and category (*up*, *down*, *flat*) were further filtered to exclude those with a read density of <2 at start, attributed to COG categories and compared to expected COG attributions from the full genome by random subsampling procedures and confidence interval limits ($p = 0.05$).

We tested whether identified genes potentially implicated in fitness in soils were on average higher expressed than in liquid suspension. For this we compared selected *down* gene lists to transcriptomic data from *P. veronii* growing in Silt (Morales et al., 2021) and counted whether the number of *down* genes with expression above the observed median ($\log_2 \text{CPM} = 3$) is higher than expected from ten randomly sampled gene sets of the genome with same size set as the number of *down* genes ($p = 0.05$ confidence interval).

To find the overlap of fitness-affected *P. veronii* genes to the essential gene set of *Acinetobacter baylyi* ADP1 (Gallagher et al., 2020) and the soil-fitness gene set of *Sphingomonas wittichii* RW1 (Roggo et al., 2013) we compared gene sets in pair-wise BLASTp using thresholds of e-value < $1e-10$, % identity of >40 (for *A. baylyi*) or ratio overlap to query length of >0.75 (for *S. wittichii* RW1).

References

1. Akkaya, Ö., Pérez-Pantoja, D.R., Calles, B., Nickel, P.I., and de Lorenzo, V. (2018). The metabolic redox regime of *Pseudomonas putida* tunes its evolvability toward novel xenobiotic substrates. *MBio* 9, e01512-18.
2. Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C., and Wingett, S. (2010). FastQC. A quality control tool for high throughput sequence data. Babraham Inst.
3. Armalytė, J., Skerniškytė, J., Bakienė, E., Krasauskas, R., Šiugždinienė, R., Kareivienė, V., Kerzienė, S., Klimienė, I., Sužiedėlienė, E., and Ružauskas, M. (2019). Microbial diversity and antimicrobial resistance profile in microbiota from soils of conventional and organic farming systems. *Front. Microbiol.* 10, 892.
4. Blanco, P., Hernando-Amado, S., Reales-Calderon, J.A., Corona, F., Lira, F., Alcalde-Rico, M., Bernardini, A., Sanchez, M.B., and Martinez, J.L. (2016). Bacterial multidrug efflux pumps: much more than antibiotic resistance determinants. *Microorganisms* 4, 14.
5. Bojanovič, K., D'Arrigo, I., and Long, K.S. (2017). Global transcriptional responses to osmotic, oxidative, and imipenem stress conditions in *Pseudomonas putida*. *Appl. Environ. Microbiol.* 83, e03236-16.
6. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.
7. Cabiscol, E., Tamarit, J., and Ros, J. (2000). Oxidative stress in bacteria and protein damage by reactive oxygen species. *Int. Microbiol. Off. J. Spanish Soc. Microbiol.* 3, 3–8.
8. Chavarría, M., Nickel, P.I., Pérez-Pantoja, D., and de Lorenzo, V. (2013). The Entner–Doudoroff pathway empowers *Pseudomonas putida* KT2440 with a high tolerance to oxidative stress. *Environ. Microbiol.* 15, 1772–1785.
9. Coronado, E., Roggo, C., and van der Meer, J.R. (2014). Identification of genes potentially involved in solute stress response in *Sphingomonas wittichii* RW1 by transposon mutant recovery. *Front. Microbiol.* 5, 585.
10. Cunliffe, M., Kawasaki, A., Fellows, E., and Kertesz, M.A. (2006). Effect of inoculum pretreatment on survival, activity and catabolic gene expression of *Sphingobium yanoikuyae* B1 in an aged polycyclic aromatic hydrocarbon-contaminated soil. *FEMS Microbiol. Ecol.* 58, 364–372.
11. Domínguez-Cuevas, P., González-Pastor, J.-E., Marqués, S., Ramos, J.-L., and de Lorenzo, V. (2006). Transcriptional tradeoff between metabolic and stress-response programs in *Pseudomonas putida* KT2440 cells exposed to toluene. *J. Biol. Chem.* 281, 11981–11991.
12. Eberlein, C., Baumgarten, T., Starke, S., and Heipieper, H.J. (2018). Immediate response mechanisms of Gram-negative solvent-tolerant bacteria to cope with environmental stress: cis-trans isomerization of unsaturated fatty acids and outer membrane vesicle secretion. *Appl. Microbiol. Biotechnol.* 102, 2583–2593.
13. Fillet, S., Vélez, M., Lu, D., Zhang, X., Gallegos, M.-T., and Ramos, J.L. (2009). TtgV represses two different promoters by recognizing different sequences. *J. Bacteriol.* 191, 1901–1909.
14. Gallagher, L.A., Bailey, J., and Manoil, C. (2020). Ranking essential bacterial processes by speed of mutant death. *Proc. Natl. Acad. Sci. U. S. A.* 117, 18010–18017.
15. Gerhardt, P., Murray, R., Costilow, R., Nester, E., Wood, W., Krieg, N., and Phillips, G. (1981). *Manual of methods for general bacteriology* (Washington: American Society for Microbiology).
16. Goryshin, I.Y., Miller, J.A., Kil, Y. V., Lanzov, V.A., and Reznikoff, W.S. (1998). Tn5/IS50 target recognition. *Proc. Natl. Acad. Sci. U. S. A.* 95, 10716–10721.
17. Green, B., Bouchier, C., Fairhead, C., Craig, N.L., and Cormack, B.P. (2012). Insertion site preference of Mu, Tn5, and Tn7 transposons. *Mob. DNA* 3, 3.
18. Hardy, L., Juan, P.-A., Coupat-Goutaland, B., and Charpentier, X. (2021). Transposon insertion sequencing in a clinical isolate of *Legionella pneumophila* identifies essential genes and determinants of natural transformation. *J. Bacteriol.* 203, e00548-20.
19. Harrison, J.J., Almlblad, H., Irie, Y., Wolter, D.J., Eggleston, H.C., Randall, T.E., Kitzman, J.O., Stackhouse, B., Emerson, J.C., Mcnamara, S., et al. (2020). Elevated exopolysaccharide levels in *Pseudomonas aeruginosa* flagellar mutants have implications for biofilm growth and chronic infections. *PLOS Genet.* 16, e1008848.
20. Huertas, M.J., Duque, E., Marqués, S., and Ramos, J.L. (1998). Survival in soil of different toluene-degrading *Pseudomonas* strains after solvent shock. *Appl. Environ. Microbiol.* 64, 38–42.
21. Jacobs, M.A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., Will, O., Kaul, R., Raymond, C., Levy, R., et al. (2003). Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci.* 100, 14339–14344.

22. Johnson, D.R., Coronado, E., Moreno-Forero, S.K., Heipieper, H.J., and van der Meer, J.R. (2011). Transcriptome and membrane fatty acid analyses reveal different strategies for responding to permeating and non-permeating solutes in the bacterium *Sphingomonas wittichii*. *BMC Microbiol.* *11*, 250.
23. Junca, H., and Pieper, D.H. (2004). Functional gene diversity analysis in BTEX contaminated soils by means of PCR-SSCP DNA fingerprinting: comparative diversity assessment against bacterial isolates and PCR-DNA clone libraries. *Environ. Microbiol.* *6*, 95–110.
24. Kota, S., Borden, R.C., and Barlaz, M.A. (1999). Influence of protozoan grazing on contaminant biodegradation. *FEMS Microbiol. Ecol.* *29*, 179–189.
25. Krell, T., Lacal, J., Guazzaroni, M.E., Busch, A., Silva-Jiménez, H., Fillet, S., Reyes-Darías, J.A., Muñoz-Martínez, F., Rico-Jiménez, M., García-Fontana, C., et al. (2012). Responses of *Pseudomonas putida* to toxic aromatic carbon sources. *J. Biotechnol.* *160*, 25–32.
26. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* *9*, 357–359.
27. Liberati, N.T., Urbach, J.M., Miyata, S., Lee, D.G., Drenkard, E., Wu, G., Villanueva, J., Wei, T., and Ausubel, F.M. (2006). An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 2833–2838.
28. de Lima-Morales, D., Chaves-Moreno, D., Jarek, M., Vilchez-Vargas, R., Jauregui, R., and Pieper, D.H. (2013). Draft genome sequence of *Pseudomonas veronii* strain 1YdBTEX2. *Genome Announc.* *1*, e00258-13.
29. de Lima-Morales, D., Chaves-Moreno, D., Wos-Oxley, M.L., Jáuregui, R., Vilchez-Vargas, R., and Pieper, D.H. (2016). Degradation of benzene by *Pseudomonas veronii* 1YdBTEX2 and 1YB2 Is catalyzed by enzymes encoded in distinct catabolism gene clusters. *Appl. Environ. Microbiol.* *82*, 167–173.
30. Liu, F., Wang, C., Wu, Z., Zhang, Q., and Liu, P. (2016). A zero-inflated Poisson model for insertion tolerance analysis of genes based on Tn-seq data. *Bioinformatics* *32*, 1701–1708.
31. Martin, M. (2011). CUTADAPT removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* *17*.
32. Martínez, J.L., Sánchez, M.B., Martínez-Solano, L., Hernández, A., Garmendia, L., Fajardo, A., and Alvarez-Ortega, C. (2009). Functional role of bacterial multidrug efflux pumps in microbial natural ecosystems. *FEMS Microbiol. Rev.* *33*, 430–449.
33. Mendonca, C.M., Yoshitake, S., Wei, H., Werner, A., Sasnow, S.S., Thannhauser, T.W., and Aristilde, L. (2020). Hierarchical routing in carbon metabolism favors iron-scavenging strategy in iron-deficient soil *Pseudomonas* species. *Proc. Natl. Acad. Sci.* *117*, 32358–32369.
34. Mirleau, P., Wogelius, R., Smith, A., and Kertesz, M.A. (2005). Importance of organosulfur utilization for survival of *Pseudomonas putida* in soil and rhizosphere. *Appl. Environ. Microbiol.* *71*, 6571–6577.
35. Molina-Santiago, C., Daddaoua, A., Fillet, S., Duque, E., and Ramos, J.-L. (2014). Interspecies signalling: *Pseudomonas putida* efflux pump TtgGHI is activated by indole to increase antibiotic resistance. *Environ. Microbiol.* *16*, 1267–1281.
36. Molina-Santiago, C., Udaondo, Z., Gómez-Lozano, M., Molin, S., and Ramos, J.-L. (2017). Global transcriptional response of solvent-sensitive and solvent-tolerant *Pseudomonas putida* strains exposed to toluene. *Environ. Microbiol.* *19*, 645–658.
37. Morales, M., Sentchilo, V., Bertelli, C., Komljenovic, A., Kryuchkova-Mostacci, N., Bourdilloud, A., Linke, B., Goesmann, A., Harshman, K., Segers, F., et al. (2016). The genome of the toluene-degrading *Pseudomonas veronii* strain 1YdBTEX2 and its differential gene expression in contaminated sand. *PLoS One* *11*, e0165850.
38. Morales, M., Sentchilo, V., Hadadi, N., and van der Meer, J.R. (2021). Genome-wide gene expression changes of *Pseudomonas veronii* 1YdBTEX2 during bioaugmentation in polluted soils. *Environ. Microbiome* *16*, 8.
39. Moreno-Forero, S.K., and van der Meer, J.R. (2015). Genome-wide analysis of *Sphingomonas wittichii* RW1 behaviour during inoculation and growth in contaminated sand. *ISME J.* *9*, 150–165.
40. Moreno-Forero SK, Rojas E, Beggah S, van der Meer JR: Comparison of differential gene expression to water stress among bacteria with relevant pollutant-degradation properties. *Environ Microbiol Rep* 2016, **8**:91-102.
41. Murínová, S., and Dercová, K. (2014). Response mechanisms of bacterial degraders to environmental contaminants on the level of cell walls and cytoplasmic membrane. *Int. J. Microbiol.* *2014*, 873081.
42. Murray, G.L., Morel, V., Cerqueira, G.M., Croda, J., Srikram, A., Henry, R., Ko, A.I., Dellagostin, O.A., Bulach, D.M., Sermswan, R.W., et al. (2009). Genome-wide transposon mutagenesis in pathogenic *Leptospira* species. *Infect. Immun.* *77*, 810–816.
43. Nguyen, T.N., Yeh, C.W., Tsai, P.C., Lee, K., and Huang, S.L. (2016). Transposon mutagenesis identifies

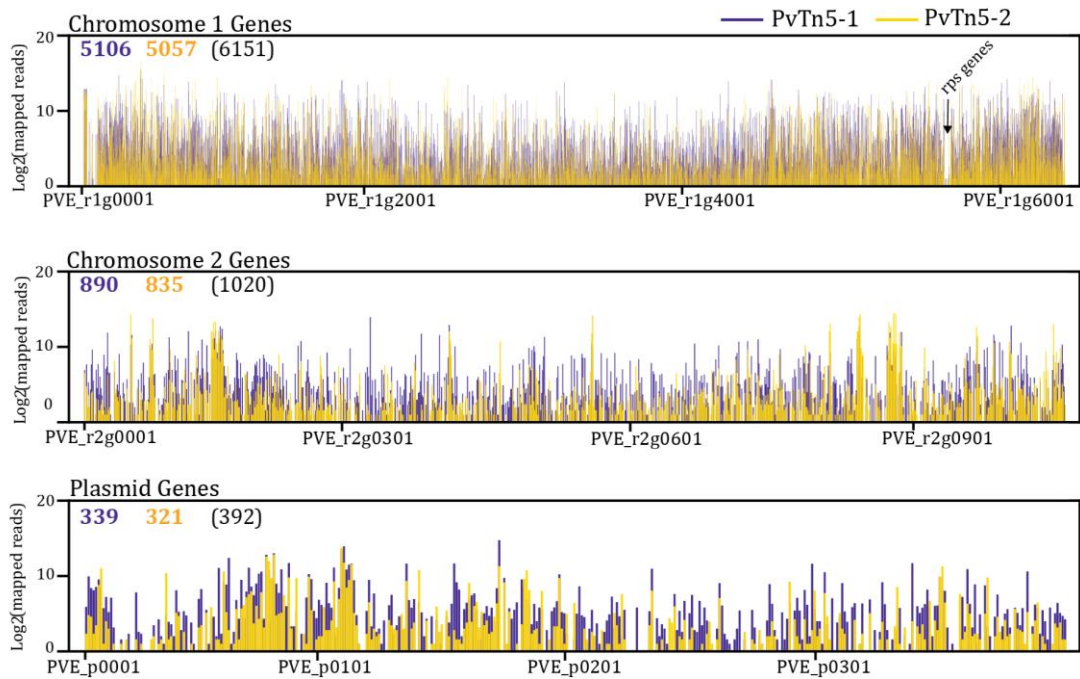
- genes critical for growth of *Pseudomonas nitroreducens* TX1 on octylphenol polyethoxylates. *Appl. Environ. Microbiol.* *82*, 6584–6592.
44. Ni, B., Colin, R., Link, H., Endres, R.G., and Sourjik, V. (2020). Growth-rate dependent resource investment in bacterial motile behavior quantitatively follows potential benefit of chemotaxis. *Proc. Natl. Acad. Sci.* *117*, 595 LP – 601.
 45. Nickel, P.I., Chavarría, M., Martínez-García, E., Taylor, A.C., and de Lorenzo, V. (2013). Accumulation of inorganic polyphosphate enables stress endurance and catalytic vigour in *Pseudomonas putida* KT2440. *Microb. Cell Fact.* *12*, 50.
 46. Nickel, P.I., Chavarría, M., Fuhrer, T., Sauer, U., and De Lorenzo, V. (2015). *Pseudomonas putida* KT2440 strain metabolizes glucose through a cycle formed by enzymes of the Entner-Doudoroff, embden-meyerhof-parnas, and pentose phosphate pathways. *J. Biol. Chem.* *290*, 25920–25932.
 47. Nickel, P.I., Fuhrer, T., Chavarría, M., Sánchez-Pascuala, A., Sauer, U., and de Lorenzo, V. (2021). Reconfiguration of metabolic fluxes in *Pseudomonas putida* as a response to sub-lethal oxidative stress. *ISME J.*
 48. van Opijnen, T., and Camilli, A. (2013). Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. *Nat. Rev. Microbiol.* *11*, 435–442.
 49. van Overbeek, L.S., Eberl, L., Givskov, M., Molin, S., and van Elsas, J.D. (1995). Survival of, and induced stress resistance in, carbon-starved *Pseudomonas fluorescens* cells residing in soil. *Appl. Environ. Microbiol.* *61*, 4202–4208.
 50. Pérez-Pantoja, D., Donoso, R., Agulló, L., Córdova, M., Seeger, M., Pieper, D.H., and González, B. (2012). Genomic analysis of the potential for aromatic compounds biodegradation in *Burkholderiales*. *Environ. Microbiol.* *14*, 1091–1117.
 51. Pieper, D.H., and Reineke, W. (2000). Engineering bacteria for bioremediation. *Curr. Opin. Biotechnol.* *11*, 262–270.
 52. Polonenko, D.R., Mayfield, C.I., and Dumbroff, E.B. (1986). Microbial responses to salt-induced osmotic stress. *Plant Soil* *92*, 417–425.
 53. Qin, A., Tucker, A.M., Hines, A., and Wood, D.O. (2004). Transposon mutagenesis of the obligate intracellular pathogen *Rickettsia prowazekii*. *Appl. Environ. Microbiol.* *70*, 2816–2822.
 54. Reva, O.N., Weinel, C., Weinel, M., Böhm, K., Stjepandic, D., Hoheisel, J.D., and Tümmler, B. (2006). Functional genomics of stress response in *Pseudomonas putida* KT2440. *J. Bacteriol.* *188*, 4079–4092.
 55. Rodríguez-Herva, J.J., García, V., Hurtado, A., Segura, A., and Ramos, J.L. (2007). The ttgGHI solvent efflux pump operon of *Pseudomonas putida* DOT-T1E is located on a large self-transmissible plasmid. *Environ. Microbiol.* *9*, 1550–1561.
 56. Roggo, C., Coronado, E., Moreno-Forero, S.K., Harshman, K., Weber, J., and van der Meer, J.R. (2013). Genome-wide transposon insertion scanning of environmental survival functions in the polycyclic aromatic hydrocarbon degrading bacterium *Sphingomonas wittichii* RW1. *Environ. Microbiol.* *15*, 2681–2695.
 57. Rojas, A., Segura, A., Guazzaroni, M.E., Terán, W., Hurtado, A., Gallegos, M.T., and Ramos, J.L. (2003). In vivo and in vitro evidence that TtgV is the specific regulator of the TtgGHI multidrug and solvent efflux pump of *Pseudomonas putida*. *J. Bacteriol.* *185*, 4755–4763.
 58. Sardesai, Y., and Bhosle, S. (2002). Tolerance of bacteria to organic solvents. *Res. Microbiol.* *153*, 263–268.
 59. Smith, G.A., and Enquist, L.W. (1999). Construction and transposon mutagenesis in *Escherichia coli* of a full-length infectious clone of pseudorabies virus, an alphaherpesvirus. *J. Virol.* *73*, 6405–6414.
 60. Sternon, J.-F., Godessart, P., Gonçalves de Freitas, R., Van der Henst, M., Poncin, K., Francis, N., Willemart, K., Christen, M., Christen, B., Letesson, J.-J., et al. (2018). Transposon sequencing of *Brucella abortus* uncovers essential genes for growth in vitro and inside macrophages. *Infect. Immun.* *86*, e00312-18.
 61. Subashchandrabose, S., Smith, S.N., Spurbeck, R.R., Kole, M.M., and Mobley, H.L.T. (2013). Genome-wide detection of fitness genes in uropathogenic *Escherichia coli* during systemic infection. *PLOS Pathog.* *9*, e1003788.
 62. Tyagi, M., da Fonseca, M.M.R., and de Carvalho, C.C.C.R. (2011). Bioaugmentation and biostimulation strategies to improve the effectiveness of bioremediation processes. *Biodegradation* *22*, 231–241.
 63. Wang, G., and Or, D. (2010). Aqueous films limit bacterial cell motility and colony expansion on partially saturated rough surfaces. *Environ. Microbiol.* *12*, 1363–1373.
 64. Wong, Y.-C., Abd El Ghany, M., Naeem, R., Lee, K.-W., Tan, Y.-C., Pain, A., and Nathan, S. (2016). Candidate Essential Genes in *Burkholderia cenocepacia* J2315 Identified by Genome-Wide TraDIS. *Front. Microbiol.* *7*, 1288.

Supplementary material

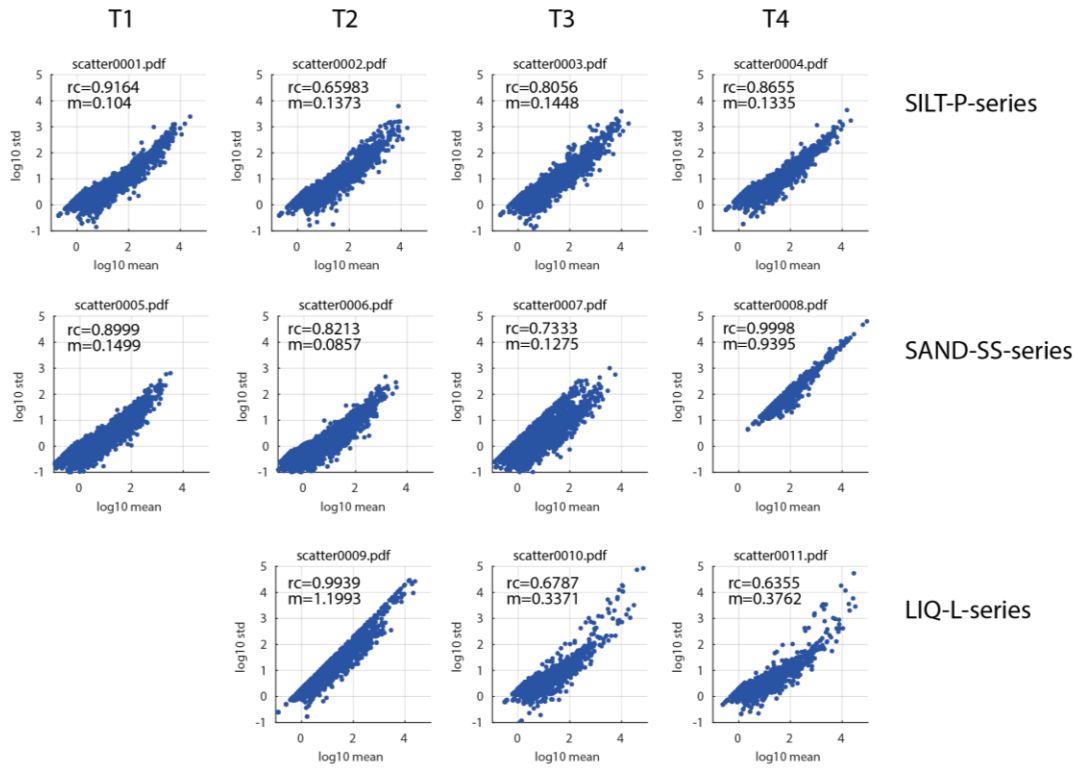
Due to the extensive length of the supplementary tables of this chapter, we provide them in digital form at the following link: <https://bit.ly/3vd9XwC>

The tables are:

- Table S1: Underrepresented *P. veronii* genes in both starting libraries.
- Table S2: Underrepresented *P. veronii* genes in the starting libraries in common with essential genes of *Acinetobacter baylyi* ADP1.
- Table S3: Inactivated *P. veronii* genes causing fitness increase common to soil growth.
- Table S4: Inactivated *P. veronii* genes causing fitness increase common to sand and liquid growth.
- Table S5: Changes in relative abundances of *P. veronii* genes implicated in aromatic compound metabolism.
- Table S6: List of genes with insertions causing fitness loss in soils.
- Table S7: Genes without any noticeable fitness effects.
- Table S8: Overlapping gene functions between *P. veronii* and *Sphingomonas wittichii* associated with fitness maintenance in soil.
- Table S9: List of primers and NEBNext adaptors used in this study.



Supplementary Figure 1. Gene insertion distribution in starting and incubated *P. veronii* transposon libraries. Average mapped number of reads (\log_2 transformed) per gene in each *P. veronii* replicon, for both starting libraries (in purple: PvTn5-1 and in yellow: PvTn5-2). The numbers inside the panels indicates the total number of genes with insertions per library and per replicon. The arrow in the top panel highlights a region without any mapped insertions, coding for ribosomal proteins.



Supplementary Figure 2. Gene insertion variability in each of the sequenced libraries.

GENERAL DISCUSSION

Understanding the adaptive cellular responses of individual bacterial strains during inoculation (i.e., lag phase, growth, and survival) helps to characterize the factors limiting bioaugmentation and will be crucial for engineering approaches that aim to introduce new functionalities into complex environments (Coronado et al., 2015; Pieper and Reineke, 2000; Sales da Silva et al., 2020). In the last decades, the variety of metabolic pathways and organisms relevant for biodegradation as well as bottlenecks for their implementation in the treatment of environmental pollution have been described (Junca and Pieper, 2004; Lee et al., 2019; Ramos et al., 2002; Segura et al., 2001; Torres et al., 2011). As a matter of fact, it is evident that we are just beginning to understand the natural diversity for biodegradation and are only starting to learn how to exploit the bioremediation potential of environmental isolates fully.

In this context, the overall aim of this work was to get a predictable knowledge of how inoculated bacteria strains react in complex environments and how these reactions determine inoculant survival. We studied the initial and long-term adaptive mechanisms of the monoaromatic compound-degrading bacterium *P. veronii* 1YdBTEX2 during inoculation and growth in non-sterile soils artificially contaminated with toluene. By combining two powerful tools, transcriptomic and transposon mutagenesis, we measured and interpreted the global reactions during adaptation and growth, the metabolic pathways responsible for the complete mineralization of the target pollutant. We also identified genes and functions essential for soil fitness in the presence of toluene. Our results describe the mechanism, capabilities, and limitations of *P. veronii* to survive in natural soils and polluted material.

BTEX compounds are the major aromatic hydrocarbons present in petroleum products. They are also the most relevant soil and groundwater contaminants (Lee et al., 2019). *P. veronii* was initially isolated from jet-fuel contaminated soil (Junca and Pieper, 2004; de Lima-Morales et al., 2013). The first task of this project was to obtain a full and gapless genome sequence to verify the organization of the BTEX catabolic pathways as well as additional key features that might be responsible for its prevalence in environments under BTEX stress. We confirmed previous studies that suggested that BTEX degradation proceeds via an initial deoxygenating reaction (*ipbAa-Ad*) followed by meta-cleavage pathways for catechol derivatives (de Lima-Morales et al., 2016). Interestingly, the *P. veronii* genome has three replicons, one of which (a mega-plasmid) encodes three redundant meta-cleavage pathways (*ipb*, *dmp*, and *nah* gene clusters). Only two of these are induced by toluene and caused significant fitness loss (*ipb* and *dmp*) when disrupted by transposon insertions. Our transcriptomic analysis allowed us to propose a candidate gene (PVE_r2g0805) to complement the function of the dihydrodiol dehydrogenase (*ipbB*) abolished by a non-sense mutation.

Besides the unique BTEX degradation pathway, we also identified that features such as siderophore synthesis, a complete denitrification pathway, and efflux systems (including toluene efflux pumps) might facilitate growth in soil.

One of the first questions we addressed in this work was whether *P. veronii* would survive the transition from liquid media to different soils and establishes itself in the presence of a target pollutant (toluene). The poor survival of the inoculated strains – which undoubtedly translates into an insufficient catalytic activity and pollutant removal – is one commonly suggested reason for which bioaugmentation does not produce the expected results (Radwan et al., 2019; Thompson et al., 2005; Widder et al., 2016). Nevertheless, the reasons for such a pitfall are not entirely understood. Our systematic study shows that, in fact, *P. veronii* can divide and maintain a stable population for extended periods in natural and polluted materials (Morales et al., 2021), thus controverting the claim of poor growth associated with inoculation of exogenous strains. Notwithstanding, our results also suggested that one of the main plights of the strain is the competition with the local resident microbial community and the risk for predation by grazing protozoa. Transcriptomic results indicated that the strain responds effectively to the local soil conditions during the different growth phases (e.g., nutrient availability, water potential, and the presence of other toxic compounds) in the presence of toluene, without displaying any sign of extreme stress or cellular death (Morales et al., 2016, 2021). Despite that, there were two scenarios in which both the wild-type strain and transposon mutants significantly diminished population sizes (Clay and Silt, respectively). Both of these soils have higher resident microbial backgrounds than e.g., Sand. Thus, we hypothesize that *P. veronii* is instead a friendly colonizer despite having at least three genomic regions coding for the type VI systems (T6SS) and two tailocin biosynthetic gene clusters. It would be interesting to explore the killing capacity of inoculated strains to resist or compete with local bacteria and eukaryotic species that feed on them. Under the evaluated conditions, we did not observe specific induction of genes of the T6SS or tailocins. Thus a better understanding of the conditions under which these machinery are activated might help to improve the long-term competitive fitness of the strain.

The findings in the current work provided solid evidence that *P. veronii* mutants with impaired flagella biosynthesis had a significant fitness increase in the soil. As discussed in Chapter 3, the two-edge sword of flagella assembly and maintenance represents an evolutionary conundrum. On the one hand, flagella allow bacteria to explore the environment, searching for nutrients and help to escape from predators or adverse conditions. On the other hand, mutants without flagellar machinery are more resistant to environmental oxidative stress due to higher metabolic energy in the form of ATP and reducing power (NADPH). Flagellin monomers are important inducers of plant and animal innate immunity. Therefore, the control

of their expression has been shown to be an important evasion strategy in plant and human pathogens (Rossez et al., 2015). One can imagine that for *P. veronii*, the flagella might act as a signal for recognition by predator amoeba, whereby turning them off or losing them present a fitness gain. We hypothesize that *P. veronii* long-term survival in the soil thus requires low or temporarily restricted expression of flagellar systems in order to reach favorable niches and optimize metabolic energy. In a later invasion stage, to avoid recognition and predation, flagella synthesis should be turned off. In this regard, more strain-specific studies need to be designed to clarify the actual role of flagella biosynthesis in the long-term survival of motile strains attempted to be used for bioremediation. It would be interesting to systematically evaluate the survival strategies of non-motile *P. veronii* mutants in conditions and scenarios not covered by the current work.

Soil is considered a complex environment. The low availability of nutrients, limited oxygen, and extreme fluctuations in water content, among others, make it a hostile environment to be colonized by foreign bacteria (Ali et al., 2020; Sales da Silva et al., 2020). As mentioned, when exogenous bacteria are inoculated into soils, they also have to compete with the local resident community for the scarce resources and avoid predation. Thus, the successful bioaugmentation of inoculated strains requires major adjustments in the physiological and metabolic programs to react fast enough to constantly changing soil conditions. In a bioremediation context, the inoculated strain should sense and use the target toxic compounds as both a potential carbon source to be metabolized, although at the same time it can act as a toxic membrane-damaging and macromolecule-disrupting agent. Therefore, an adjustment and precise response to all these stimuli require a fine-tuned and dedicated transcriptional response that involves a large number of genes. As we showed, the transition from liquid media to Sand involves significant transcriptomic changes of at least one-third of the genes encoded in the genome (Chai et al., 2016; Morales et al., 2016; Moreno-Forero and van der Meer, 2015). Perhaps one of the major challenges of the approaches used in this study is the interpretation of such massive changes, especially when a considerable proportion of the genes that seem to play an important role in the adaptation are poorly annotated. Therefore, with the currently available tools, it is impossible to give a biological meaning to their role in the general adaptation to and growth in soil.

Another challenge is the correct selection of adequate control conditions. Liquid suspended growth is commonly used as a control to evaluate changes associated with soil inoculations. However, our findings support other studies that indicate that liquid controls may not be the best choice. In an attempt to find a better alternative to be used as a control condition, we evaluated an inert artificial porous matrix (AMP) as a proxy of soil porosity. In AMP, we observe

a reduced growth of the strain and changes in the regulation of toluene metabolism, most likely due to higher toluene toxicity or differences in oxygen availability.

Additionally, the transcriptional changes were not as dramatic as the ones observed when using liquid cultures. That allows a better interpretation of the potential effects originating from different soil types from porosity itself. Our results suggest that more attention should be given to the correct selection of adequate control conditions.

Additionally, bulk genomic techniques such as RNA-seq and Tn5-seq can only assess average properties across thousands of cells, thereby ignoring the stochasticity of gene expression (Rossez et al., 2015), losing the resolution of the different subpopulations states and dynamics during soil invasion. Although at the moment, it is challenging to measure subpopulation responses and to compare such subpopulations across samples, such information is crucial for understanding the heterogeneity of the cells' responses during growth in soil and for comparative analysis of samples from different conditions.

Recent technological advances facilitate the study of biological systems at single-cell level resolution (e.g., single-cell RNA sequencing). These advances are coupled with the development of powerful computational methods to identify subpopulations of cells present in complex samples.

After observing that the strain was able to colonize the different soils and materials used in this study, we wondered whether the transcriptomic responses could be generalized in a so-called "soil-transcriptomic" program. We compared transcriptomic responses in different soils and materials at different growth phases, covering a broad range of conditions. Our results showed that under the tested conditions, the adaptation of *P. veronii* does not induce a unique program. We observed that *P. veronii* rather maintains a robust metabolic program characterized by a core of common functions (e.g., nitrate respiration, osmoregulation, amino acid recycling, oxidative stress, and copper homeostasis) regardless of the environment. Some of the key responses and gene functions were previously pointed out for other soil inoculants (e.g., osmotic regulation program, solvent efflux and several catabolic gene clusters and regulators). Some of them were also highlighted as important functions for soil survival by mutant screening.

Additionally, we also identified specific functions associated with each environment that reflected different needs, such as different cofactor regeneration, the selective induction of transporter systems, or diverse stress defense systems. We also observed that in later growth phases, the strain showed signs of nutrient limitations such as nitrogen starvation in Jonction or sulfur limitations in Silt, indicating that nutrient limitation determines long-term colonization failure in polluted soils.

As discussed throughout the thesis, *P. veronii* cells adapt seemingly easily to the local conditions by using similar core functions to maintain their central metabolism irrespective of their environment. Due to its large redundant genome, the same functions can be carried out by different genes, which gives some versatility when it comes to colonizing new niches, an adaptive response to maximize inclusive fitness in uncertain times. The screening of conditionally essential genes showed, as expected that genes involved in toluene metabolism and efflux systems conferred fitness advantage to aromatic compounds. We found very little overlap between genes whose mutation caused fitness loss in both soils (e.g., urea and short-chain fatty acid metabolism, oxidative stress defense, and nutrient/element transport). Nevertheless, the most evident result was that mutants with impaired motility had a significant advantage for long-term survival in the soils.

In conclusion, we demonstrated that it is possible to study and interpret global reactions of *P. veronii* during colonization of non-sterile natural soils, despite the complexity of the reactions and the plethora of changes that adaptation implies. One of the major bottlenecks for the interpretation of our data remains the poor annotation of hundreds of genes that were highly induced during growth in soils or whose inactivation was associated with significant fitness loss and the limitations inherent to the bulked sample analysis. The characterization of genes with hypothetical functions will be crucial for better understanding bacteria mechanisms for soil adaptations. Overall, the results of this project give a global picture of the strategies and capabilities of *P. veronii* to grow and survive in contaminated soils. It also highlights that although some of the mechanisms are commonly shared with other strains (e.g., *Sphingomonas*), the analysis of the physiological response and limitations have to be done at the strain level due to differences in global regulatory mechanisms and gene redundancies.

Taken together this knowledge may help to develop protocols for 'preparing' bioaugmentation inoculums by triggering specific programs or optimize physiology to increase survival (e.g., to optimize growth media composition, to add supplements to overcome limitations of the on-site bioremediation, or prepare an efficient microbial consortium). Potentially structural components of the 'programs' can be improved and optimized by genetic engineering to produce more robust and equipped strains (superbug) to cope in a specific environment with, for instance, a new catabolic potential. At the moment, though, the empirical approach of screening and selecting strong candidates from the polluted sites and optimizing on-site bioremediation protocols seems a reasonable starting point. Although studies like this one designed to understand the interaction between xenobiotics and organisms on the fate, survival, and activities of microorganisms in contaminated soils described the environmental behavior of potential biodegraders, these findings have to be combined with the biochemical and genetic engineering studies to provide

the ground for successful interventions into environmental processes and, thereby, lead to optimized strategies for bioremediation.

References

1. Ali, N., Dashti, N., Khanafer, M., Al-Awadhi, H., and Radwan, S. (2020). Bioremediation of soils saturated with spilled crude oil. *Sci. Rep.* *10*, 1116.
2. Chai, B., Tsoi, T. V., Iwai, S., Liu, C., Fish, J.A., Gu, C., Johnson, T.A., Zylstra, G., Teppen, B.J., Li, H., et al. (2016). *Sphingomonas wittichii* strain RW1 genome-wide gene expression shifts in response to dioxins and clay. *PLoS One* *11*, 1–14.
3. Coronado, E., Valtat, A., and van der Meer, J.R. (2015). *Sphingomonas wittichii* RW1 gene reporters interrogating the dibenzofuran metabolic network highlight conditions for early successful development in contaminated microcosms. *Environ. Microbiol. Rep.* *7*, 480–488.
4. Junca, H., and Pieper, D.H. (2004). Functional gene diversity analysis in BTEX contaminated soils by means of PCR-SSCP DNA fingerprinting: comparative diversity assessment against bacterial isolates and PCR-DNA clone libraries. *Environ. Microbiol.* *6*, 95–110.
5. Lee, Y., Lee, Y., and Jeon, C.O. (2019). Biodegradation of naphthalene, BTEX, and aliphatic hydrocarbons by *Paraburkholderia aromaticivorans* BN5 isolated from petroleum-contaminated soil. *Sci. Rep.* *9*, 860.
6. Morales, M., Sentchilo, V., Bertelli, C., Komljenovic, A., Kryuchkova-Mostacci, N., Bourdilloud, A., Linke, B., Goesmann, A., Harshman, K., Segers, F., et al. (2016). The genome of the toluene-degrading *Pseudomonas veronii* strain 1YdBTEX2 and its differential gene expression in contaminated sand. *PLoS One* *11*, e0165850.
7. Morales, M., Sentchilo, V., Hadadi, N., and van der Meer, J.R. (2021). Genome-wide gene expression changes of *Pseudomonas veronii* 1YdBTEX2 during bioaugmentation in polluted soils. *Environ. Microbiome* *16*, 8.
8. Moreno-Forero, S.K., and van der Meer, J.R. (2015). Genome-wide analysis of *Sphingomonas wittichii* RW1 behaviour during inoculation and growth in contaminated sand. *ISME J.* *9*, 150–165.
9. Pieper, D.H., and Reineke, W. (2000). Engineering bacteria for bioremediation. *Curr. Opin. Biotechnol.* *11*, 262–270.
10. Radwan, S.S., Al-Mailem, D.M., and Kansour, M.K. (2019). Bioaugmentation failed to enhance oil bioremediation in three soil samples from three different continents. *Sci. Rep.* *9*, 19508.
11. Ramos, J.L., Duque, E., Gallegos, M.-T., Godoy, P., Ramos-Gonzalez, M.I., Rojas, A., Teran, W., and Segura, A. (2002). Mechanisms of solvent tolerance in gram-negative bacteria. *Annu. Rev. Microbiol.* *56*, 743–768.
12. Rossez, Y., Wolfson, E.B., Holmes, A., Gally, D.L., and Holden, N.J. (2015). Bacterial flagella: twist and stick, or dodge across the kingdoms. *PLoS Pathog.* *11*, e1004483–e1004483.
13. Sales da Silva, I.G., Gomes de Almeida, F.C., Padilha da Rocha e Silva, N.M., Casazza, A.A., Converti, A., and Asfora Sarubbo, L. (2020). Soil bioremediation: overview of technologies and trends. *Energies* *13*.
14. Segura, A., Duque, E., Hurtado, A., and Ramos, J.L. (2001). Mutations in genes involved in the flagellar export apparatus of the solvent-tolerant *Pseudomonas putida* DOT-T1E strain impair motility and lead to hypersensitivity to toluene shocks. *J. Bacteriol.* *183*, 4127 LP – 4133.
15. Thompson, I.P., Van Der Gast, C.J., Ciric, L., and Singer, A.C. (2005). Bioaugmentation for bioremediation: the challenge of strain selection. *Environ. Microbiol.* *7*, 909–915.
16. Torres, S., Pandey, A., and Castro, G.R. (2011). Organic solvent adaptation of Gram positive bacteria: Applications and biotechnological potentials. *Biotechnol. Adv.* *29*, 442–452.
17. Widder, S., Allen, R.J., Pfeiffer, T., Curtis, T.P., Wiuf, C., Sloan, W.T., Cordero, O.X., Brown, S.P., Momeni, B., Shou, W., et al. (2016). Challenges in microbial ecology: building predictive understanding of community function and dynamics. *ISME J.* *10*, 2557–2568.

CURRICULUM VITAE

Marian MORALES

EDUCATION

PhD candidate in Life Sciences, University of Lausanne	2014 – 2021
Master in Engineering, University of Valle (Colombia)	2009 – 2013
Bachelor in Microbiology, University of Valle (Colombia)	2003 – 2008

LIST OF PUBLICATIONS

Genome-wide gene expression changes of *Pseudomonas veronii* 1YdBTEX2 during bioaugmentation in polluted soils.

Morales M., Sentschilo V., Hadadi N., van der Meer, J.R.

Environmental Microbiome 16, 8 (2021). <https://doi.org/10.1186/s40793-021-00378-x>

Mechanistic insights into bacterial metabolic reprogramming from omics-integrated genome-scale models.

Hadadi N., Pandey V., Chiappino-Pepe A., Morales M., Gallart-Ayala H., Mehl F., Ivanisevic J., Sentschilo V., van der Meer JR.

npj Syst Biol Appl 6, 1 (2020). <https://doi.org/10.1038/s41540-019-0121-4>

The genome of the toluene-degrading *Pseudomonas veronii* strain 1YdBTEX2 and its differential gene expression in contaminated Sand.

Morales M., Sentschilo V., Bertelli C., Komljenovic A., Kryuchkova-Mostacci N., Bourdilloud A., Linke B., Goesmann A., Harshman K., Segers F., Delapierre F., Fiorucci D., Seppey M., Trofimenco E., Berra P., El Taher A., Loiseau C., Roggero D., Sulfiotti M., Etienne A., Ruiz Buendia G., Pillard L., Escoriza A., Moritz R., Schneider C., Alfonso E., Ben Jeddou F., Selmoni O., Resch G., Greub G., Emery O., Dubey M., Pillonel T., Robinson-Rechavi M., van der Meer JR.

PLOS ONE 11(11): e0165850. (2016). <https://doi.org/10.1371/journal.pone.0165850>

Chapter 7 - Ozonolysis.

Travaini R., Marangon-Jardim C., Colodette J., Morales-Otero M., and Bolado-Rodríguez S. Pretreatment of Biomass. Elsevier, (2015). ISBN 9780128000809.

<https://doi.org/10.1016/B978-0-12-800080-9.00007-4>.

Sugarcane bagasse ozonolysis pretreatment: effect on enzymatic digestibility and inhibitory compound formation

Travaini R., Morales-Otero MD., Coca M., Da-Silva R., Bolado S. Bioresource Technology 133 (2013) 332–339. <https://doi.org/10.1016/j.biortech.2013.01.133>.

LIST OF CONFERENCES

Swiss Meeting in Environmental Microbiology
7th swiss Microbial ecology Meeting SME 2019,
Lausanne, Switzerland
30.01.19-01.02.2019

3rd International Sytem X.ch
Conference on Systems Biology,
Zürich, Switzerland
4-7.09.2017

Annual Meeting of the SSM
74th Annual Meeting SSM and SSM satellite meeting, 13-15.
[Oral and Poster Presentation]
Bern, Switzerland
06.2016

Acknowledgements

First and foremost I would like to thank my supervisor Prof. Jan Roelof van der Meer for letting me do my Ph.D. in his lab and for giving me all his support. His guidance was crucial to reach and cross the finishing line. Thanks for hearing all my complaints, for all the encouraging words and for your trust. Thanks for your supervision, for always making me look at every result from a different perspective.

My sincere gratitude goes also to Professors Dietmar Piper, Nico Boon, Michel Chapuisat, and Tadeusz Kawecki for the interesting discussions and comments before and during my Ph.D. exam.

Furthermore, I am deeply grateful to Luc Patiny for believing in me, encouraging me, and for opening the first doors of this great adventure to me. Thank you so much for helping me take the first step. Thank you for welcoming me into your home with your lovely family when I first arrived in Switzerland.

I would like to thank all past and present members of Jan's lab - especially, Manupriyam, Diogo, Roxane, Nicolas, Andrea V., Xavier, Clémence, Andrea D., and Senka for all their contributions, support, and time for discussions. Thanks to them, it was never boring in the lab. A special acknowledgment goes to Vladimir, for his wonderful guidance regarding not only laboratory matters, but also intellectual guidance of my projects. Thank you for giving me a hand when mine was broken. Thank you for always being there for me especially when I needed it the most. Thank you for your friendship. T&T.

I acknowledge every single person at DMF that helped me to accomplish this research on a scientific, social, administrative, and personal level. I was blessed working in such a constructive, stimulating, and healthy environment. I thank The Swiss Initiative in Systems Biology, SystemsX, for financially supporting this work.

I met amazing people that became my family. Thanks to Silvia M., Monica, Annabelle, Noushin, Shiam, Ania, Silvia B., Rita, Björn, German, Marc, Flo, Domi, Lucie, and Damiano for being such amazing friends, you guys make my life in Switzerland much easier.

I am grateful of having such a supportive family. Knowing that you trust me and you are proud of me no matter what, encourages me to go for more and more each time. Los amo! Finally, I would like to thank Marco for his tremendous support and patience. For being a never-ending source of inspiration. My life would not be the same without you!

Lausanne, August 6 2021

