**Overflow metabolism and redox homeostasis underly swarming diversity in pathogenic bacteria**

**Metabolic constrains control the cooperative swarming diversity in pathogenic bacteria**

Guillem Santamaria1,3+, Chen Liao1+, Zhe Wang2, Kyu Rhee2, Francisco Pinto3, Jinyuan Yan1\*, Joao B. Xavier1\*

1Program for Computational and Systems Biology, Memorial Sloan-Kettering Cancer Center, New York, NY, USA

2Department of Medicine, Weill Cornell Medical College, New York, NY

3Departamento de Biologia, Universidade de Lisboa, Portugal

+these authors contributed equally

\*Correspondence: [yanj2@mskcc.org](mailto:yanj2@mskcc.org) and xavierj@mskcc.org

**Abstract**

Many species of bacteria have an incredible ability to move cooperatively over surfaces in swarms. Although swarming cooperation is deemed important to bacterial fitness and virulence, strains of the same species can vary widely in their swarming phenotype. Here we compare the metabolomes of 28 clinical *Pseudomonas aeruginosa* isolates from hospitalized patients to find a mechanism for intraspecies swarming diversity. We found that the isolates incapable of producing rhamnolipids—a surfactant necessary for swarming—had perturbed tricarboxylic acid (TCA) cycle and amino acid pathways, and grew exponentially slower in glycerol. Both metabolomics signatures and computer simulations using a genome-scale model advanced a hypothesis that the non-producers of rhamnolipids have elevated oxidative stress that impairs the TCA cycle, slows down growth, and shuts down carbon overflow via rhamnolipids which may save precious carbon resources for cell maintenance. Validation experiments supported the hypothesis, revealing that rhamnolipids as cooperative secreted products link intracellular redox homeostasis—a individual-level trait—to swarming—a population-level behavior. Along with the diversity of accessory genomes, redox balance partly explains the phenotypic diversity of swarming found across *P. aeruginosa*, a major cause of human infection.

**Significance**

Bacteria have remarkable displays of cooperation. The pathogen *Pseudomonas aeruginosa* swarms by cooperatively secreting surfactants called rhamnolipids to lubricate surfaces. To understand why some *P. aeruginosa* strains swarm and others do not, we combined metabolomics, computational modeling and experiments to study diverse swarming behaviors observed across 28 isolates of *Pseudomonas aeruginosa* obtained from infected patients. We demonstrated that rhamnolipid production is allowed when strains are able to maintain redox homeostasis with lower ROS (reactive oxygen species) generation. As a consequence, the rhamnolipid producers can become non-producers in a growth medium that induces greater ROS production. Rhamnolipid biosynthesis—serving as an overflow of metabolism into a cooperative secretion—couples the need that single cells have to reduce internal stress with the benefit that swarming brings to the entire population. This mechanism links single cell physiology and a collective behavior key to the fitness and virulence of pathogenic bacteria.

**Synopsis**

**A picture containing clock

Description automatically generated**

This study combined metabolomics, computational modeling and experiments to systematically investigate the swarming diversity of *Pseudomonas aeruginosa* obtained from infected patients, yielding insights into the genetic and metabolic contexts that control the bacterial swarming behavior

* Rhamnolipid secretion is a necessary but not sufficient condition of swarming
* *P. aeruginosa* strains that are unable to produce rhamnolipids in the glycerol minimal medium have perturbed metabolism that leads to slower growth and elevated oxidative stress
* These rhamnolipid producers can become non-producers in the same medium with succinate as the sole carbon source that induces greater ROS (reactive oxygen species) production
* The requirement of redox homeostasis for rhamnolipid production serves as an effective mechanism that prevents overflow of precious carbon for population-level benefit unless the benefit outweighs its associated cost
* The study represents an important step towards linking single cell physiology with collective behavior key to the fitness and virulence of pathogenic bacteria.

**Introduction**

Swarming enables millions of bacteria to travel together across centimeter-long distances in a few hours (Deforet *et al*, 2014). This cooperative trait is typically found in bacteria of the phyla Alpha-protebacteria and Gamma-proteobacteria, suggesting that swarming is a complex phenotype that cannot be evolved by simply acquiring a few set of genes: Swarming requires flagella or pili to propel the cells, cell-cell interactions to maintain population integrity during migration, and the collective secretion of biosurfactants to lubricate the surface (Kearns, 2010). Studies over the past two decades revealed many molecular details of swarming motility through genetic (e.g., knockouts, transgenesis) and environment (e.g., nutrient composition and viscosity of culture medium) perturbations (Köhler *et al*, 2000; Mattingly *et al*, 2018) as well as through experimental evolution (van Ditmarsch *et al*, 2013). Many genes that modulate swarming impact key cellular activities including tricarboxylic acid (TCA) cycle and stress response (Inoue *et al*, 2007; Tremblay & Déziel, 2010) and therefore have pleiotropic effects. Despite the many insights, we lack a systems-level understanding of this cooperative bacterial behavior (Kim & Surette, 2004; Inoue *et al*, 2007).

*Pseudomonas aeruginosa*—an opportunistic human pathogen and a major cause of hospital-acquired infections (Klevens *et al*, 2007)—has a remarkable swarming ability that produces long straight segments (tendrils) in its fractal-like swarming pattern. Similar to other swarming species, *P. aeruginosa* requires both flagella and pili to move, LasR-LasI and RhlR-RhlI quorum sensing systems to communicate within population, and rhamnolipid production to lubricate the surface (Köhler *et al*, 2000). Rhamnolipids are a mixture of biosurfactants that consist of 3-(3-hydroxyalkanoyloxy) alkanoic acids (HAAs), mono-rhamnolipids and di-rhamnolipids (Abdel-Mawgoud *et al*, 2010). RhlA is the only enzyme required to drive the conversion of fatty acid biosynthesis intermediates (β-hydroxyacyl-ACP) to HAA (Zhu & Rock, 2008), whereas RhlB and RhlC each conjugates one molecule of rhamnose to HAAs to produce mono-rhamnolipids and di-rhamnolipids in sequential steps (Chong & Li, 2017). The genes encoding RhlA and RhlB are located in the same operon *rhlAB*, which is regulated by the quorum-sensing cascade headed by LasR-LasI and followed by RhlR-RhlI (Medina *et al*, 2003). Other than quorum-sensing signals, *rhlAB* expression is also controlled by nutrient cues such as an excess of carbon source (Xavier *et al*, 2011; Boyle *et al*, 2015; Mellbye & Schuster, 2014). The combination of quorum-sensing and metabolic signals regulating swarming prevents cheating by rhamnolipid non-producers such as the Δ*rhlA* mutant by prudently ensuring that rhamnolipids are only produced at the right time when the production benefits outweigh its costs. Strains that produce rhamnolipids too soon or overproduce rhamnolipids, such as a strain engineered with the inducible PBAD-*rhlA* construct, have less carbon available for growth and are susceptible to cheating by non-producers (Xavier *et al*, 2011; de Vargas Roditi *et al*, 2013).

Besides the role of metabolic prudence, which regulates rhamnolipid production and makes swarming robust to cheating, other aspects of metabolism might have influenced the evolution of swarming across the *P. aeruginosa* phylogenetic tree. Metabolism is the currency of all physiological processes that support life (Smith & Morowitz, 2004) and a major determinant of social behavior (Biro & Stamps, 2010). One previous study (Boyle *et al*, 2017) attempted to compare the intracellular metabolites of the laboratory strain PA14 and mutants of that strain with different swarming phenotypes. The Δ*rhlA* mutant which lacks the ability to produce rhamnolipids and swarming, had less than 10 metabolites perturbed compared to the wild-type. The Δ*cbrA* mutant, which lacks a global regulator of carbon-nitrogen balance, had impaired swarming and lower growth rate with dozens of perturbed metabolites. Notably, the Δ*cbrA* mutant still had the flagella needed to move and secreted even more rhamnolipids than the wild-type. More interestingly, its swarming phenotype could be rescued by compensatory point mutations in the RNA chaperone *hfq* through rewiring of metabolic network to reconstruct a unique metabolome which was distinct from both the Δ*cbrA* and the wild-type.

Here we employed a comparative approach in a wide search for metabolic constraints on *P. aeruginosa* swarming. We profiled the metabolomes of 28 clinical isolates from hospitalized cancer patients in Memorial Sloan-Kettering Cancer Center (MSKCC) (Yan *et al*, 2017, 2019). Compared to the mutants of PA14 (Boyle *et al*, 2017), clinical isolates evolved independently in environments that are largely unknown to us, and they display a wide range of swarming phenotypes that cannot be easily explained by their wide genomic diversity (Yan *et al*, 2017). By combining metabolomics with growth curve analysis, we unveiled the characteristics of rhamnolipid non-producers including reduced TCA cycle activity, increased levels of amino acids, and slower growth, and further collected experimental evidences these metabolic changes are caused by increased oxidative stress. Our results suggest that rhamnolipid production as a form of overflow metabolism signifies the redox state of *P. aeruginosa* cells and their ability to detoxify reactive oxygen species (ROS) produced by energy metabolism, particularly the TCA cycle. The mechanism we propose in this study links redox homeostasis with cooperative secretions, highlighting the importance of bacterial metabolism to the evolution of social behaviors.

**Results**

**Swarming phenotype varies across the phylogenetic tree.** We started by analyzing the shapes of swarming colonies of 28 clinical isolates obtained from infected cancer patients whose genomes vary widely (Fig. 1). We used image analysis to measure five morphological features, including area (the percentage of the plate occupied by the colony), maximum length (maximum diagonal of the rectangle fitting the colony), circularity, eccentricity and length of image skeleton (Materials and Methods). A principal component analysis of the morphological features revealed that maximum length and circularity captured most of the shape diversity (Supplementary. Fig. S1). A swarming score as a linear combination of the two features showed no apparent correlation with phylogeny of the core genomes (Moran’s I test, p-value=0.93) or with the tissue of origin of each isolate (#do we need a p-value here?#). Since production of rhamnolipids is required for swarming motility, we measured the ability of strains to produce rhamnolipids in a synthetic media with glycerol as the sole carbon source (Boyle *et al*, 2015). As expected, all strains able to swarm also produced rhamnolipids. Not all rhamnolipid producers swarmed, however, which confirms that rhamnolipids are necessary but not sufficient for swarming (Caiazza *et al*, 2005; Déziel *et al*, 2003). Rhamnolipid production was also uncorrelated with phylogeny (Fig. 1B**,** Moran’s I test, p-value=0.14). To better assess the link between swarming and rhamnolipid production, we built an evolutionary model to reconstruct their ancestral states along the phylogenetic tree. The model predicted that the common ancestor of all the strains can both swarm and produce rhamnolipids, suggesting that the diversity of swarming found across the isolates is explained by some strains losing their ability to swarm or produce rhamnolipids, not by independent lineages evolving those functions independently. The model also shows that rhamnolipid production is more conserved than swarming: the immediate ancestors of all non-producers are inferred to be rhamnolipid producers (Supplementary Figure S2A). By contrast, the loss of swarming occurred earlier than loss of rhamnolipids in several lineages (Supplementary Figure S2B), which agrees with the fact that swarming depends on other factors besides rhamnolipid production.

**Swarming diversity is partially explained by the genes absent in rhamnolipid non-producers.** Since the phylogeny inferred by core genes was unable to explain the observed swarming diversity, we asked whether accessory genes, particularly those missing in rhamnolipid non-producers, could explain the loss of swarming. Here we focus on accessory genes in the non-producers, rather than non-swarmers, because the latter constitutes the majority of strains with too many accessory genes to study in details. By contrast, 8 rhamnolipid non-producers all lose swarming and they represent a much smaller group for genetic analysis. Using reciprocal blast to infer orthologous relationships (Materials and Methods), we identified 304 genes (Supplementary File 1) only absent in those non-producers that may directly contribute to the loss of rhamnolipid production (89 well annotated genes are listed in Table 1): F5677 lacks the flagella motor switch protein *fliM*, W36662 lacks the type IV pili assembly protein *pilC* and its sensor kinase *pilS*, and F63912, W36662, W60856 lack three quorum-sensing genes *lasR*, *rhlR*, and *rhlI*. Mutant of those genes including Δ*fliM*, Δ*pilS*, Δ*rhlR* and Δ*rhlI* displayed strong swarming-deficient phenotypes (Köhler *et al*, 2000; Kamatkar & Shrout, 2011; Overhage *et al*, 2007), except for *pilC* (Luo *et al*, 2015). The mutant of a two-component response regulator gene *pirR* that is missing in H27930 was found to repress swarming in multiple culture media (Kollaran *et al*, 2019). S86968 lacks *nfxB*, whose mutation caused global dysregulation of physiology and metabolism in *P. aeruginosa* including impaired swarming (Stickland *et al*, 2010).

Besides these functions, genes involved in carbon catabolism, biofilm, chemotaxis, antibiotic resistance, virulence and redox regulation were also found missing in these non-producers. M55212 misses the catabolite repression control gene *crc*, although the *crc* transposon mutant is able to swarm as well as the wild-type (Yeung *et al*, 2011). H27930 lacks the methyl-accepting chemotaxis genes *pctA* and *pctB* and the biofilm biosynthesis gene *pelA*. Both M1608 and S86968 lack the pyridoxal phosphate (vitamin B6) biosynthetic gene *pdxA* and the *hcnABC* genes coding for the virulence factors hydrogen cyanide (Pessi & Haas, 2000). F5677 cannot express PhzA2 for biosynthesis of phenazine—a well-known virulence factor, and also misses qscR—a quorum-sensing control repressor. The absence of *amrB* that encodes a membrane protein of the AmrAB-mediated efflux system in M1608 suggests that the strain may be sensitive to the aminoglycoside antibiotic (Westbrock-Wadman *et al*, 1999). The missing genes for redox regulation include soxR (absent in F5677)—a redox sensitive transcriptional regulator, *katE* (absent in M1608)—a catalase that degrades H2O2, and *gor* (absent in M1608)—a flavoprotein that oxidizes glutathione, which plays an important role in protection against H2O2 damage.

**Rhamnolipid producers grow faster in glycerol.** The missing genes in the 8 non-rhamnolipid-producers (Table 1) include transcriptional regulators (e.g., *rhlR*, *lasR*, *crc*, *nfxB*, *soxR*) known to elicit growth and global metabolic responses when absent. These global changes parallel the evolution of swarming and rhamnolipid production and thus fundamentally reflect the cellular conditions associated with the adaptive loss of the two phenotypes. We asked whether the bacterial growth rate could explain the two phenotypes. To address this question, we tracked the growth curves of all clinical isolates over 48h in the same glycerol media used to measure rhamnolipid production. The growth curves showed different patterns in the length of lag time and exponential growth-rates (Fig. 1C). Based on the Euclidean distance among the entire growth curves, most of the swarming strains (except for PAO1, F30658 and F23197) clustered together (Supplementary Fig. S3) and the strains in this cluster were also strong swarmers with high swarming scores. However, rhamnolipid producers clustered with non-producers, suggesting that rhamnolipid production cannot be classified by the entire growth curves and more sophisticated methods to extract local growth curve features might be needed.

We then used both unsupervised and supervised learning approaches to identify growth features associated with rhamnolipid production. For the unsupervised approach, we used a non-negative matrix decomposition method (Lee & Seung, 1999) which decomposes each growth curve as a weighted sum of three basis functions (i.e., features). Although the growth curves of the rhamnolipid producers (orange lines) and non-producers (blue lines) largely overlapped (Fig. 2A), there was a significant difference in the weights associated with basis 1 between the two groups (Fig. 2B), suggesting an association between rhamnolipid production and bacterial growth. To pinpoint features associated with rhamnolipid production, we divided each growth curve into three phases based on its shape (Fig. 2C and Supplementary Fig. S4) and defined 7 quantitative features to characterize each growth phase (Fig. 2D and Supplementary File 2). Using Random Forest classification, we found that the top two features with the highest explanatory power were the maximum and averaged specific growth rates in phase I when growth speeds up (Fig. 2E). Considering the majority of rhamnolipids are produced in phase II when growth slows down (Boyle *et al*, 2015), our finding revealed a strong temporal link between exponential growth rate in phase I and rhamnolipid production in phase II. We hypothesized that the link might be mediated by metabolic homeostasis: the disruption of homeostasis both slows down growth and disfavors overflow of excess carbon which would be preferentially used for cell maintenance and stress responses under metabolic imbalanced conditions.

**Rhamnolipid non-producers have perturbed TCA cycle and amino acid pathways.** Similar to our growth curve analysis, we adopted both unsupervised and supervised learning approaches to test the associations of cell metabolism with swarming and rhamnolipid production. In the same glycerol minimal medium, we collected intracellular metabolomics of all our strains (except for M55212 and F23197 which grew too slow) during the transition between phase I and phase II when rhamnolipid production begins. Using LC-MS we identified a total of 92 compounds (Fig. 3A) spanning a wide range of abundances. After data normalization and imputation (Supplementary Fig. S5), hierarchical clustering showed the consistency of data across all three replicates for each strain (Fig. 3B), except for one replicate of H47921 (sample 25, an outlier removed from further analysis). The clustering of the metabolomics data yielded three major clusters with one cluster containing all strong swarmers (cluster in the middle) without rhamnolipids non-producing strains, supporting the link between swarming and intracellular metabolism (Boyle *et al*, 2017). However, no clear metabolic pattern was found among rhamnolipids non-producing strains. Despite four out of the six non-producers clustered in the same subgroup, they all share some similar metabolomics pattern with producers from the same cluster.

The lack of a clear correlation between the metabolomic profile of *P. aeruginosa* strains and their ability to produce rhamnolipids led us to search for specific metabolites using supervised methods. We fitted our metabolomic data (explanatory variables) to categorized rhamnolipid production (response variable) using Orthogonal Projections to Latent Structures-Discriminant Analysis (OPLS-DA) (Materials and Methods). [#please check if my rephrase below makes sense#] The resulted model had a reasonable well fit (R2 = 0.82, Q2 = 0.66, *p-*value = 5e-4) (Fig. 4A) with the single predictor component that separated the group of rhamnolipid non-producers from the producers group explaining 5% of the variance across all data. The rest 95% of the data variance was explained by three orthogonal components that capture variation within both groups, suggesting that rhamnolipid producers and non-producer groups are highly heterogeneous in metabolomics. The domination of within-group variance over across-group variance also explains why it is difficult to identify discriminative compounds using unsupervised method. The loading values of the predictive component reflect how much each compound contributes to the separation between rhamnolipid producers and non-producers with negative loadings indicating higher metabolite abundance in non-producers (Fig. 4B and Supplementary Fig. 6). We identified 24 (#please confirm if this number is correct#) metabolites that have significantly different loading values between rhamnolipid producers and non-producers using a univariate Mann-Whitney U test (bars with black outline in Fig. 4B and Supplementary Fig. 6). These metabolites were then used to identify enriched metabolic pathways that differentiate the two groups through FELLA (Picart-Armada *et al*, 2018) (Materials and Methods, #should we have a supplementary file for FELLA output?#). The most discriminative pathways were the TCA cycle and amino acid metabolism (Fig. 4B and). Three out of six metabolites in the TCA cycle were significantly changed in the univariate test: fumarate, succinate and citrate. Fumarate and malate showed a positive loading value, while those of cis-aconitate, citrate, alpha-ketoglutarate and succinate were negative (Fig. 4B). Pyruvate remained relatively constant across all the strains, implying that the differential responses in the TCA cycle were independent from the changes in its upstream central carbon metabolism. Besides the TCA cycle metabolites, the majority of annotated compounds in the metabolism of branched chain amino acids (leucine/isoleucine, valine) and sulfur-containing amino acids (cysteine/methionine) had higher abundance in the rhamnolipid non-producers relative to the producers. A striking exception to this trend was N-Formylmethionine (fMet), which had lower abundance in non-producers. Diminished level of fMet was also found in the Δ*rhlA* mutant of *P. aeruginosa* (Supplementary Fig. 7)*.* Since that mutant had similar growth rate as the wild-type (Boyle *et al*, 2017), we conclude that the association between fMet and rhamnolipid production is not simply due to a growth defect.

**Strains that do not produce rhamnolipids have increased oxidative stress.** The above comparative analyses of growth curve and metabolomics suggest that the rhamnolipid non-producers might have higher oxidative stress levels possibly elicited by the following interlinked reasons. The TCA cycle harbors five enzymes with Fe-S clusters (aconitase A, aconitase B, succinate dehydrogenase subunit B, fumarase A, fumarase B (Py & Barras, 2010)) and thus represents one of the most vulnerable pathways subject to attacks of ROS. It is likely that higher oxidative stress in the non-producers reduces flux through the TCA cycle, which would explain their slower growth. The significantly opposite associations of succinate and fumarate we found in rhamnolipid production can also be explained by the reduced activity of succinate dehydrogenase (SDH) under oxidative stress conditions. SDH is a membrane-bound dehydrogenase linked to the respiratory chain—a major site of ROS production in the cell—and also a member of the TCA cycle that catalyzes the oxidation of succinate into fumarate (Hederstedt & Rutberg, 1981). Since SDH contains [2Fe-2S], [3Fe-4S], and [4Fe-4S] clusters (Ayala), oxidative stress that damages Fe-S clusters could decrease SDH activity *in vivo*. Some intermediate metabolites in amino acid biosynthetic pathways that are also substrates of Fe-S containing enzymes accumulated in the non-producers. For example, the [4Fe-4S] cluster-containing 3-isopropylmalate dehydratase subunit LeuC is a subunit of the isopropylmalate isomerase that catalyzes conversions from alpha-Isopropylmalate to 2-Isopropylmaleate and further to 3-isopropylmalate, where both alpha-Isopropylmalate and 2-Isopropylmaleate had significantly higher levels in the non-producers. Another example is the glutamate synthetases, which are composed of a large chain (encoded by *gltB*) and a small chain (encoded by *gltD*). Both subunits contain Fe-S clusters and catalyze the production of glutamate from 2-Oxoglutarate and glutamine. We found slightly higher 2-Oxoglutarate and glutamine as well as lower glutamine in the non-producers. Also, with the exception of glutamate, proline and tryptophan, all other amino acids detected in our data (threonine, alanine, aspartate, glutamine, cysteine, lysine, histidine, tyrosine, and arginine) were higher in non-producers. This observation suggests that non-producers may be attempting to recycle amino acids by extensive protein degradation against oxidative stresses.

These many reasons suggestthat rhamnolipids might not be secreted under high oxidative stress. To investigate this hypothesis we modified a high-quality genome-scale model of *Pseudomonas* metabolism, iJN1411 (Nogales *et al*, 2020, 2017), to simulate the maximum growth rate and rhamnolipids secretion potential under different redox stress levels (Materials and Methods). We simulated growth in the culture medium used in our metabolomics experiments with glycerol and ammonium as the sole carbon and nitrogen sources. Since rhamnolipids are produced when carbon is in excess (Boyle *et al*, 2015), we set the carbon-to-nitrogen (C:N) ratio to 10.0, which exceeds the minimum C:N threshold (6.3) that our model predicts to permit rhamnolipids secretion at fast growth rates (Supplementary Fig. 8). The redox stress levels were computationally varied by changing the flux levels of three redox molecules: NADH (reduced nicotinamide adenine dinucleotide), NADPH (reduced nicotinamide adenine dinucleotide phosphate) and GSH (reduced glutathione); these molecules are responsible for the bulk of cellular electron transfer and likely to be the main sources of ROS (Xiao & Loscalzo, 2019). For all three redox molecules, we found that the maximum growth rate is maintained at an intermediate flux range (redox homeostasis, white area) while any flux rate outside the range (gray shading) gradually reduced growth rate (Fig. 5, upper panels). Except for extremely small GSH flux, what generally accompanies the compromised growth rate is the abrupt shutdown in the potential of secretion of rhamnolipid precursors (HAA), mono- and di-rhamnolipids as well as many (but not all) central carbon metabolites (Fig. 5, lower panels). Importantly, none of these secretion fluxes reduce growth rate, nor do they provide growth benefits (i.e., each individual flux can be as low as zero without reducing maximum growth rate). This observation agrees with the fact that the *ΔrhlA* mutant of *P. aeruginosa* type strain PA14 displayed the same grow rate as its wild-type strain in the same glycerol minimum medium (van Ditmarsch & Xavier, 2011). Notably, most of the overflowed carbon-rich molecules, including HAA, mono- and di-rhamnolipids, are only secreted when growth rate maintains at its maximum value under carbon excessive condition. The finding that molecules other than rhamnolipids can be secreted under maximum growth suggests that *P. aeruginosa* cells have flexibility to overflow carbon through alternative metabolic pathways. Moreover, the fact that *P. aeruginosa* maintains the secretion of rhamnolipids among these options supports that this mechanism may be selected under pressure other than single-cell fitness. Collectively, these simulations support a link between rhamnolipids secretion and redox stress, where fast growth and rhamnolipids secretion are both metabolically constrained by cellular redox homeostasis.

**Experiments support the link between growth, oxidative stress and rhamnolipid production.** We further validated the link between oxidative stress and rhamnolipid production by two additional experiments. In the first experiment,we monitored growth and rhamnolipid production of all our clinical isolates when they grew in the same minimal medium but with succinate as the sole carbon source (Supplementary Fig. S9). Since succinate enters the carbon metabolism directly through the TCA cycle, we expected that rhamnolipids might not be secreted due to higher ROS level generated by TCA cycle, which represses rhamnolipid production. Indeed, none of the strains that produced rhamnolipids in glycerol secreted rhamnolipids in succinate. We similarly extracted 7 phase-dependent growth curve features of our strains grown in succinate (Supplementary File 3), showing that the strainshad faster exponential phase (phase I) growth rate and shorter delay before phase I commences. However, the strains could hardly maintain the high growth rate and reached much lower cell density compared to their growth in glycerol (Supplementary Fig. S10), implying fast accumulation of oxidative stress. More interestingly, the duration of phase II for these strains was much shorter in succinate than in glycerol (Supplementary Fig. S10), suggesting a stronger stress response that would shut down growth as well as carbon overflow, and force quick entry into stationary phase for cell maintenance.

In the second experiment, we simultaneously monitored cell density and hydrogen peroxide (H2O2) level (Materials and Methods) in the glycerol minimal medium during the growth of our isolates (Fig. 6A,B). H2O2 is a representative ROS that can diffuse freely between cell and the environment. H2O2 was detected in the absence of cells, indicating that it can be produced by oxidization of the culture medium. We computed the H2O2 level difference between w/o and w/ cells and showed that all strains except for M1608 can degrade H2O2 over the first 18 hours (Fig. 6B). Most importantly, the non-producers (blue lines) degraded less H2O2 compared to the mild (green lines) and strong producers (red lines), supporting our hypothesis that non-producers are more susceptible to oxidative stress. Unsurprisingly, the worst H2O2 degrader M1608 lacked *katE* and *gor* (Table 1)*,* both of which are important for H2O2 degradation. To factor out the possibility that the non-producers degraded less H2O2 due to lower cell density, we calculated the H2O2 removal rate per cell and observed similar relative trend of H2O2 degradation between non-producers and producers (Fig. 6C). Formally, we used a linear mixed-effect model to test the effect of rhamnolipid production on per-cell H2O2 removal rate (Materials and Methods). The effect size of mild and strong producers relative to the non-produces are 1.25 and 1.41 respectively with both p-values < 0.001, confirming again that the rhamnolipid producers can detoxify H2O2 better.

**Discussion**

Taken together, our data leads to metabolic features that explain the swarming diversity in clinical isolates (Fig. 7): The rhamnolipid producers typically grow fast and maintain redox homeostasis due to their greater abilities to degrade ROS; by contrast, the non-producers grow slowly with reduced TCA cycle activity due to increased oxidative stress. By comparing growth curve and metabolomes of our clinical strains, we revealed that rhamnolipid non-producers generally have perturbed TCA cycle and amino acid metabolism. We demonstrate that these metabolic perturbations coincide with disrupted redox homeostasis, which is possibly caused by loss of genes that encode ROS detoxification functions (e.g., M1608 lacks *katE* and *gor*). All our clinical strains were isolated from hospitalized patients, a host-associated environment where we may expect redox stresses such as ROS imposed by the immune system in their fight against pathogens (Puertollano *et al*, 2011). Non-producers of rhamnolipids seem less capable of dealing with such oxidative stresses, and it remains unclear what selected for this loss of function *in vivo*. Rhamnolipid production is, nonetheless, relatively frequent across the phylogenetic tree. The genes for rhamnolipid biosynthesis (*rhlA*, *rhlB*, *rhlC*) remain conserved across all of our clinical isolations even in non-producers, suggesting that the loss of rhamnolipid production results from broader metabolic adaptations such as oxidative stress responses.

We found that rhamnolipid producers could remove environmental H2O2 faster than non-producers. Since H2O2 exposure significantly decreased expression of the key quorum-sensing genes (*LasI*, *LasR*, *RhlI*, *RhlR*, *PqsA* and PqsR) (Mohamed *et al*, 2020), we hypothesized that, for the non-producers with intact quorum-sensing genes, their expressions might be inhibited in response to oxidative stress, which consequently abolishes rhamnolipid production and swarming (Fig. 7). Considering quorum sensing regulates hundreds of *P. aeruginosa* genes(Lequette *et al*, 2006), shutting down quorum-sensing control might be a mechanism to save energy and divert precious resources to cell maintenance and oxidative stress responses. It was reported that Lon protease can repress both LasR/LasI and RhlR/RhlI quorum sensing systems (Takaya *et al*, 2008); we suspect that oxidative stress inhibits quorum-sensing responses through Lon protease, further abolishing rhamnolipid production.

Rhamnolipid production resembles overflow metabolism, a mechanism generally associated with fast growth of bacteria under conditions of high glucose consumption. Overflow metabolism occurs when cells simultaneously operate on both energy-efficient (e.g., respiration) and -inefficient (e.g., fermentation) pathways and secret metabolic byproducts that could otherwise be used for catabolism or anabolism. This phenomenon has been observed in different cell types, including aerobic fermentation in *Escherichia coli* (Farmer & Jones, 1976), the Crabtree effect in *Saccharomyces cerevisiae* (De Deken, 1966) and the Warburg effect in cancer cells (Vander Heiden *et al*, 2009). Many molecular mechanisms have been proposed and validated to account for the seemingly wasteful usage of low energy-yield pathway in presence of a higher-energy-yield alternative (Basan *et al*, 2015; Szenk *et al*, 2017). For example, Szenk *et al.* proposed that pure respiratory flux operating at levels demanded by fast growth without fermentation would be toxic due to accumulated NADH that cannot be recycled to NAD+ by respiration under surface limitation (Szenk *et al*, 2017). It was further shown that overexpression of NADH oxidase in both *E. coli* (Vemuri *et al*, 2006) and *S. cerevisiae* (Vemuri *et al*, 2007) can reduce the NADH/NAD+ ratio and suppress the overflow metabolism, which agreed with our findings that rhamnolipid production is redox-dependent. One apparent difference between *P. aeruginosa* rhamnolipid production and overflow metabolism observed in other cell types is that rhamnolipids are controlled by quorum-sensing in addition to nutrient cues, which result in a cell-density-dependence (Boyle *et al*, 2015).

Should rhamnolipid production provide a fitness benefit for *P. aeruginosa* under any environmental conditions? We have previously shown that rhamnolipid production is dispensable for normal growth of *P. aeruginosa* under laboratory condition: the Δ*rhlA* mutant (rhamnolipids defector) had almost the same growth rate (van Ditmarsch & Xavier, 2011) and very similar metabolomic profiles (Boyle *et al*, 2017), suggesting that the carbon used for rhamnolipid production does not interfere with biomass production. This is consistent with our flux-balance analysis of Pseudomonas metabolism where shutting down rhamnolipid production does not affect growth and excess carbon can be alternatively secreted in form of other metabolites such as acetate (Fig. 5). Notably, Δ*rhlA* had higher gamma-Glutamylcysteine (Supplementary Fig. S7), the immediate precursor of glutathione. Since glutathione is a well-known antioxidant that protects cell from ROS damage (Ezraty *et al*, 2017), Δ*rhlA* may exhibit slight oxidative burden, which would explain why the cheater has a marginally lower fitness and failed to invade the wild-type in swarming competitions (de Vargas Roditi *et al*, 2013; Monaco *et al*, 2020). Our mechanism proposes that increased oxidative burden may be related to altered redox ratio (NAD(P)H/NAD(P)+), considering that fatty acid biosynthesis that provides precursors for rhamnolipid production regenerates NAD(P)+ from NAD(P)H. Notably, the active shutdown of rhamnolipid production in redox imbalanced condition ensures that precious carbon and enzyme resources were not wasted but diverted to oxidative stress response mechanisms.

This study sheds light on how metabolite secretion can impact both single-cell physiology and social behavior. Microorganisms secrete many different kinds of metabolic byproducts, including extracellular enzymes, toxins and cell-cell signaling molecules (Schmidt *et al*, 2019) and the byproducts of one microorganism can impact another to drive interaction. Our model proposes that the use of a given molecule in social interactions is likely not arbitrary: molecules that conciliate individual-level and population-level interests are more likely to drive microbial social interactions.

**Materials and Methods**

lysogen broth () at at 250 rpm

(Xavier *et al*, 2011; van Ditmarsch *et al*, 2013) swarmingSwarming medium contains 0.5% agar (Bacto) and is supplemented with 5g/L casamino acids, 1mM MgSO4, 0.1mM CaCl2 and 1X salt buffer (12g/L Na2HPO4 (Fisher Scientific), 15g/L KH2PO4 (Fisher Scientific) and 2.5g/L NaCl, pH6.7).

**Rhamnolipid production determination.** The production of rhamnolipids was assessed by drop-collapse assay (Jain & Collins; Chen *et al*, 2007). We placed 50 μL of the culture’s supernatant on a polystyrene surface (the lid of a 96 well plate). The presence of rhamnolipids decreases the surface tension of the liquid, making the drop collapse.

**Image analysis and determination of swarming score.** Images of the swarming plates were obtained with a Chemidoc gel doc imager (Bio Rad). The extraction of the morphological features from the images of the swarming plates were analyzed using Matlab bwmorph function. The features extracted from the images were perimeter of the colony, maximum length (the longitude of the rectangle that fits the colony), area percentage of the plate occupied by the colony, circularity (measured as 4\*π/P2), skeleton (#definition of skeleton#) and eccentricity (the eccentricity of the fitted ellipse). The feature values of each strain were normalized by the corresponding values of the PA14 strain. The swarming score () for each clinical isolate is defined as a linear combination of the maximum length () and circularity () of its swarming colony (the two most discriminative features that separate strong swarmers from weak swarmers)

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

where

**Cross-strain reciprocal BLAST.** To identify orthologous gene across all *P. aeruginosa* strains, systematic cross-strain reciprocal BLAST was performed. Protein sequences of each strain and the type strain PA14 were mutually queried against and the BLAST hits (genes) with <80% coverage were removed. Two genes are orthologous if they were mutually the best hit to each other and genes (from other strains) that are orthologous to the same PA14 gene were considered orthologous to each other.

**Phylogenetic analysis.** The phylogenetic tree of clinical *P. aeruginosa* strains (Fig. 1A) was constructed from core genomes as previously described (Yan *et al*, 2017). Moran’s I test was carried out using the ape package in R (Paradis & Schliep, 2019). The ancestral state of swarming and rhamnolipid production was reconstructed using corHMM package (Beaulieu *et al*, 2013).

**Metabolite extraction.** All *P. aeruginosa* strains were grown until OD600 = 0.2 (end of exponential phase of growth) in glycerol minimal media. Bacteria was then loaded into 0.25 μm nylon membranes (Millipore) using vacuum, transferred to pre-warmed hard agar plates with the same medium composition and incubated at 37ºC during 2.5 h. The filters were then passed to 35 mm polystyrene dishes (Falcon) with 1 mL of 2:2:1 methanol:acetonitrile:H2O quenching buffer and incubated there during 15 minutes on dry ice. Cells were removed by scraping and the lysate containing quenching buffer was transferred to 1.5 mL tubes and centrifuged at 16000 rpm for 10 minutes at 4ºC. Supernatant transferred to fresh tubes and stored at -80ºC.

**Metabolomic data preprocessing.** The extracts were profiled using liquid-chromatography coupled to mass spectrometry (LC-MS), identifying a total of 92 compounds (Fig. 3A). Some compounds contained missing values.These missing values in metabolite abundance can be (1) truly missing; (2) present in a sample but its level is below detection limit; (3) present in a sample at a level above the detection limit but missing due to failure of algorithms in data processing. Here we assume that a metabolite with missing values in all three replicates is truly missing in the sample and removed from our analysis (Supplementary Fig. 5). However, if the missing values were only found in one or two replicates, the missing values were imputed by the average of the non-missing values. After that imputation all compounds with missing values were removed (Fig. S4).

The peak areas were normalized using Cross-Contribution Compensating Multiple Standard Normalization (CCMN) (Redestig *et al*, 2009) with NormalizeMets R package (De Livera *et al*, 2018). This method relies on the use of multiple internal standards, but as the LC-MS was done without using them we used instead a set of metabolites assumed to be constant across all the strains. They were selected with a Kuskal-Wallis test, adjusting the *p-*value with Benjamini-Hochberg method. The ones with a *p-*value above 0.05 were considered constant (pyruvate, methylglyoxal, (S)-2-Acetolactate, Tyramine, D-Glucose, (S)-Lactate, N-acetyl-L-glutamate 5-semialdehyde, 4-Aminobutyraldehyde and Glycine), therefore after the normalization step they were removed (in black in Fig. 3A).

**Hierarchical clustering of metabolomic data.** The Hierarchical Clustering Analysis of the normalized metabolomic data was performed using gplots R package (Warnes *et al.* 2009), with Euclidean distance and Ward’s aggregation method. The clustering was done with all the metabolites from preprocessed metabolomic table, despite the experimentalists were not confident about the identity of 16 compounds (indicated in red in Figure 2A): we assumed that if the peaks appeared in all the strains they could not be artifacts of the LC-MS. These compounds are not shown in the heatmap of Figure 3B, and were not used in downstream analyses. Fumarate and Guanosine were initially categorized as ambiguous by the experimentalists, but we rescued them as our clinical isolates only had enzymatic genes related to them among all the possible compounds with the same molecular weight, according to KEGG database.

**Metabolic pathway enrichment.** The differential metabolites between rhamnolipid producers and non producers was determined by a Mann-Whitney test, with *p*-values adjusted with Benjamini-Hochberg method and a significance level of 0.05. These compounds were fed to FELLA algorithm (Picart). FELLA retrieves a graph consisting all the entries in KEGG database for, in our case, *Pseudomonas aeruginosa* strain UCBPP-PA14, and applies over it diffusion algorithms, using as input for the network the differential compounds (Vandin *et al* 2011). The output of FELLA consists in all the subnetwork of the entries predicted to have a high connectivity with the differential compounds. We filtered the entries of this subnetwork to keep only the metabolic pathways.

**OPLS-DA model.** OPLS-DA model of metabolomics data was built using ropls R package (Thévenot *et al*, 2015), fixing the number of orthogonal components to 3. R2 and Q2, key parameters for assessing the validity of the model, were assessed with 7-fold cross validation. The significance of the model was determined by permutation test (n = 2000). The *p*-value corresponds to the proportion of Q2perm above Q2. With a *p-*value below 0.05 we considered the model significant. The loadings of the predictive component of the model were extracted to determine how each metabolite contributes to the separation according to the phenotype.

**Genome-scale modeling.** Custom Python codes were developed with the COBRApy package (Ebrahim *et al*, 2013) to carry out all metabolic flux modeling and simulations in the paper. Since iJN1411 model was developed for *Pseudomonas putida*, we removed genes and associated reactions that are missing in all our strains but present in the iJN1411 model. The futile cycles involving NADH, NADPH, and GSH were also removed. The modified iJN1411 model was further expanded by adding rhamnolipid biosynthesis pathway involving 9 new metabolites and 12 new reactions. The metabolites, reactions and gene included in the final model are shown in Supplementary File 3.

The boundary fluxes of the model were set to mimic the composition of the glycerol minimum medium. For C:N=10, the lower bounds of glycerol and ammonium fluxes were set to -10 and -3 respectively. For C:N=3, the lower bounds were set to -3 and -3 respectively. The flux unit is mmol/gDW/h throughout the paper. To constrain the total producing flux of NADH (the same for NADPH and GSH) at a certain value , we first defined a binary variable for each NADH-involving reaction to indicate whether NADH is produced by this reaction. Given the stoichiometric coefficient of NADH in this reaction () and its flux value (), the mathematical constraints for was set by for and otherwise. Therefore, the constraint that equalizes the total NADH producing flux and a constant is simply . However, both and are variables and such quadratic constraint has not yet been supported by COBRApy. We overcame this difficulty by defining and linearized the product with the following two inequalities: and , where and are the lower and upper bounds of . The two constraints ensures that when and when . The minimum/maximum flux values of byproduct secretion was simulated by flux variability analysis at maximum growth rate.

**H2O2 measurement**. The H2O2 level in the extracellular medium was quantified with Amplex® Red Hydrogen Peroxide/Peroxidase Assay Kit (Invitrogen, Carlsbad, USA, Catalog no. A22188). 500 μL of #what# was spinned down after overnight growth in LB medium, washed twice in PBS and normalized to OD 1. Each reaction was done in a final volume of 100 μL, with a final concentration of 50 μM of Amplex® Red reagent, 0.1 U/mL of HRP (Horseradish Peroxidase) and 0.2 cell OD, in glycerol synthetic medium, in BD Falcon (BD Biosciences, San Jose, CA) 96 well flat-bottom plates. The first column of the plate corresponded to the reaction without cells (the volume was substituted by PBS), and the last column instead of cells contained H2O2 (final concentration 10 μM). OD600 and fluorescence 530/590 nm was measured in 10 minutes intervals (48 h 37ºC). The H2O2 removal was determined by subtracting the values of emission from each strain to the values of emission of the wells without cells neither H2O2: H2O2 is produced in the medium; as the cells can detoxify it the difference between the values will indicate the amount removed. The cells that suffer from oxidative stress will add H2O2 to the medium, thus being the difference negative. We used OD600 values for obtaining H2O2 removal per cell.

**References**

Abdel-Mawgoud AM, Lépine F & Déziel E (2010) Rhamnolipids: diversity of structures, microbial origins and roles. *Appl. Microbiol. Biotechnol.* **86:** 1323–1336

Ayala C Fe-S cluster assembly pathways in bacteria. *Castro*

Basan M, Hui S, Okano H, Zhang Z, Shen Y, Williamson JR & Hwa T (2015) Overflow metabolism in Escherichia coli results from efficient proteome allocation. *Nature* **528:** 99–104

Beaulieu JM, O’Meara BC & Donoghue MJ (2013) Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. *Syst. Biol.* **62:** 725–737

Biro PA & Stamps JA (2010) Do consistent individual differences in metabolic rate promote consistent individual differences in behavior? *Trends Ecol. Evol. (Amst.)* **25:** 653–659

Boyle KE, Monaco H, van Ditmarsch D, Deforet M & Xavier JB (2015) Integration of metabolic and quorum sensing signals governing the decision to cooperate in a bacterial social trait. *PLoS Comput. Biol.* **11:** e1004279

Boyle KE, Monaco HT, Deforet M, Yan J, Wang Z, Rhee K & Xavier JB (2017) Metabolism and the evolution of social behavior. *Mol. Biol. Evol.* **34:** 2367–2379

Caiazza NC, Shanks RMQ & O’Toole GA (2005) Rhamnolipids modulate swarming motility patterns of *Pseudomonas aeruginosa.* *J. Bacteriol.* **187:** 7351–7361

Chen CY, Baker SC & Darton RC (2007) The application of a high throughput analysis method for the screening of potential biosurfactants from natural sources. *J. Microbiol. Methods*

Chong H & Li Q (2017) Microbial production of rhamnolipids: opportunities, challenges and strategies. *Microb. Cell Fact.* **16:** 137

De Deken R (1966) The Crabtree Effect: A Regulatory System in Yeast. *J. Gen. Microbiol.* **44:** 149–156

De Livera AM, Olshansky G, Simpson JA & Creek DJ (2018) NormalizeMets: assessing, selecting and implementing statistical methods for normalizing metabolomics data. *Metabolomics* **14:** 54

Deforet M, van Ditmarsch D, Carmona-Fontaine C & Xavier JB (2014) Hyperswarming adaptations in a bacterium improve collective motility without enhancing single cell motility. *Soft Matter* **10:** 2405–2413

Déziel E, Lépine F, Milot S & Villemur R (2003) *rhlA* is required for the production of a novel biosurfactant promoting swarming motility in *Pseudomonas aeruginosa*: 3-(3-hydroxyalkanoyloxy)alkanoic acids (HAAs), the precursors of rhamnolipids. *Microbiology (Reading, Engl.)* **149:** 2005–2013

van Ditmarsch D, Boyle KE, Sakhtah H, Oyler JE, Nadell CD, Déziel É, Dietrich LEP & Xavier JB (2013) Convergent evolution of hyperswarming leads to impaired biofilm formation in pathogenic bacteria. *Cell Rep.* **4:** 697–708

van Ditmarsch D & Xavier JB (2011) High-resolution time series of *Pseudomonas aeruginosa* gene expression and rhamnolipid secretion through growth curve synchronization. *BMC Microbiol.* **11:** 140

Ebrahim A, Lerman JA, Palsson BO & Hyduke DR (2013) COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Syst. Biol.* **7:** 74

Ezraty B, Gennaris A, Barras F & Collet J-F (2017) Oxidative stress, protein damage and repair in bacteria. *Nat. Rev. Microbiol.* **15:** 385–396

Farmer IS & Jones CW (1976) The energetics of Escherichia coli during aerobic growth in continuous culture. *Eur. J. Biochem.* **67:** 115–122

Hederstedt L & Rutberg L (1981) Succinate dehydrogenase--a comparative review. *Microbiol Rev* **45:** 542–555

Inoue T, Shingaki R, Hirose S, Waki K, Mori H & Fukui K (2007) Genome-wide screening of genes required for swarming motility in Escherichia coli K-12. *J. Bacteriol.* **189:** 950–957

Jain DK & Collins DL A drop-collapsing test for screening surfactant-producing microorganisms. *Thompson*

Kamatkar NG & Shrout JD (2011) Surface hardness impairment of quorum sensing and swarming for Pseudomonas aeruginosa. *PLoS One* **6:** e20888

Kearns DB (2010) A field guide to bacterial swarming motility. *Nat. Rev. Microbiol.* **8:** 634–644

Kim W & Surette MG (2004) Metabolic differentiation in actively swarming Salmonella. *Mol. Microbiol.* **54:** 702–714

Klevens RM, Edwards JR, Richards CL, Horan TC, Gaynes RP, Pollock DA & Cardo DM (2007) Estimating health care-associated infections and deaths in U.S. hospitals, 2002. *Public Health Rep.* **122:** 160–166

Köhler T, Curty LK, Barja F, van Delden C & Pechère JC (2000) Swarming of *Pseudomonas aeruginosa* is dependent on cell-to-cell signaling and requires flagella and pili. *J. Bacteriol.* **182:** 5990–5996

Kollaran AM, Joge S, Kotian HS, Badal D, Prakash D, Mishra A, Varma M & Singh V (2019) Context-Specific Requirement of Forty-Four Two-Component Loci in Pseudomonas aeruginosa Swarming. *iScience* **13:** 305–317

Lee DD & Seung HS (1999) Learning the parts of objects by non-negative matrix factorization. *Nature* **401:** 788–791

Lequette Y, Lee J-H, Ledgham F, Lazdunski A & Greenberg EP (2006) A distinct QscR regulon in the Pseudomonas aeruginosa quorum-sensing circuit. *J. Bacteriol.* **188:** 3365–3370

Luo Y, Zhao K, Baker AE, Kuchma SL, Coggan KA, Wolfgang MC, Wong GC & O’Toole GA (2015) A hierarchical cascade of second messengers regulates *Pseudomonas aeruginosa* surface behaviors. *MBio* **6:**

Mattingly AE, Kamatkar NG, Borlee BR & Shrout JD (2018) Multiple Environmental Factors Influence the Importance of the Phosphodiesterase DipA upon *Pseudomonas aeruginosa* Swarming. *Appl. Environ. Microbiol.* **84:**

Medina G, Juárez K, Valderrama B & Soberón-Chávez G (2003) Mechanism of Pseudomonas aeruginosa RhlR transcriptional regulation of the rhlAB promoter. *J. Bacteriol.* **185:** 5976–5983

Mellbye B & Schuster M (2014) Physiological framework for the regulation of quorum sensing-dependent public goods in Pseudomonas aeruginosa. *J. Bacteriol.* **196:** 1155–1164

Mohamed FA, Shaker GH & Askoura MM (2020) Oxidative Stress Influences Pseudomonas aeruginosa Susceptibility to Antibiotics and Reduces Its Pathogenesis in Host. *Curr Microbiol*

Monaco H, Sereno T, Liu K, Reagor C, Deforet M & Xavier JB (2020) Spatial-temporal dynamics of a microbial cooperative behavior robust to cheating. *BioRxiv*

Nogales J, Gudmundsson S, Duque E, Ramos JL & Palsson BO (2017) Expanding The Computable Reactome In *Pseudomonas putida* Reveals Metabolic Cycles Providing Robustness. *BioRxiv*

Nogales J, Mueller J, Gudmundsson S, Canalejo FJ, Duque E, Monk J, Feist AM, Ramos JL, Niu W & Palsson BO (2020) High-quality genome-scale metabolic modelling of Pseudomonas putida highlights its broad metabolic capabilities. *Environ. Microbiol.* **22:** 255–269

Overhage J, Lewenza S, Marr AK & Hancock REW (2007) Identification of genes involved in swarming motility using a *Pseudomonas aeruginosa* PAO1 mini-Tn5-lux mutant library. *J. Bacteriol.* **189:** 2164–2169

Paradis E & Schliep K (2019) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35:** 526–528

Pessi G & Haas D (2000) Transcriptional control of the hydrogen cyanide biosynthetic genes hcnABC by the anaerobic regulator ANR and the quorum-sensing regulators LasR and RhlR in Pseudomonas aeruginosa. *J. Bacteriol.* **182:** 6940–6949

Picart S FELLA: an R package to enrich metabolomics data. *Armada*

Picart-Armada S, Fernández-Albert F, Vinaixa M, Yanes O & Perera-Lluna A (2018) FELLA: an R package to enrich metabolomics data. *BMC Bioinformatics* **19:** 538

Puertollano MA, Puertollano E, de Cienfuegos GÁ & de Pablo MA (2011) Dietary antioxidants: immunity and host defense. *Curr. Top. Med. Chem.* **11:** 1752–1766

Py B & Barras F (2010) Building Fe-S proteins: bacterial strategies. *Nat. Rev. Microbiol.* **8:** 436–446

Redestig H, Fukushima A, Stenlund H, Moritz T, Arita M, Saito K & Kusano M (2009) Compensation for systematic cross-contribution improves normalization of mass spectrometry based metabolomics data. *Anal. Chem.* **81:** 7974–7980

Schmidt R, Ulanova D, Wick LY, Bode HB & Garbeva P (2019) Microbe-driven chemical ecology: past, present and future. *ISME J.* **13:** 2656–2663

Smith E & Morowitz HJ (2004) Universality in intermediary metabolism. *Proc. Natl. Acad. Sci. USA* **101:** 13168–13173

Stickland HG, Davenport PW, Lilley KS, Griffin JL & Welch M (2010) Mutation of nfxB causes global changes in the physiology and metabolism of Pseudomonas aeruginosa. *J. Proteome Res.* **9:** 2957–2967

Szenk M, Dill KA & de Graff AMR (2017) Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the Membrane Real Estate Hypothesis. *Cell Syst.* **5:** 95–104

Takaya A, Tabuchi F, Tsuchiya H, Isogai E & Yamamoto T (2008) Negative regulation of quorum-sensing systems in Pseudomonas aeruginosa by ATP-dependent Lon protease. *J. Bacteriol.* **190:** 4181–4188

Thévenot EA, Roux A, Xu Y, Ezan E & Junot C (2015) Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *J. Proteome Res.* **14:** 3322–3335

Tremblay J & Déziel E (2010) Gene expression in Pseudomonas aeruginosa swarming motility. *BMC Genomics* **11:** 587

Vander Heiden MG, Cantley LC & Thompson CB (2009) Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324:** 1029–1033

de Vargas Roditi L, Boyle KE & Xavier JB (2013) Multilevel selection analysis of a microbial social trait. *Mol. Syst. Biol.* **9:** 684

Vemuri GN, Altman E, Sangurdekar DP, Khodursky AB & Eiteman MA (2006) Overflow metabolism in Escherichia coli during steady-state growth: transcriptional regulation and effect of the redox ratio. *Appl. Environ. Microbiol.* **72:** 3653–3661

Vemuri GN, Eiteman MA, McEwen JE, Olsson L & Nielsen J (2007) Increasing NADH oxidation reduces overflow metabolism in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. USA* **104:** 2402–2407

Westbrock-Wadman S, Sherman DR, Hickey MJ, Coulter SN, Zhu YQ, Warrener P, Nguyen LY, Shawar RM, Folger KR & Stover CK (1999) Characterization of a Pseudomonas aeruginosa efflux pump contributing to aminoglycoside impermeability. *Antimicrob. Agents Chemother.* **43:** 2975–2983

Xavier JB, Kim W & Foster KR (2011) A molecular mechanism that stabilizes cooperative secretions in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **79:** 166–179

Xiao W & Loscalzo J (2019) Metabolic responses to reductive stress. *Antioxid. Redox Signal.*

Yan J, Deforet M, Boyle KE, Rahman R, Liang R, Okegbe C, Dietrich LEP, Qiu W & Xavier JB (2017) Bow-tie signaling in c-di-GMP: Machine learning in a simple biochemical network. *PLoS Comput. Biol.* **13:** e1005677

Yan J, Estanbouli H, Liao C, Kim W, Monk JM, Rahman R, Kamboj M, Palsson BO, Qiu W & Xavier JB (2019) Systems-level analysis of NalD mutation, a recurrent driver of rapid drug resistance in acute Pseudomonas aeruginosa infection. *PLoS Comput. Biol.* **15:** e1007562

Yeung ATY, Bains M & Hancock REW (2011) The sensor kinase CbrA is a global regulator that modulates metabolism, virulence, and antibiotic resistance in *Pseudomonas aeruginosa.* *J. Bacteriol.* **193:** 918–931

Zhu K & Rock CO (2008) RhlA converts beta-hydroxyacyl-acyl carrier protein intermediates in fatty acid synthesis to the beta-hydroxydecanoyl-beta-hydroxydecanoate component of rhamnolipids in *Pseudomonas aeruginosa.* *J. Bacteriol.* **190:** 3147–3154

|  |  |
| --- | --- |
| **Non-producers** | **Missing genes** |
| **F5677** | *arsB, arsC, arsH, arsR, chiC, dacB, fliM, gcd, qscR, gntK, gntR, ilvA1, kdsB, pbpC, pchA, phzA2, soxR* |
| **F63912** | *lasR* |
| **H27930** | *arnA, pelA, pcaB, pcaD, pfeA, icmF3, rrsmE, pcd, pirR, pctA, pctB, osmE* |
| **M1608** | *tauD, opdG, pdxA, hcnC, hcnB, hcnA, glgA, ybhO, glgB, katE, glgP, ligD, ada, opdO, kynB, fusA2, cynR, cynT, cynS, cmrA, gor, galU, amrB, gnyD, gnyB, gnyH, gnyA, gnyL, hmgA, fahA, maiA, bdhA, atoB, scoB, ppiC1, pqqD, pqqC, pqqB, exaC, exaB, exaA, pqqF, braZ, rbsK, rbsR, rbsC, rbsA, rbsB, pcd* |
| **M55212** | *crc, rph* |
| **S86968** | *kdsB, cysP, hcnC, hcnB, hcnA, pdxA, opdG, nfxB, cupB3* |
| **W36662** | *rhlR, rhlI, cupC1, pilC, pilS* |
| **W60856** | *lasR, pirA* |

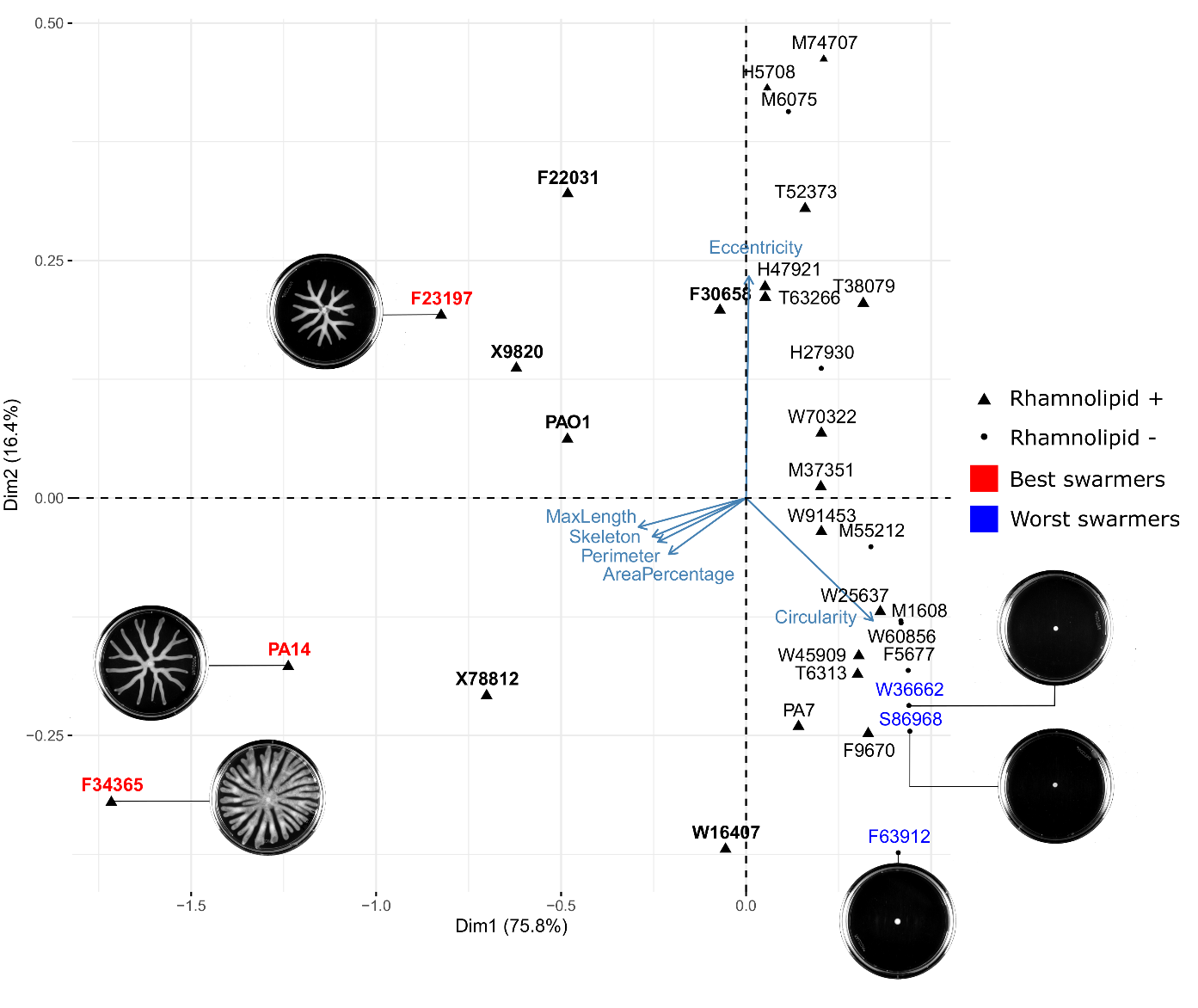
Table 1 Genes that are missing in non-rhamnolipid-producers but present in all producers

**Figures**

A picture containing screenshot

Description automatically generated

**Figure 1.** Diversity of swarming across the *P. aeruginosa* phylogeny of core genomes. **A.** Phylogeny of clinical isolates obtained from patients with cancer at MSKCC (Yan *et al*, 2017, 2019) together with reference strains PAO1, PA14 and PA7. The tissue where each isolate was originally isolated is labeled by circle colors. **B**. Swarming and rhamnolipid production phenotypes. The swarming ability of these strains is represented by the actual swarming images on the agar plates (left column). The ability of producing rhamnolipids of these strains is indicated by circle sizes (right column). Both phenotypes do not correlate with phylogeny. All swarmers can secrete rhamnolipids but some rhamnolipid producers cannot swarm. **C**. Growth curves in a synthetic minimal medium using glycerol as the sole carbon source..



**Supplementary Figure 1.** Principle component analysis and biplot of the swarming phenotype. Six morphological features were extracted from quantitative imaging analysis (see Methods in the main text), including maximum length (MaxLength), skeleton, perimeter, area of swarming colony (AreaPercentage), circularity and eccentricity. The MaxLength and circularity were the top two features that separate swarmers (bold font-weight) from non-swarmers (normal font-weight). The top 3 strains that have visually the largest (best swarmers) and smallest (worst swarmers) swarming colonies are highlighted in red and blue respectively with their swarming images shown alongside. The abilities of these strains to produce rhamnolipids are also indicated by triangles (producers) and dots (non-producers).

A screenshot of a cell phone

Description automatically generated

**Supplementary Figure 2.** Phylogenetic ancestor state reconstruction of rhamnolipid production (A) and swarming phenotypes (B). For both (A) and (B), pie charts at the ancestor nodes of branches represent relative likelihood proportion of each possible phenotypic state. We rooted the tree with PA7, a *P. aeruginosa* isolate that is often used as an outlier to root phylogenetic trees.

A screenshot of a cell phone

Description automatically generated

**Supplementary Figure 3.** Clustergram of growth curves of *P. aeruginosa* clinical isolates and three type strains PA14, PAO1 and PA7 in glycerol minimal medium. Euclidean distance was used as the measure of similarity. The strains that are able to swarm are indicated in bold. Most of the swarmers are clustered together. The swarmers that do not fall into this cluster include F30658 and PAO1, both of which are mild swarmers, as well as F23197 whose growth curve has a relatively longer lag time.

A picture containing shirt

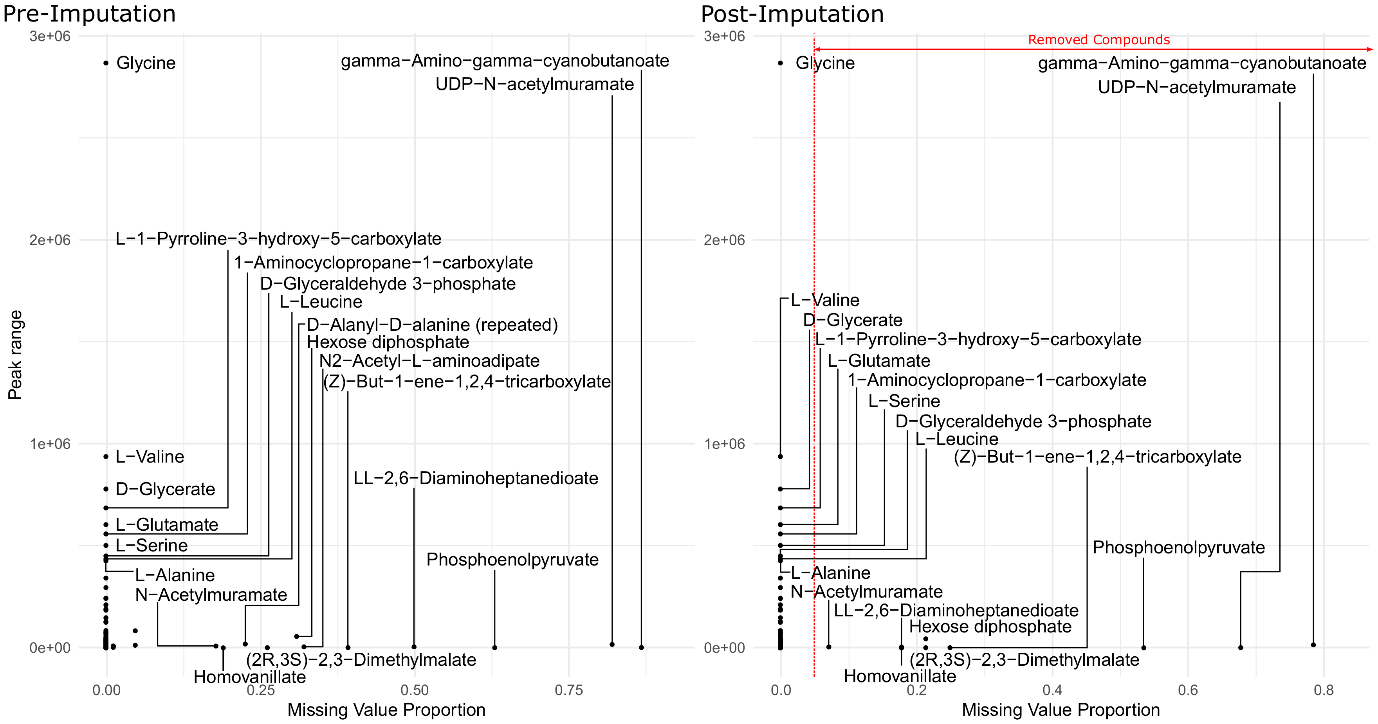
Description automatically generated

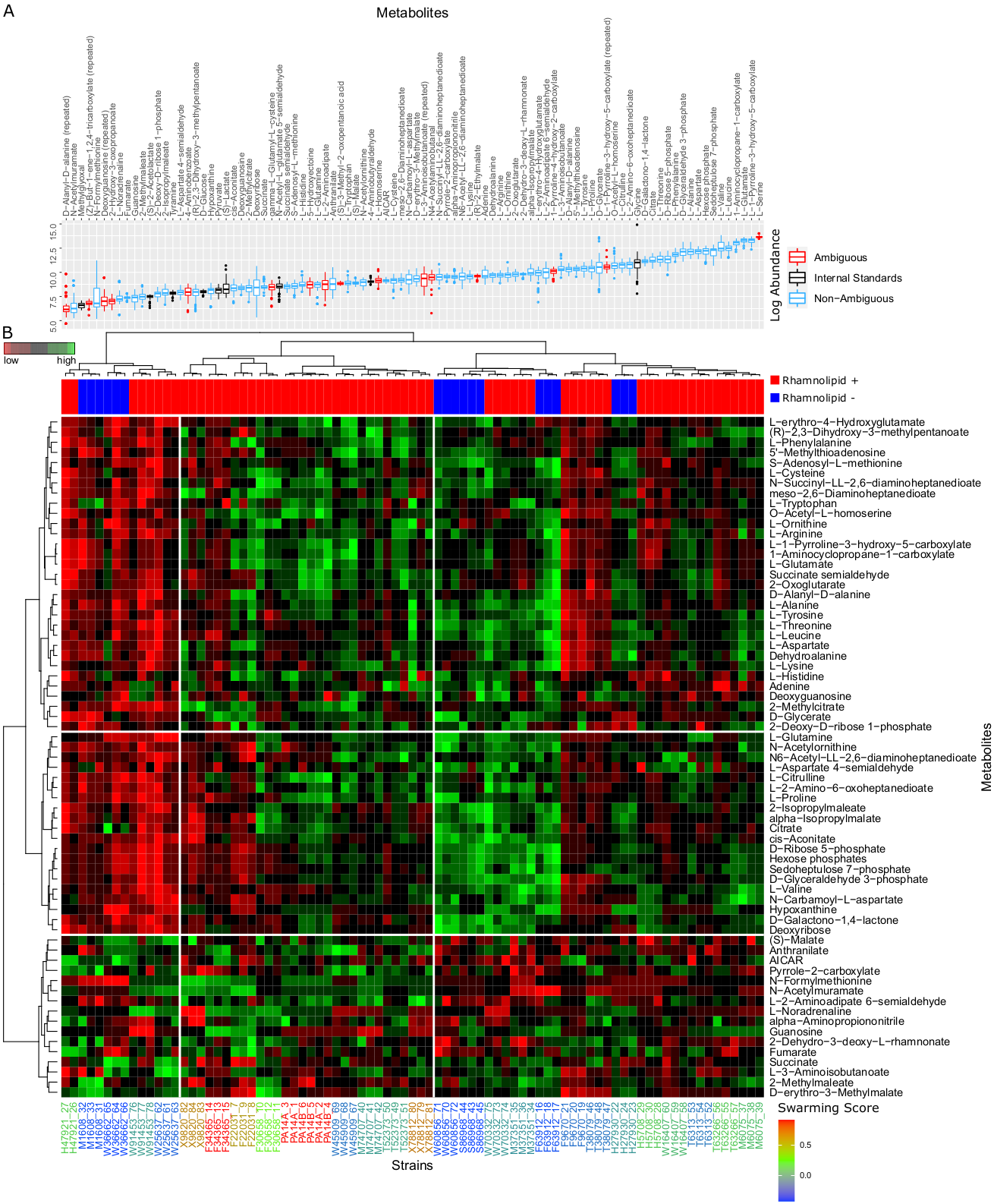
**Figure 2.** Growth curve features distinguish rhamnolipid producers from non-producers. (**A**,**B**) Unsupervised feature selection using non-negative matrix factorization (NNMF), which decomposes growth curves of all *Pseudomonas* isolates into three additive basis functions (features) such that each growth curve can be approximately represented by the weighted sum of these functions. (**A**) Growth curves from both rhamnolipid producers (orange) and non-producers (blue). (**B**) Decomposed components (basis function multiplied by weights; left panels) and weights (right panels) from NNMF grouped by rhamnolipid (RL) production. The shaded areas represent 95% bootstrap confidence interval of the mean. (**C-E**) Supervised feature selection using Random Forest classifier. (**C,D**) Feature extraction method. Each growth curve (excluding the initial lag phase) was divided into three phases (**C**) and each phase was described by 7 quantitative features (**D**). (**F**) Ranking of feature importance in classifying rhamnolipid producers. Inset: boxplot of maximum specific growth rate of phase I grouped by rhamnolipid production. Welch’s t-test was used in (**B**) and (**E**) for significance testing. \*\*\*\*, p-value 0.0001; \*, p-value 0.05; ns, p-value > 0.05.

A picture containing text, map

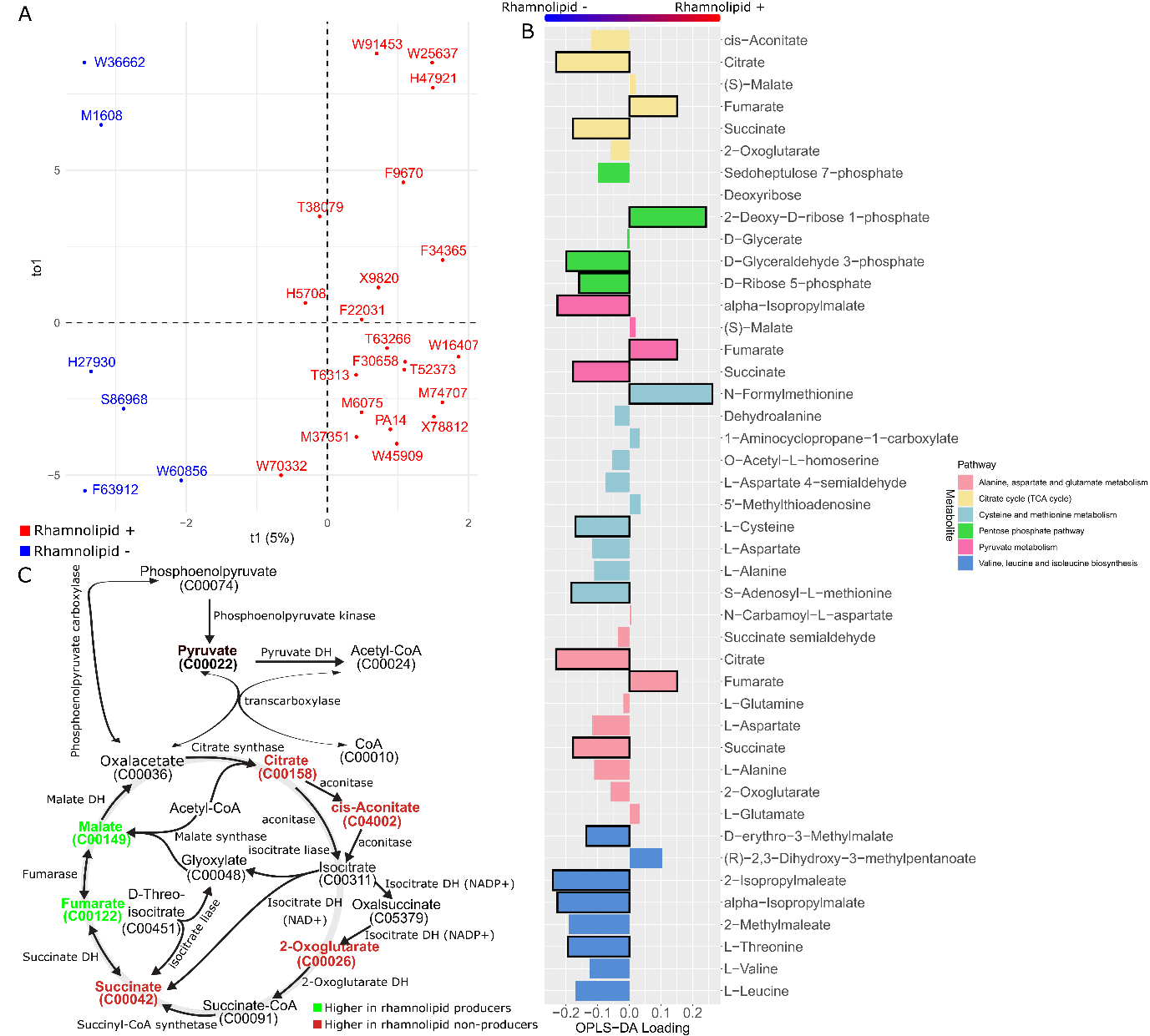
Description automatically generated

**Supplementary Figure 4.** Growth curve of *P. aeruginosa* strains in glycerol minimal medium. Phase I, II, and III are colored by red, blue and green respectively.

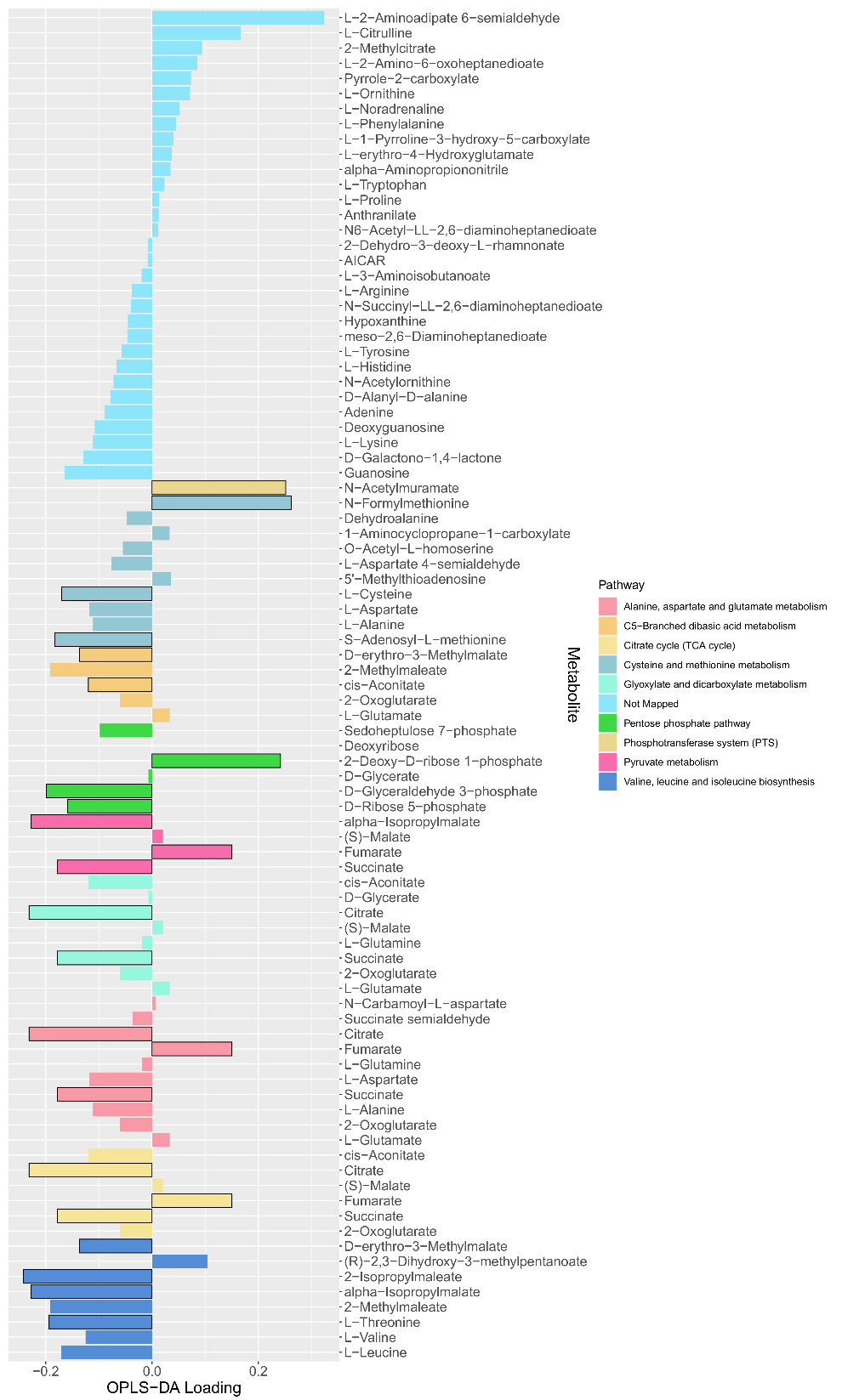


**Supplementary Figure S5.** Metabolite peak area before (left panel) and after (right panel) imputation. The missing value proportion of a metabolite in the x-axis represents the frequency of missing values of the metabolite across all replicates of all strains in our study. The peak range of a metabolite in the y-axis is defined as the maximum peak area minus the minimum peak area among all non-missing values of the same metabolite across all our strains. 

**Figure 3.** A.Levels of all the metabolites identified by LC-MS. The values shown here are absolute pre-normalization of all the compounds in the uncurated metabolite table (logarithm of the peak area) used for the HCA normalization. The blue-coded metabolites had a non-ambiguous peak identification. The black-coded metabolites were used as Internal Standard in the Cross-Contribution Compensating Multiple Standard Normalization. These were assumed to be constant across strains by a Kruskal-Wallis test and a level of significance of 0.05 (after Benjamini-Hochberg *p-*value correction). The red-coded metabolites were removed after the HCA because of uncertain identity. Many metabolites were renamed during the curation, producing some duplicated names, which are labeled as “repeated”, and were removed. B.Hierarchical Cluster Analysis of the metabolic profiles.We included strain M6075, which was not in our phylogenetic assay due to unsatisfactory genomic sequencing and assembly. The intracellular metabolites measured for each strain show that swarmers have similar metabolic profiles. The metabolic profile of each strain was determined by LC-MS and normalized with CCMN method. 3 replicates were done for each strain. This heatmap is only showing the compounds that the experimentalists were confident about their identity. The data was clustered using Euclidean distance and Ward D aggregation method. The swarming behavior of each strain is color coded. The high swarmers are red-coded and low swarmers blue coded. The strains that swarm better appear all grouped in the same sector. F30658, which is a mild swarmer, is the only swarmer strain that is not grouped to the other ones. Regarding the metabolites, some sectors of the heatmap are higher in some groups of strains, but there is not any obvious pattern differentiating the strains with high swarming score.



**Figure 4. Differential metabolomics between rhamnolipid producers and non-producers.** (**A)** Scores plot of the OPLS model. The strains appear separated according to whether they produce or not rhamnolipids across the predictor component (t1). (**B**). Loading values for the predictor component of a selected number of metabolites, mapped to the pathways they are involved in according to KEGG database, among the ones found to be potentially altered according to FELLA algorithm. (**C**). TCA Cycle map. The compounds that are included in our metabolomic dataset are highlighted in bold. The colored ones are the ones that are correlated to any of the rhamnolipid production phenotypes.



**Supplementary Figure 6.** The loading values of all predictive metabolites of the OPLS-DA model. The differential metabolites between producers and non-producers were determined by a Mann Whitney test (adjusted *p-*values with Bonferroni-Hochberg method) with a level of significance of 0.05 (bars with black outline) and used as input for a metabolic pathway enrichment with FELLA algorithm. The colors indicate the mapped metabolite pathway for each metabolite.

A close up of a computer

Description automatically generated

**Supplementary Figure 7.** Volcano plot of metabolomics data between wild-type *P. aeruginosa* UCBPP-PA14 strain and its *rhlA* mutant (replotted with permission from (Boyle *et al*, 2017)).

A close up of a sign

Description automatically generated

**Figure 5** Growth and rhamnolipid secretion are constrained by redox status. The redox stress levels are perturbed by altering fluxes of NADH (reduced nicotinamide adenine dinucleotide; **A**), NADPH (reduced nicotinamide adenine dinucleotide phosphate; **B**) and GSH (reduced glutathione; **C**). Upper panels are predicted maximum growth rates and lower panels are predicted maximum byproduct secretion fluxes. C:N indicates the carbon-to-nitrogen ratio between glycerol and ammonium in the culture medium. Abbreviations: HAA: 3-(3-hydroxyalkanoyloxy) alkanoate; monoRL: monorhamnolipid; diRL: dirhamnolipid; aKG: alpha-ketoglutarate.

A picture containing kite, flying, fireworks

Description automatically generated

**Supplementary Figure 8.** Theoretical estimation of threshold carbon (glycerol):nitrogen (ammonium) ratio above which carbon is in excess in the sense that carbon release through rhamnolipids and central carbon metabolites does not compromise biomass production. Abbreviations: HAA: 3-(3-hydroxyalkanoyloxy)alkanoate; monoRL: monorhamnolipid; diRL: dirhamnolipid; aKG: alpha-ketoglutarate.A picture containing text

Description automatically generated

**Supplementary Figure 9.** Growth curve of *Pseudomonas aeruginosa* strains in succinate minimal medium. Phase I, II, and III are colored by red, blue and green respectively.

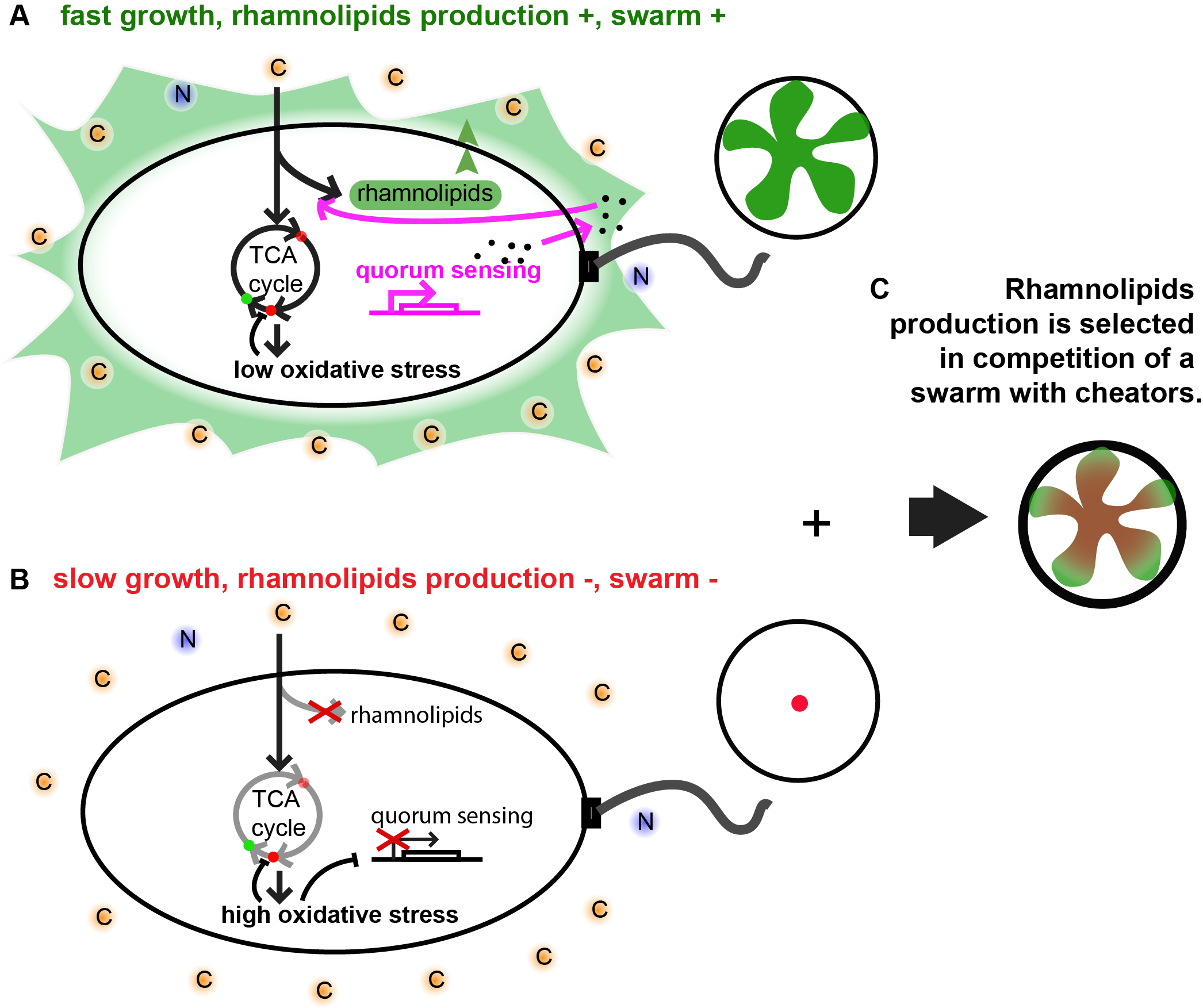
A picture containing outdoor, large, black, standing

Description automatically generated

**Supplementary Figure 10.** Comparison of growth curve features of *Pseudomonas aeruginosa* strains grown in glycerol and succinate minimal medium.A close up of a map

Description automatically generated

**Figure 6.**  Comparison of the ability of removing environmental hydrogen peroxide among strong rhamnolipid producers (RL++), weak producers (RL+), and non-producers (RL-). (A) The total amount of hydrogen peroxide removed from the environment. Negative values indicate net cellular production of hydrogen peroxide released to the environment. (B) The specific hydrogen peroxide removal rate. The lines represent the mean values and the shading areas represent standard deviation.



**Figure 7**. Metabolic model of rhamnolipid production in *P. aeruginosa*. A) During aerobic growth, cells rely on the TCA cycle to invest in biomass. If the cell could release accumulated oxidative stress, growth rate can be maintained high. The membrane redox stress is partially released and therefore the flux between succinate (red dot) and fumarate (green dot) is less reduced, resulting in higher fumarate level than in rhamnolipid non-producer. When cells reach certain density and nitrogen becomes limited, rhamnolipid production is turned on by quorum sensing. (B) If TCA cycle is significantly slowed down by redox stress, quorum sensing fails to be activated and rhamnolipids is not produced.. C) When rhamnolipids producing cells are competed with cheators, the producers do not lose the completion due to the metabolic constrains, which imposes a selection for the cooperative secretion. red dots: metabolites whose levels are lower in rhamnolipid producers; green dots: metabolites whose levels are higher in rhamnolipid producers.

|  |  |  |  |
| --- | --- | --- | --- |
| **KEGG.id** | **Entry.type** | **KEGG.name** | **p.score** |
| pau00020 | pathway | Citrate cycle (TCA cycle) - Pseudomonas aerug... | 2.051E-02 |
| pau00030 | pathway | Pentose phosphate pathway - Pseudomonas aerug... | 2.657E-04 |
| pau00250 | pathway | Alanine, aspartate and glutamate metabolism -... | 3.502E-02 |
| pau00270 | pathway | Cysteine and methionine metabolism - Pseudomo... | 3.106E-04 |
| pau00290 | pathway | Valine, leucine and isoleucine biosynthesis -... | 1.000E-06 |
| pau00620 | pathway | Pyruvate metabolism - Pseudomonas aeruginosa ... | 3.766E-02 |
| pau00630 | pathway | Glyoxylate and dicarboxylate metabolism - Pse... | 1.105E-02 |
| pau00660 | pathway | C5-Branched dibasic acid metabolism - Pseudom... | 8.242E-03 |
| pau01503 | pathway | Cationic antimicrobial peptide (CAMP) resista... | 3.891E-02 |
| pau02060 | pathway | Phosphotransferase system (PTS) - Pseudomonas... | 4.222E-02 |
| M00001 | module | Glycolysis (Embden-Meyerhof pathway), glucose... | 1.613E-02 |
| M00002 | module | Glycolysis, core module involving three-carbo... | 4.022E-03 |
| M00003 | module | Gluconeogenesis, oxaloacetate => fructose-6P | 2.299E-03 |
| M00004 | module | Pentose phosphate pathway (Pentose phosphate ... | 3.993E-03 |
| M00005 | module | PRPP biosynthesis, ribose 5P => PRPP | 1.000E-06 |
| M00007 | module | Pentose phosphate pathway, non-oxidative phas... | 1.030E-04 |
| M00009 | module | Citrate cycle (TCA cycle, Krebs cycle) | 7.517E-04 |
| M00010 | module | Citrate cycle, first carbon oxidation, oxaloa... | 4.789E-04 |
| M00011 | module | Citrate cycle, second carbon oxidation, 2-oxo... | 8.189E-03 |
| M00012 | module | Glyoxylate cycle | 1.577E-03 |
| M00019 | module | Valine/isoleucine biosynthesis, pyruvate => v... | 5.143E-03 |
| M00032 | module | Lysine degradation, lysine => saccharopine =>... | 2.928E-04 |
| M00049 | module | Adenine ribonucleotide biosynthesis, IMP => A... | 3.245E-02 |
| M00149 | module | Succinate dehydrogenase, prokaryotes | 4.051E-02 |
| M00165 | module | Reductive pentose phosphate cycle (Calvin cyc... | 8.362E-04 |
| M00167 | module | Reductive pentose phosphate cycle, glyceralde... | 2.917E-04 |
| M00173 | module | Reductive citrate cycle (Arnon-Buchanan cycle... | 1.377E+09 |
| M00308 | module | Semi-phosphorylative Entner-Doudoroff pathway... | 2.511E-02 |
| M00344 | module | Formaldehyde assimilation, xylulose monophosp... | 5.467E-03 |
| M00374 | module | Dicarboxylate-hydroxybutyrate cycle | 2.552E-02 |
| M00432 | module | Leucine biosynthesis, 2-oxoisovalerate => 2-o... | 1.000E-06 |
| M00535 | module | Isoleucine biosynthesis, pyruvate => 2-oxobut... | 1.000E-06 |
| M00570 | module | Isoleucine biosynthesis, threonine => 2-oxobu... | 8.290E-04 |
| M00620 | module | Incomplete reductive citrate cycle, acetyl-Co... | 1.000E-06 |
| M00633 | module | Semi-phosphorylative Entner-Doudoroff pathway... | 3.830E-02 |
| M00740 | module | Methylaspartate cycle | 4.398E-02 |
| 1.1.1.85 | enzyme | 3-isopropylmalate dehydrogenase | 1.000E-06 |
| 1.2.1.12 | enzyme | glyceraldehyde-3-phosphate dehydrogenase (pho... | 1.578E-03 |
| 1.2.1.9 | enzyme | glyceraldehyde-3-phosphate dehydrogenase (NAD... | 4.815E-04 |
| 1.3.5.1 | enzyme | succinate dehydrogenase | 5.028E+08 |
| 1.3.5.4 | enzyme | fumarate reductase (quinol) | 5.028E+08 |
| 1.4.1.9 | enzyme | leucine dehydrogenase | 1.032E-02 |
| 1.8.4.15 | enzyme | protein dithiol oxidoreductase (disulfide-for... | 3.891E-02 |
| 2.2.1.1 | enzyme | transketolase | 1.838E+08 |
| 2.2.1.2 | enzyme | transaldolase | 5.939E-04 |
| 2.3.3.1 | enzyme | citrate (Si)-synthase | 1.000E-06 |
| 2.3.3.13 | enzyme | 2-isopropylmalate synthase | 1.000E-06 |
| 2.4.2.1 | enzyme | purine-nucleoside phosphorylase | 1.000E-06 |
| 2.4.2.2 | enzyme | pyrimidine-nucleoside phosphorylase | 1.000E-06 |
| 2.5.1.49 | enzyme | O-acetylhomoserine aminocarboxypropyltransfer... | 4.137E-02 |
| 2.6.1.2 | enzyme | alanine transaminase | 1.000E-06 |
| 2.6.1.42 | enzyme | branched-chain-amino-acid transaminase | 1.985E-02 |
| 2.6.1.66 | enzyme | valine---pyruvate transaminase | 1.000E-06 |
| 2.7.1.15 | enzyme | ribokinase | 1.791E-02 |
| 2.7.1.221 | enzyme | N-acetylmuramate 1-kinase | 1.000E-06 |
| 2.7.2.3 | enzyme | phosphoglycerate kinase | 1.747E-02 |
| 2.7.6.1 | enzyme | ribose-phosphate diphosphokinase | 1.000E-06 |
| 2.7.9.2 | enzyme | pyruvate, water dikinase | 7.185E-03 |
| 2.8.3.18 | enzyme | succinyl-CoA:acetate CoA-transferase | 1.672E-02 |
| 3.1.3.105 | enzyme | N-acetyl-D-muramate 6-phosphate phosphatase | 1.000E-06 |
| 3.1.3.11 | enzyme | fructose-bisphosphatase | 3.302E-02 |
| 3.1.3.48 | enzyme | protein-tyrosine-phosphatase | 1.524E-02 |
| 3.2.2.10 | enzyme | pyrimidine-5'-nucleotide nucleosidase | 1.000E-06 |
| 3.2.2.4 | enzyme | AMP nucleosidase | 1.000E-06 |
| 3.4.21.107 | enzyme | peptidase Do | 1.929E-02 |
| 3.4.24.40 | enzyme | serralysin | 3.891E-02 |
| 3.5.1.28 | enzyme | N-acetylmuramoyl-L-alanine amidase | 1.000E-06 |
| 3.5.4.2 | enzyme | adenine deaminase | 1.786E-02 |
| 3.6.1.13 | enzyme | ADP-ribose diphosphatase | 7.396E+08 |
| 3.7.1.20 | enzyme | 3-fumarylpyruvate hydrolase | 2.495E-02 |
| 4.1.1.31 | enzyme | phosphoenolpyruvate carboxylase | 4.009E-02 |
| 4.1.2.13 | enzyme | fructose-bisphosphate aldolase | 2.966E-03 |
| 4.1.2.14 | enzyme | 2-dehydro-3-deoxy-phosphogluconate aldolase | 2.363E-02 |
| 4.1.2.48 | enzyme | low-specificity L-threonine aldolase | 9.416E-04 |
| 4.1.3.1 | enzyme | isocitrate lyase | 1.814E-02 |
| 4.2.1.2 | enzyme | fumarate hydratase | 1.000E-06 |
| 4.2.1.20 | enzyme | tryptophan synthase | 3.153E-02 |
| 4.2.1.3 | enzyme | aconitate hydratase | 1.000E-06 |
| 4.2.1.33 | enzyme | 3-isopropylmalate dehydratase | 1.000E-06 |
| 4.2.1.35 | enzyme | (R)-2-methylmalate dehydratase | 1.000E-06 |
| 4.2.3.1 | enzyme | threonine synthase | 3.155E-02 |
| 4.3.1.1 | enzyme | aspartate ammonia-lyase | 1.000E-06 |
| 4.3.1.19 | enzyme | threonine ammonia-lyase | 1.000E-06 |
| 4.3.2.1 | enzyme | argininosuccinate lyase | 1.647E-02 |
| 4.3.2.2 | enzyme | adenylosuccinate lyase | 9.353E-04 |
| 5.1.3.1 | enzyme | ribulose-phosphate 3-epimerase | 1.624E-02 |
| 5.2.1.8 | enzyme | peptidylprolyl isomerase | 3.891E-02 |
| 5.3.1.1 | enzyme | triose-phosphate isomerase | 1.698E+08 |
| 5.3.1.6 | enzyme | ribose-5-phosphate isomerase | 2.338E-02 |
| 5.4.2.2 | enzyme | phosphoglucomutase (alpha-D-glucose-1,6-bisph... | 4.898E+09 |
| 6.2.1.5 | enzyme | succinate---CoA ligase (ADP-forming) | 6.932E-03 |
| 6.3.1.2 | enzyme | glutamine synthetase | 3.639E-02 |
| 6.4.1.1 | enzyme | pyruvate carboxylase | 4.182E+07 |
| R00086 | reaction | ATP phosphohydrolase | 3.936E-02 |
| R00122 | reaction | ADP phosphohydrolase | 3.029E-02 |
| R00177 | reaction | ATP:L-methionine S-adenosyltransferase | 1.894E-02 |
| R00182 | reaction | AMP phosphoribohydrolase | 1.000E-06 |
| R00220 | reaction | L-serine ammonia-lyase | 2.595E-02 |
| R00258 | reaction | L-Alanine:2-oxoglutarate aminotransferase | 4.726E-02 |
| R00351 | reaction | acetyl-CoA:oxaloacetate C-acetyltransferase (... | 1.000E-06 |
| R00352 | reaction | acetyl-CoA:oxaloacetate C-acetyltransferase [... | 1.000E-06 |
| R00402 | reaction | succinate:NAD+ oxidoreductase | 1.000E-06 |
| R00405 | reaction | Succinate:CoA ligase (ADP-forming) | 4.397E+09 |
| R00432 | reaction | Succinate:CoA ligase (GDP-forming) | 2.166E-02 |
| R00446 | reaction | L-lysine:NAD+ 6-oxidoreductase (deaminating) | 1.000E-06 |
| R00479 | reaction | isocitrate glyoxylate-lyase (succinate-formin... | 3.522E-02 |
| R00490 | reaction | L-aspartate ammonia-lyase (fumarate-forming) | 1.000E-06 |
| R00510 | reaction | cytidine-5'-monophosphate phosphoribohydrolas... | 1.000E-06 |
| R00653 | reaction | N-formyl-L-methionine amidohydrolase | 1.000E-06 |
| R00751 | reaction | L-threonine acetaldehyde-lyase (glycine-formi... | 1.000E-06 |
| R00782 | reaction | L-cysteine hydrogen-sulfide-lyase (deaminatin... | 1.000E-06 |
| R00891 | reaction | L-serine hydro-lyase (adding hydrogen sulfide... | 1.000E-06 |
| R00892 | reaction | L-cysteine:NAD+ oxidoreductase | 1.384E-02 |
| R00893 | reaction | L-Cysteine:oxygen oxidoreductase | 7.899E-03 |
| R00894 | reaction | L-glutamate:L-cysteine gamma-ligase (ADP-form... | 3.659E-03 |
| R00895 | reaction | L-Cysteine:2-oxoglutarate aminotransferase | 1.786E-02 |
| R00897 | reaction | O3-acetyl-L-serine:hydrogen-sulfide 2-amino-2... | 4.683E-03 |
| R00901 | reaction | L-cysteine hydrogen-sulfide-lyase (adding sul... | 4.539E-03 |
| R00994 | reaction | (2R,3S)-3-methylmalate:NAD+ oxidoreductase | 1.000E-06 |
| R00996 | reaction | L-threonine ammonia-lyase (2-oxobutanoate-for... | 1.000E-06 |
| R00999 | reaction | O-Succinyl-L-homoserine succinate-lyase (deam... | 2.925E-02 |
| R01001 | reaction | L-cystathionine cysteine-lyase (deaminating | 3.769E-04 |
| R01015 | reaction | D-glyceraldehyde-3-phosphate aldose-ketose-is... | 1.000E-06 |
| R01049 | reaction | ATP:D-ribose-5-phosphate diphosphotransferase | 1.000E-06 |
| R01051 | reaction | ATP:D-ribose 5-phosphotransferase | 1.000E-06 |
| R01053 | reaction | Ribose-5-phosphate:ammonia ligase (ADP-formin... | 1.804E+07 |
| R01054 | reaction | ADP-ribose ribophosphohydrolase | 9.884E+08 |
| R01055 | reaction | uracil hydro-lyase (adding D-ribose 5-phospha... | 2.245E-03 |
| R01056 | reaction | D-ribose-5-phosphate aldose-ketose-isomerase | 1.000E-06 |
| R01057 | reaction | D-Ribose 1,5-phosphomutase | 1.000E-06 |
| R01058 | reaction | D-glyceraldehyde 3-phosphate:NADP+ oxidoreduc... | 1.000E-06 |
| R01059 | reaction | ATP:D-glyceraldehyde 3-phosphotransferase | 1.000E-06 |
| R01061 | reaction | D-glyceraldehyde-3-phosphate:NAD+ oxidoreduct... | 1.357E-04 |
| R01063 | reaction | D-glyceraldehyde-3-phosphate:NADP+ oxidoreduc... | 1.690E-04 |
| R01066 | reaction | 2-deoxy-D-ribose-5-phosphate acetaldehyde-lya... | 1.494E-04 |
| R01067 | reaction | D-Fructose 6-phosphate:D-glyceraldehyde-3-pho... | 5.048E+08 |
| R01068 | reaction | D-fructose-1,6-bisphosphate D-glyceraldehyde-... | 8.170E+09 |
| R01069 | reaction | D-tagatose 1,6-bisphosphate D-glyceraldehyde-... | 4.029E-04 |
| R01070 | reaction | beta-D-fructose-1,6-bisphosphate D-glyceralde... | 8.852E+08 |
| R01082 | reaction | (S)-malate hydro-lyase (fumarate-forming) | 1.000E-06 |
| R01083 | reaction | N6-(1,2-dicarboxyethyl)AMP AMP-lyase (fumarat... | 2.469E-03 |
| R01085 | reaction | 3-fumarylpyruvate fumarylhydrolase | 1.563E-02 |
| R01086 | reaction | 2-(Nomega-L-arginino)succinate arginine-lyase... | 5.126E-03 |
| R01087 | reaction | Maleate cis-trans-isomerase | 2.433E+09 |
| R01113 | reaction | Glutathione:L-cystine oxidoreductase | 3.593E-02 |
| R01128 | reaction | 5'-Inosinate phosphoribohydrolase | 1.000E-06 |
| R01213 | reaction | acetyl-CoA:3-methyl-2-oxobutanoate C-acetyltr... | 1.000E-06 |
| R01215 | reaction | L-Valine:pyruvate aminotransferase | 3.535E-02 |
| R01244 | reaction | Adenine aminohydrolase | 2.846E-02 |
| R01270 | reaction | Nicotinamide D-ribonucleotide phosphoribohydr... | 1.000E-06 |
| R01322 | reaction | Citrate:CoA ligase (ADP-forming) | 3.624E-04 |
| R01324 | reaction | citrate hydroxymutase | 1.000E-06 |
| R01325 | reaction | citrate hydro-lyase (cis-aconitate-forming) | 1.000E-06 |
| R01364 | reaction | 4-fumarylacetoacetate fumarylhydrolase | 2.477E-02 |
| R01440 | reaction | D-Xylulose-5-phosphate:formaldehyde glycolald... | 4.322E+09 |
| R01465 | reaction | L-threonine:NAD+ oxidoreductase | 1.000E-06 |
| R01466 | reaction | O-phospho-L-homoserine phosphate-lyase (addin... | 6.537E+09 |
| R01561 | reaction | adenosine:phosphate alpha-D-ribosyltransferas... | 2.961E-03 |
| R01570 | reaction | thymidine:phosphate deoxy-alpha-D-ribosyltran... | 1.000E-06 |
| R01621 | reaction | D-xylulose 5-phosphate D-glyceraldehyde-3-pho... | 1.000E-06 |
| R01641 | reaction | sedoheptulose-7-phosphate:D-glyceraldehyde-3-... | 1.000E-06 |
| R01652 | reaction | 4-Methyl-2-oxopentanoate + CO2 <=> (2S)-2-Iso... | 4.391E-03 |
| R01827 | reaction | sedoheptulose-7-phosphate:D-glyceraldehyde-3-... | 1.594E-04 |
| R01830 | reaction | beta-D-Fructose 6-phosphate:D-glyceraldehyde-... | 3.088E+08 |
| R01863 | reaction | inosine:phosphate alpha-D-ribosyltransferase | 7.366E-03 |
| R01867 | reaction | (S)-dihydroorotate:fumarate oxidoreductase | 1.035E-03 |
| R01876 | reaction | uridine:phosphate alpha-D-ribosyltransferase | 2.008E-03 |
| R01969 | reaction | Deoxyguanosine:orthophosphate ribosyltransfer... | 1.000E-06 |
| R02147 | reaction | guanosine:phosphate alpha-D-ribosyltransferas... | 1.837E-02 |
| R02164 | reaction | succinate:quinone oxidoreductase | 1.000E-06 |
| R02294 | reaction | N-Ribosylnicotinamide:orthophosphate ribosylt... | 4.332E-02 |
| R02296 | reaction | cytidine:orthophosphate alpha-D-ribosyltransf... | 3.703E-02 |
| R02297 | reaction | Xanthosine:orthophosphate ribosyltransferase | 4.035E-02 |
| R02313 | reaction | N6-(L-1,3-Dicarboxypropyl)-L-lysine:NAD+ oxid... | 1.000E-06 |
| R02315 | reaction | N6-(L-1,3-Dicarboxypropyl)-L-lysine:NADP+ oxi... | 1.000E-06 |
| R02317 | reaction | (S)-2-amino-6-oxohexanoate hydro-lyase (spont... | 1.000E-06 |
| R02340 | reaction | (1S,2R)-1-C-(indol-3-yl)glycerol 3-phosphate ... | 7.357E-03 |
| R02409 | reaction | Coenzyme A:oxidized-glutathione oxidoreductas... | 4.474E-02 |
| R02484 | reaction | deoxyuridine:orthophosphate 2-deoxy-D-ribosyl... | 1.000E-06 |
| R02557 | reaction | Deoxyadenosine:orthophosphate ribosyltransfer... | 1.000E-06 |
| R02722 | reaction | L-serine hydro-lyase [adding 1-C-(indol-3-yl)... | 3.011E-03 |
| R02748 | reaction | Deoxyinosine:orthophosphate ribosyltransferas... | 1.000E-06 |
| R02749 | reaction | 2-deoxy-D-ribose 1-phosphate 1,5-phosphomutas... | 1.000E-06 |
| R03102 | reaction | L-2-aminoadipate-6-semialdehyde:NAD+ 6-oxidor... | 1.000E-06 |
| R03103 | reaction | L-2-aminoadipate-6-semialdehyde:NADP+ 6-oxido... | 1.000E-06 |
| R03217 | reaction | O-Acetyl-L-homoserine succinate-lyase (adding... | 3.059E-02 |
| R03260 | reaction | O-Succinyl-L-homoserine succinate-lyase (addi... | 2.438E-03 |
| R03524 | reaction | L-cysteine hydrogen-sulfide-lyase (adding HCN... | 3.504E-02 |
| R03896 | reaction | (R)-2-Methylmalate hydro-lyase (2-methylmalea... | 1.149E+08 |
| R03898 | reaction | 2-methylmaleate hydratase | 1.000E-06 |
| R03968 | reaction | 2-Isopropylmalate hydro-lyase | 1.000E-06 |
| R04001 | reaction | 3-Isopropylmalate hydro-lyase | 1.000E-06 |
| R04390 | reaction | L-2-aminoadipate-6-semialdehyde:NADP+ 6-oxido... | 2.104E-02 |
| R04426 | reaction | (2R,3S)-3-isopropylmalate:NAD+ oxidoreductase | 1.000E-06 |
| R04559 | reaction | 1-(5'-Phosphoribosyl)-5-amino-4-(N-succinocar... | 7.225E-03 |
| R05598 | reaction | Benzylsuccinate fumarate-lyase | 1.083E-02 |
| R05605 | reaction | 2-dehydro-3-deoxy-6-phospho-D-gluconate D-gly... | 2.199E-03 |
| R05636 | reaction | 1-Deoxy-D-xylulose-5-phosphate pyruvate-lyase... | 3.417E-03 |
| R06531 | reaction | ATP:L-threonine O-phosphotransferase | 2.361E-04 |
| R06590 | reaction | sedoheptulose-7-phosphate:D-glyceraldehyde-3-... | 1.654E-02 |
| R06789 | reaction | (2Z,4E,7E)-2-Hydroxy-6-oxonona-2,4,7-triene-1... | 1.163E-02 |
| R06903 | reaction | 2-Methylnaphthalene + Fumarate <=> 2-Naphtylm... | 2.194E-02 |
| R06987 | reaction | propanoyl-CoA:formate C-propanoyltransferase | 2.218E-02 |
| R07159 | reaction | D-glyceraldehyde-3-phosphate:ferredoxin oxido... | 1.000E-06 |
| R07274 | reaction | O-phospho-L-serine:hydrogen-sulfide 2-amino-2... | 1.948E-04 |
| R07399 | reaction | Acetyl-CoA + Pyruvate + H2O <=> (R)-2-Methylm... | 2.986E-03 |
| R07456 | reaction | D-ribulose 5-phosphate,D-glyceraldehyde 3-pho... | 3.194E+09 |
| R08323 | reaction | 7-Mercaptoheptanoic acid + L-Threonine + ATP ... | 5.991E-03 |
| R08555 | reaction | (R)-lactate hydro-lyase (adding N-acetyl-D-gl... | 9.758E-03 |
| R08559 | reaction | protein-N(pi)-phosphohistidine:N-acetylmurama... | 1.000E-06 |
| R10049 | reaction | 2-oxopropanal:D-fructose 1,6-bisphosphate gly... | 4.546E-02 |
| R10052 | reaction | (2R,3S)-3-isopropylmalate:NAD+ oxidoreductase | 1.000E-06 |
| R10090 | reaction | citrate:N6-acetyl-N6-hydroxy-L-lysine ligase ... | 6.270E-03 |
| R10170 | reaction | (2R,3S)-3-isopropylmalate hydro-lyase (2-isop... | 1.000E-06 |
| R10270 | reaction | (5R)-5-phosphooxy-L-lysine phosphate-lyase (d... | 4.432E-04 |
| R10343 | reaction | succinyl-CoA:acetate CoA-transferase | 1.608E-03 |
| R10660 | reaction | fumarate CoM:CoB oxidoreductase (succinate fo... | 6.353E+09 |
| R10677 | reaction | citrate:L-glutamate ligase (ADP-forming) | 2.358E-02 |
| R10699 | reaction | L-lysine:8-amino-7-oxononanoate aminotransfer... | 1.201E-03 |
| R10860 | reaction | D-glyceraldehyde-3-phosphate:NAD+ oxidoreduct... | 1.000E-06 |
| R10888 | reaction | C20914 + L-Threonine <=> C20915 | 1.339E-02 |
| R10891 | reaction | Tabtoxin + H2O <=> Tabtoxinine-beta-lactam + ... | 1.168E-03 |
| R10892 | reaction | tabtoxinine-beta-lactam:L-threonine ligase (A... | 8.832E-04 |
| R10940 | reaction | fumarate:L-2,3-diaminopropanoate ligase (AMP-... | 6.289E-03 |
| R10973 | reaction | D-ribose-2,5-bisphosphate 2-phosphohydrolase | 2.912E-03 |
| R11024 | reaction | ATP:N-acetyl-D-muramate 1-phosphotransferase | 1.000E-06 |
| R11679 | reaction | (S)-2-amino-6-oxohexanoate:NADP+ oxidoreducta... | 1.000E-06 |
| R11785 | reaction | N-acetyl-D-muramate 6-phosphate phosphohydrol... | 1.000E-06 |
| R12308 | reaction | L-2,3-diaminopropanoate:citrate ligase (2-[(L... | 2.769E-03 |
| R12353 | reaction | D-Ornithine + Citrate + ATP <=> N5-Citryl-D-o... | 4.337E-04 |
| R12354 | reaction | N5-Citryl-D-ornithine + Citrate + ATP <=> Sta... | 3.695E-02 |
| R12596 | reaction | Spermidine + Citrate + ATP <=> N-Citryl-sperm... | 2.685E-04 |
| C00009 | compound | Orthophosphate | 9.902E-03 |
| C00042 | compound | Succinate | 1.000E-06 |
| C00058 | compound | Formate | 1.695E-02 |
| C00073 | compound | L-Methionine | 3.505E-03 |
| C00097 | compound | L-Cysteine | 1.000E-06 |
| C00109 | compound | 2-Oxobutanoate | 1.133E-02 |
| C00117 | compound | D-Ribose 5-phosphate | 1.000E-06 |
| C00118 | compound | D-Glyceraldehyde 3-phosphate | 1.000E-06 |
| C00122 | compound | Fumarate | 1.000E-06 |
| C00141 | compound | 3-Methyl-2-oxobutanoic acid | 2.304E-02 |
| C00158 | compound | Citrate | 1.000E-06 |
| C00188 | compound | L-Threonine | 1.000E-06 |
| C00672 | compound | 2-Deoxy-D-ribose 1-phosphate | 1.000E-06 |
| C00673 | compound | 2-Deoxy-D-ribose 5-phosphate | 2.442E-02 |
| C02226 | compound | 2-Methylmaleate | 3.627E-04 |
| C02504 | compound | alpha-Isopropylmalate | 1.000E-06 |
| C02631 | compound | 2-Isopropylmaleate | 1.000E-06 |
| C02713 | compound | N-Acetylmuramate | 1.000E-06 |
| C03145 | compound | N-Formylmethionine | 1.000E-06 |
| C04076 | compound | L-2-Aminoadipate 6-semialdehyde | 1.000E-06 |
| C04411 | compound | (2R,3S)-3-Isopropylmalate | 1.000E-06 |
| C06032 | compound | D-erythro-3-Methylmalate | 1.000E-06 |
| C16698 | compound | N-Acetylmuramic acid 6-phosphate | 2.110E-03 |

**Supplementary Table 1:** output of the FELLA algorithm. A set of differential compounds was determined with a Mann-Whitney U test between rhamnolipid producers and rhamnolipid non producers (*p-*value adjusted with Benjamini-Hochberg method, alpha = 0.05). These compounds were used for applying diffusion algorithms over a graph consisting in all entries for *Pseudomonas aeruginosa* strain UCBPP-PA14 in KEGG database. The entries shown in the table are the ones with a significative probability of receiving part of the simulated flux.