**Metabolic constraints drive the swarming diversity in pathogenic bacteria**

**Abstract**

Many species of bacteria have the incredible ability to form swarms and move cooperatively across surfaces. Swarming motility is important to the fitness and virulence of bacteria, yet strains of the same bacterial species vary widely: What underlies swarming diversity? Here we presented comparative analyses of genetics, metabolomics and growth curve data to investigate why 28 clinical *P. aeruginosa* strains isolated from hospitalized patients differ in their abilities to swarm and produce rhamnolipids—a necessary but not sufficient requirement for swarming. Despite the diversity can be partially attributed to the loss of genes essential to both phenotypes, the comparative analyses of metabolomics and growth curve data reveal increased oxidative stress in the non-rhamnolipid-producers with perturbed tricarboxylic acid (TCA) cycle, amino acid metabolism, and slower exponential growth rate. Genome-scale model simulations and validation experiments confirm the links between growth, redox homeostasis and carbon overflow. Our findings suggest that redox homeostasis is a fundamental metabolic constraint of rhamnolipid production and expands the “metabolic prudence” concept proposed a decade ago: production of rhamnolipids is prudently controlled by integrating metabolic cues from nutrients, cell population density as well as cellular redox status. By studying rhamnolipid production in *P. aeruginosa*, this work provides a specific example of using systems biology approach to link bacterial metabolism with social traits, and sheds further lights on the general impacts of metabolic byproduct secretion on bacterial physiology, ecology and evolutionary strategies.

**Significance**

Bacteria display and maintain rather complex behavior as a population by social interactions including cooperation. A central question about bacterial social behavior has asked how it is regulated and constrained by metabolic state of individual cells. The pathogen *Pseudomonas aeruginosa* has a peculiar metabolism that makes the environment-dependent decision to cooperate within populations by secreting rhamnolipids—the public good required for swarming. Here we combine computational approaches with experiments to investigate how cell metabolism informs the decision of swarming as well as rhamnolipid production and drives the observed phenotypic diversity across 28 clinical *P. aeruginosa* strains isolated from patients. *P. aeruginosa* senses intracellular redox state and integrates the redox signal with nutrient cues as well as the information of population density to determine if it should produce rhamnolipids. In a synthetic minimum medium with glycerol as the sole carbon source, the non-rhamnolipid-producers have perturbed metabolic activity, diminished ability to detoxify hydrogen peroxide, and grow slower. When the producers are cultured in succinate minimal medium, they all become non-producers since succinate as a TCA cycle intermediate drives production of reactive oxygen species and promotes oxidative stress. Our work sheds new lights on the complex metabolic sensing and regulatory processes that govern the cellular decision to cooperate.

**Introduction**

Swarming motility is a flagella-driven movement of millions of bacterial cells to spread together over solid or semi-solid surfaces, and travel centimeter-long distances in a few hours (Deforet *et al*, 2014). The motility is narrowly conserved within certain bacterial species that are typically in the phyla Alpha-proteobacteria and Gamma-proteobacteria (Kearns, 2010), suggesting that swarming is a complex phenotype that cannot be evolved by simply acquiring a few set of genes. The biological complexity of swarming is also reflected in the multi-scale nature of its underlying functional requirements, including flagellar or pili biosynthesis to propel forward, cell-cell interactions to maintain population integrity during migration, and collective secretion of biosurfactants as wetting agents to lubricate the surface and lower surface tension (Kearns, 2010). Over the past two decades, considerable progress has been made in understanding the molecular mechanisms 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 laboratory evolution (van Ditmarsch *et al*, 2013). Despite invaluable insights, these studies mainly focused on single genetic or environmental determinants but lack a systems-level understanding of the genetic and metabolic influences on the swarming phenotype (Kim & Surette, 2004; Inoue *et al*, 2007). Indeed, the molecular functions of hundreds of swarming-associated genes span a wide range of cellular activities including tricarboxylic acid (TCA) cycle and stress response (Inoue *et al*, 2007; Tremblay & Déziel, 2010), indicating that swarming is an emergent property that cannot be analyzed solely in terms of single genes or metabolic pathways.

Swarming bacteria are often associated with pathogenesis by showing enhanced antibiotic resistance and virulence (Overhage *et al*, 2008; Kearns, 2010; Wang *et al*, 2004). *Pseudomonas aeruginosa*—an opportunistic human pathogen and a major cause of hospital infections (Klevens *et al*, 2007)—has 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 the presence of excess carbon in relative to metabolic capacity (Xavier *et al*, 2011; Boyle *et al*, 2015; Mellbye & Schuster, 2014). The combination of quorum-sensing and metabolic signals prevent evolutionary exploitation of non-cooperative cheaters (e.g., an Δ*rhlA* mutant) by prudently ensuring that massive rhamnolipids are only produced at the right timing when the production benefits outweigh its costs, which would otherwise slows down growth and makes the cooperative trait susceptible to cheating (Xavier *et al*, 2011).

The metabolic prudence described above imposes one metabolic constraint on swarming. However, it remains generally unclear how cell metabolism informs the decision of swarming and what are other constraints that prevent swarming when violated. Since metabolism is the currency of all physiological processes that support life (Smith & Morowitz, 2004) and a major determinant of social behavior such as swarming (Biro & Stamps, 2010), searching for other metabolic constraints can help elucidate the principles of social behavior evolution by natural selection and uncover the links between individual-level processes and population-level fitness. Deletion of *rhlA* gene, which abolishes rhamnolipid production and swarming completely, had only minimal effects on cell metabolism (~10 intracellular metabolites were perturbed) (Boyle *et al*, 2017). Indirect perturbations of cell metabolism, such as changing carbon and nitrogen sources (Köhler *et al*, 2000) or antibiotic treatment (Bru *et al*, 2019), reveal intimate links between swarming and certain metabolic pathways. Direct perturbations through global metabolic network could not only abolish swarming but devastate metabolic homeostasis with major changes in the intracellular metabolome. For example, mutations in the gene *cbrA*—a global regulator that modulates carbon-nitrogen balance, virulence and antibiotic resistance (Yeung *et al*, 2011)—impaired swarming, reduced growth rate of *P. aeruginosa*, and disrupted metabolism by altering dozens of intracellular metabolites (Boyle *et al*, 2017). Notably, the *ΔcbrA* mutant still had the necessary flagella to swim and expressed even more rhamnolipids than its parent wild-type strain. More interestingly, the swarming phenotype can 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 (Boyle *et al*, 2017).

To discover new metabolic constraints of *P. aeruginosa* swarming, we collected a panel of 28 clinical strains isolated from hospitalized cancer patients in Memorial Sloan-Kettering Cancer Center (MSKCC) (Yan *et al*, 2017, 2019). Compared to laboratory-evolved strains, the main advantage of using clinical isolates is that they were evolved in many body sites of different patients and thus more likely to generate swarming diversity. Indeed, these clinical isolates constitute a wide range of swarming phenotypes, including 8 strains that neither swarm nor produce rhamnolipids. The phylogenetic relatedness of core genomes were unable to explain the observed diversity of swarming and rhamnolipid production, which can be partially explained by the loss of genes essential to both phenotypes (e.g., genes encoding flagellar proteins and quorum-sensing regulators) in the accessory genomes. By comparative analyses of metabolomics and growth curve data acquired in the liquid culture with glycerol as the sole carbon source, we found so far unknown characteristics shared within the non-producers of rhamnolipids including reduced tricarboxylic acid (TCA) cycle activity, increased levels of amino acids, and slower growth. We hypothesized that these metabolic changes are caused by increased oxidative stress and validated our theory with mathematical modeling and validation experiments. We provide strong evidences that rhamnolipid production is constrained by redox homeostasis with rhamnolipid producers more capable of detoxifying reactive oxygen species (ROS) produced by energy metabolism and particularly the TCA cycle. Our theory revealed the links between growth, redox homeostasis and carbon overflow, adding to the expanding literature on the studies of mechanistic explanation and evolution of bacterial overflow metabolism.

**Results**

**Swarming diversity is uncorrelated with phylogeny of core genomes.** Swarming colonies in the casamino-acid agar plate (see Methods) showed markedly different morphological shapes across the 28 clinical isolates and 3 reference type strains, with a half of strains unable to swarm in addition to a few strong swamers (Fig. 1). To quantify the swarming morphology, we used image analysis (see Methods) to measure five morphological features, including the swarming area (the percentage of the plate occupied by the colony), the maximum length (maximum diagonal of the rectangle fitting the colony), the circularity, the eccentricity and the length of a skeleton obtained from the image. A principal component analysis revealed that the maximum length and circularity of the colonies are the two features that best describe the diversity of swarming shapes (Supplementary. Fig. S1). By defining a swarming score as a linear combination of the two features (see Methods), we show that swarming does not correlate with phylogeny of the core genomes nor with the tissue of origin of each isolate (Moran’s I test, p-value=0.93). Since production of rhamnolipids is required for swarming motility, we further measured rhamnolipid production using a drop collapse assay (see Methods) in a synthetic media with glycerol as the sole carbon source and provided in excess for rhamnolipid production (Boyle *et al*, 2015). As expected, all strains able to swarm produced rhamnolipids but rhamnolipid producers may not swarm, confirming that rhamnolipid production is a necessary but not sufficient requirement for swarming. Similar to swarming, rhamnolipid production is also uncorrelated with phylogeny (Fig. 1B**,** Moran’s I test, p=0.14).

How do rhamnolipid production and swarming evolve in the lineage of clinical isolates? To better understand the evolutionary relationship of the two phenotypes, we built an evolutionary model to reconstruct their ancestral states along the phylogenetic tree (see Methods). The common ancestor of all the strains can both swarm and produce rhamnolipids, suggesting loss of functions in the direction of evolution. Compared to swarming, rhamnolipid production is more evolutionarily conserved and its functional loss occurs more recently, as the immediate ancestors of all isolates that do not produce rhamnolipids are rhamnolipids producers (Supplementary Figure S2A). By contrast, the loss of swarming motility occurs earlier in the phylogenetic tree (Supplementary Figure S2B), which agrees with the fact that swarming depends on other factors than rhamnolipid production.

**Swarming diversity is partially explained by the genes absent in non-swarmers.** Since the core genomes were unable to explain the observed swarming diversity, we studied the role of accessory genomes, particularly those genes that are only missing in non-rhamnolipid-producers (Table 1). Notably, genes encoding rhamnolipid production pathways (*rhlA*, *rhlB*, *rhlC*) are intact in all non-producers. However, 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*. Except for *pilC* (Luo *et al*, 2015), all other genes have been experimentally validated as swarming regulators (Köhler *et al*, 2000; Kamatkar & Shrout, 2011; Overhage *et al*, 2007), where strong swarming-deficient phenotypes were observed for the Δ*fliM*, Δ*pilS*, Δ*rhlR* and Δ*rhlI* mutants. The mutant of a two-component response regulator gene *pirR* was found to repress swarming in multiple culture media and may explain the swarming deficiency of H27930 (Kollaran *et al*, 2019). The inability of S86968 to swarm may be attributed to the lack of *nfxB*, whose mutation caused global dysregulation of physiology and metabolism in *P. aeruginosa* including impaired swarming (Stickland *et al*, 2010).

Besides these genes, other metabolic 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 are fast-growing strains.** The missing genes in the 8 non-rhamnolipid-producers (Table 1) involve many transcriptional regulators (e.g., *rhlR*, *lasR*, *crc*, *nfxB*, *soxR*) that are 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 first tested how much growth can explain the two phenotypes. To address this question, we tracked the growth curves of all clinical isolates in the same glycerol media used to measure rhamnolipids production and monitored their population density changes over a time course of 48h. Visually, these growth curves show different patterns in the length of lag time and exponential growth-rates (Fig. 1C). Based on the Euclidean distance among the entire growth curves, we found that most of the swarming strains (except for PAO1, F30658 and F23197) clustered together (Supplementary Fig. S3) and the strains in this cluster are also strong swarmers with high swarming scores. However, rhamnolipids producers are mixed 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 may be needed.

We then used both unsupervised and supervised learning approaches to extract the growth features and tested their associations with rhamnolipid production. For the unsupervised approach, we used a non-negative matrix decomposition method (Lee & Seung, 1999) which approximates 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 overlap (Fig. 2A), there is 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 identify the exact interpretable feature with which rhamnolipid production is mostly associated, 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 1). Using Random Forest classification, we found that the top two features with the highest explanatory power are 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 reveals a strong temporal link between exponential growth rate in phase I and rhamnolipid production in phase II. We hypothesized that the link may be mediated by metabolic homeostasis as its disruption both slows down growth and disfavors overflow of excess carbon which would be preferentially used for cell maintenance and stress response under metabolic imbalanced conditions.

**TCA cycle and amino acid metabolism are perturbed in non-rhamnolipid-producers.** Similar to our growth curve analysis, we also 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 (see Methods), except for M55212 and F23197 which grew too slow, during the transition between phase I and phase II when rhamnolipids production begins. After data normalization and imputation (see Methods and Supplementary Fig. S5), hierarchical clustering shows the consistency of corrected data across all three replicates for each strain (Fig. 3), except for one replicate of H47921 (sample 25) which was removed from further analysis. The clustering of the metabolomics data yielded three major patterns with one pattern shared by all strong swarmers, supporting the link between swarming and intracellular metabolism (Boyle *et al*, 2017). However, no clear metabolic pattern was found among non-producers of rhamnolipids, despite four out of the six non-producers clustered in the same subgroup. The failure to identify associations between the whole metabolome and rhamnolipids production suggests that the associations may be detectable at specific metabolites carrying out key metabolic functions.

To identify such signature metabolites, we fitted our metabolomic data (explanatory variables) to categorized rhamnolipids production phenotype (response variables) using Orthogonal Projections to Latent Structures-Discriminant Analysis (OPLS-DA) (see Methods). The output model (t1) can separate the strains according to whether they produce rhamnolipids or not with good fit (R2 = 0.822, Q2 = 0.664, *p-*value = 5e-4) (Fig. 4A), explaining 5% of the variance, meaning that rhamnolipid producers and non-producer groups are heterogeneous (95% of the variance of the data is orthogonal to the predictor component), which explains why it is difficult to identify the compound using unsupervised method. To identify metabolic pathways that differentiate rhamnolipids producers from non-producers, we performed a metabolic pathway enrichment analysis using FELLA (Picart-Armada *et al*, 2018) with a univariate analysis, by mapping the significantly differentially produced metabolites to the *P. aeruginosa* type strain UCBPP-PA14 in the KEGG database. The entries with high connectivity with the provided compounds are returned by the algorithm [Vandin et al. 2011] (Supp. Tab. 1). We grouped the metabolites by the identified pathways and found that the most discriminative pathways were the TCA cycle and amino acid metabolism (Fig. 4B and Supplementary Fig. 6). The TCA cycle showed higher fumarate, malate but lower cis-aconitate, citrate, alpha-ketoglutarate, succinate in the rhamnolipid producers relative to the non-producers. It is worth mentioning that pyruvate level remains 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 non-rhamnolipid-producers relative to the producers. A striking exception to this trend is N-Formylmethionine (fMet), which had lower abundance in the non-producers. Diminished level of fMet was also found in the Δ*rhlA* mutant of *P. aeruginosa* (Supplementary Fig. 7)*.* Since the mutant had similar growth rate as the wild-type (Boyle *et al*, 2017), the association between fMet and rhamnolipid production is not a growth effect.

**Strains that do not produce rhamnolipids have increased oxidative stress.** The above comparative analyses of growth curve and metabolomics consistently suggest that the non-rhamnolipid-producers generally have higher oxidative stress levels and possibly elicited stress responses for a multitude of 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 explains their slower growth.
* The significantly opposite associations of succinate and fumarate with 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 can presumably decreases 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.
* Except for glutamate, proline and tryptophan, all other detected amino acids (threonine, alanine, aspartate, glutamine, cysteine, lysine, histidine, tyrosine, and arginine) were higher in the non-producers to various degrees. One possibility is that the non-producers attempted to recycle amino acids by extensive protein degradation against oxidative stresses.
* #there may be more. Keep digging#

**Rhamnolipid production requires redox homeostasis in simulations.** To verify our hypothesis that rhamnolipids are not secreted under oxidative stress conditions, we modified a high-quality genome-scale model of *Pseudomonas* metabolism, iJN1411 (Nogales *et al*, 2020, 2017), to simulate the maximum growth rate and rhamnolipid secretion potential under different redox stress levels (see Methods). The culture medium used for simulations is the same as the experimental setup with glycerol and ammonium as the sole carbon and nitrogen sources respectively. Since rhamnolipids are produced when carbon is in excess, we set the carbon-to-nitrogen (C:N) ratio to 10.0, which exceeds the minimum C:N threshold (6.3) that permits rhamnolipid secretion at even fastest growth rate (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), which 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) while any deviation from the range (gray shading) gradually reduces 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 are mandatory and provides growth benefit (i.e., each individual flux can be as low as zero without reducing maximum growth rate), suggesting that carbon release through different pathways are equivalent from a metabolic network perspective. Collectively, these simulation results confirm the link between growth, rhamnolipid secretion, and redox stress, where growth and rhamnolipid secretion are both metabolically constrained by cellular redox homeostasis.

**Validation experiments confirm 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 some rhamnolipids producers become non-producers due to stronger TCA cycle-mediated ROS generation which negatively regulates rhamnolipid production. Indeed, all the strains that produce rhamnolipids in glycerol became non-producers in succinate. We similarly extracted 7 phase-dependent growth curve features of our strains grown in succinate (Supplementary File 2), showing that theyhad faster exponential phase (phase I) growth rate, shorter delay before exponential phase commences, and lower maximum cell density compared to their growth in glycerol (Supplementary Fig. S10). More interestingly, the duration of phase II for these strains is much shorter in succinate than in glycerol (Supplementary Fig. S10), suggesting stronger stress response that shuts down growth and forces quick entry into stationary phase for cell maintenance.

In the second experiment, we monitored hydrogen peroxide (H2O2) level in the glycerol minimal medium during the growth of our isolates (Fig. 6). 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. The presence of cells from rhamnolipid producers (red and green lines) degraded H2O2 and reduced its environmental level; however, the non-producers (blue lines) polluted the environment with even more H2O2 than the amount they were able to degrade (Fig. 6A). Not surprisingly, the worst H2O2 degrader M1608 lacked *katE* and *gor,* 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 net release of H2O2 by non-producers in the lag phase (Fig. 6B). However, they become as capable of degrading H2O2 as the producers in the exponential phase, suggesting their antioxidant systems have slower adaptive responses to the new culture medium and the immediate oxidative shock caused by oxidizing the medium components.

Fig. 7 summarizes our major finding and further proposes a molecular mechanism that explains why some clinical isolates produce rhamnolipids and swarm while the others do not. 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 and increased oxidative stress. 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 were inhibited in response to oxidative stress, which consequently abolishes rhamnolipid production and swarming. 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 the negative regulation of rhamnolipid production by oxidative stress is mediated by the Lon protease, which is upregulated by oxidative stress and inhibits quorum-sensing responses.

**Discussion**

In this study, we showed diversity of swarming and rhamnolipid production across clinical isolates of *P. aeruginosa* and provided mechanistic explanations that underlie the observed diversity. The diversity is partially attributed to the loss of genes essential to both phenotypes, such as those encoding flagellar proteins and quorum-sensing regulators (Table 1). Compared to rhamnolipid producers, the non-producers generally have slower growth and perturbed TCA cycle and amino acid metabolism. We demonstrate that these metabolic perturbations in the non-producers are consequences of disrupted redox homeostasis, which are possibly caused by their genetic deficiency in detoxifying ROS. Our findings thus suggest that the differential redox states across all *P. aeruginosa* strains are the fundamental metabolic factors associated with the observed diversity. Since our clinical strains were isolated from hospitalized patients, they are expected to experience a redox stress environment because ROS are commonly used by the immune system such as the phagocytes as weapons to fight against pathogens (Puertollano *et al*, 2011). We hypothesized that these non-producers of rhamnolipids incapable of dealing with oxidative stress quickly lost their abilities to produce rhamnolipids during adaptative evolution in patients. This hypothesis is consistent with our phylogenetic analysis, showing that the phenotype remained relatively stable across the phylogenetic tree but only evolved recently. Our genetic analysis shows that genes in the rhamnolipid biosynthesis pathway (*rhlA*, *rhlB*, *rhlC*) are highly conserved throughout all of our clinical isolations even in non-producers, suggesting that the evolutionary loss of rhamnolipid production capability is not for its own sake but a consequence of broader metabolic adaptations such as oxidative stress responses.

Our study revealed a strong and growth-independent association between fMet abundance and rhamnolipid production in *P. aeruginosa*. The non-rhamnolipid-producing strains showed significantly lower fMet levels compared to the producing isolates, and so does the Δ*rhlA* mutant compared to its parental strain PA14. fMet is known for its key role in translation initiation as it is the first residue of all nascent peptides synthesized by bacterial ribosomes. The formyl-group on the initiating methionine residue is co-translationally removed by the peptide deformylase and this step is critical for cleavage of the terminal methionine from the nascent protein by Met-aminopeptidase in a later stage and, more importantly, the protein’s stability and function. Besides, the terminal fMet acts as degradation signals for protein quality control such that, during peptide elongation, the misfolded proteins without timely removal of fMet can be degraded (Piatkov *et al*, 2015). It was reported that *Bacillus subtills* became sensitive to H2O2 and defective for swarming when lacking Formyl-Methionine Transferase (FMT)—the enzyme attaching a formyl group to methionine loaded on tRNAfmet (Cai *et al*, 2017). Another relevant finding has shown that methionine can be mis-incorporated into non-methionine residues of proteins under oxidative stress and protect mammalian cells from ROS inactivation by reacting with ROS molecules first before they attack sensitive methionine residues whose oxidation leads to protein functional loss (Lee *et al*, 2014). Therefore, one possibility of lower fMet levels in non-rhamnolipid producers is that a higher proportion of the methionine pool in the non-producers was strategically posited at non-methionine residues of proteins as a protective mechanism against oxidative stress, leaving less methionine to be formylated by FMT. But, in general, the mechanistic links among fMet, swarming/rhamnolipid production, and oxidative stress remains unclear and warrants further research.

We demonstrated in this study that rhamnolipid producers are generally fast-growing strains and it was previously shown that rhamnolipids are not produced unless carbon is in excess (Boyle *et al*, 2015). Therefore, rhamnolipid production shares similar characteristics with overflow metabolism, a mechanism that is 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 of acetate 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. These mechanisms are generally based on cost-benefit analysis, despite the cost associated with large respiratory flux can be incurred under different types of resource limitations (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 *Escherichia coli* (Vemuri *et al*, 2006) and *Saccharomyces 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. Still, there is much to be understood about the relationship between *P. aeruginosa* rhamnolipid production and overflow metabolism observed in other cell types. One apparent difference is that rhamnolipids biosynthesis is controlled by quorum-sensing in addition to nutrient cues, which result in cell-density-dependence of rhamnolipid production. Indeed, *P. aeruginosa* produces rhamnolipids when cells exit the fast-growing exponential phase and enter phase II, during which the cells continue to grow but at a lower rate (Boyle *et al*, 2015).

Does rhamnolipid production provide fitness advantage for *P. aeruginosa* cells 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 (rhamnolipid defector) had almost the same growth rate (van Ditmarsch & Xavier, 2011) and very similar metabolomics (Boyle *et al*, 2017) with its wild-type parent strain in the liquid culture, suggesting that the carbon used for rhamnolipids 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). Still, the defector 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), we suspect that the defector may exhibit slight oxidative burden, which also explains the outcome of swarming competition in the agar plate where the mutant was disfavored in mixture with the wild-type even though the former can free ride on the rhamnolipids produced by the latter (de Vargas Roditi *et al*, 2013). The 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. The fitness benefit of cooperative production of rhamnolipids can also be consequence of population-level selection since the biosurfactant secretion enables a swarming colony to spread the agar plate for new territories and nutrients that sustain cell proliferation. Although the defectors can free-ride on the rhamnolipids produced by the wild-type, they do not have as much spatial and temporal access to the rhamnolipids as the wild-type since rhamnolipids are first produced within the neighborhood of the wild-type cells. It was shown that even marginal advantage of the wild-type strain in the ability to expand colony can be exaggerated in space and lead to discernable effect on its final proportion (Liao *et al*, 2017). Collectively, the higher fitness of the rhamnolipid-mediated cooperation is likely an example of multilevel selection that has been previously suggested (de Vargas Roditi *et al*, 2013).

Notably, the fitness benefits associated with rhamnolipid production would vanish without integrating nutrient-sensing and quorum-sensing signals. Indeed, constitutive rhamnolipid producers compromised growth and were disfavored in the swarming competition with the defectors (de Vargas Roditi *et al*, 2013). Secondary metabolism such as rhamnolipid production typically occurs in bacteria during stationary phase when nutrients become limited. We previously demonstrated that nitrogen is the limiting nutrient during the transition from phase I to phase II (Boyle *et al*, 2015). Nitrogen limitation especially is ubiquitous in the microbial world and high carbon-to-nitrogen ratio was found in the intestine of 30 mammals including humans and it shapes microbiome composition (Reese *et al*, 2018). Under nitrogen limitation, quorum-sensing system further allows synchronization and amplifications of the nutrient cues among populations. Therefore, the integrated signals provide a temporally prudent control of costly metabolic byproduct secretion: *P. aeruginosa* cells prioritize resource investment to growth in nutrient-rich medium, while they divert excess carbon to rhamnolipids secretion as swarming public goods when nitrogen becomes limited and sufficient cells are present. In this paper, we extended the concept of metabolic prudence proposed a decade ago (Xavier *et al*, 2011): the secretion of rhamnolipids was controlled not only by carbon availability and cell density, but also by redox homeostasis. ROS are reactive intermediates that damage macromolecules, including DNA, RNA, lipids and proteins. Particularly, the cysteine and methionine residues in proteins are susceptible to ROS modifications which alter protein structure and function. For example, *E. coli* cells subject to H2O2 challenge are limited in methionine biosynthesis due to oxidative inactivation of MetE—the enzyme catalyzing the final step of *de novo* methionine biosynthesis. The active shutdown of rhamnolipids production in redox imbalanced condition ensures that precious carbon and enzyme resources were not wasted but diverted to oxidative stress response mechanisms.

Finally, our study sheds broad light on how metabolite secretion impacts microbial physiology, ecology and evolutionary strategies. Microorganisms secrete many different kinds of metabolic byproducts, including nutrients, extracellular enzymes, toxins and cell-cell signaling molecules (Schmidt *et al*, 2019). In a multispecies community, the byproducts produced by the member species can impact each other and drive polymicrobial interactions. One such interaction is cross-feeding, where the metabolic byproducts of one cell type are nutrients that can be utilized by other cell types. Different from cross-feeding, rhamnolipids are not consumable but enable bacteria to swarm over surfaces. It has been experimentally shown that rhamnolipids from *P. aeruginosa* can help a sulfate-reducing bacteria, *Desulfovibrio vulgaris*, to disperse (Wood *et al*, 2018). Owing to its metabolic versatility, *P. aeruginosa* can also synthesize and accumulate large amount of polyhydroxyalkanoid acids (PHAs), and previous work has showed that rhamnolipids and PHAs can be simultaneously produced (Hori *et al*, 2002). Unlike secretable rhamnolipids which are shared by the entire population as swarming public goods, PHA molecules are reservoirs of carbon and energy sources and stored inside of cells to cope with changing environmental conditions especially carbon starvation (Pham *et al*, 2004). The biosynthesis of PHAs is closely linked to rhamnolipids production via competition for the common precursor; therefore, the alternative survival strategies between rhamnolipids and PHA production allow *P. aeruginosa* cells to quickly adapt to new environmental stress with minimal metabolic rewiring.

**Materials and Methods**

**Rhamnolipid production determination.** The production of rhamnolipids was assessed by drop-collapse assay. 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 (Jain & Collins; Chen *et al*, 2007).

**Growth curve assay.** The clinical isolates were inoculated in 3 mL of LB and incubated 37ºC overnight with shaking. 500 μL of cell culture was pelleted and washed 3 times with PBS. 0.0025 OD600 units were inoculated into glycerol minimal medium in BD Falcon (BD Biosciences, San Jose, CA) 96 well flat-bottom plates, with 150 μL of suspension per well. The plate was incubated during 48 hours at 37ºC in a Tecan Infinite M1000 or Tecan Infinite M1000 Pro plate reader (Männedorf, Switzerland), with an orbital shaking of 4 mm of amplitude. OD600 was measured in 10 minutes intervals.

The clinical isolates along with were grown in the same synthetic medium of metabolomics extraction (with glycerol as sole carbon source) at 37ºC for 48 h (initial OD600 was 0.0025). The growths curves were determined by measuring OD600 each 10 minutes.

**Swarming assay** Swarming assays were performed as described previously (Xavier *et al*, 2011; van Ditmarsch *et al*, 2013). The clinical isolates were inoculated in 3 mL of LB and incubated at 37ºC overnight, with shaking. 500 μL of the culture was pelleted and washed twice with PBS. 2 μL of this suspension was spotted on the surface of casamino acids soft agar plates, without penetrating the agar with the pipette tip. The plates were incubated at 37ºC ~24 hours. Two replicates were done per strain. In each batch of swarming assays PA14 was used as control.

**Swarming score** Swarming score () for each clinical isolate is defined as a linear combination of the maximum length () and circularity () of its swarming colony

where and are the mean values of and across all clinical isolates respectively.

**Imaging** Images of the swarming plates were obtained with a Chemidoc gel doc imager (Bio Rad).

**Image analysis and determination of swarming score** 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 and eccentricity (the eccentricity of the fitted ellipse).

The analysis of the morphological features was performed in R. The average of the replicates for each feature was obtained. The values of each feature in each experiment were normalized to the value of PA14 control in that experiment. A PCA biplot was generated and maximum length and circularity were found to be the features that most spreaded swarmers and non swarmers. These two features were the one used for obtaining the swarming score (Equation 1). The coefficients in the equation correspond to the rotation of the two features that maximizes the variance of swarming score.

**Metabolic extraction and metabolomic data preprocessing** After extraction and profiling using liquid-chromatography coupled to mass spectrometry (LC-MS), we identified a total of 99 compounds (**Fig. 3A**), and from these we chose 70 compounds that passed several criteria for further analysis (see Methods).

The 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 (including pyruvate, XX, XX), therefore after the normalization step they were removed (in black in Fig. 2A).

**Hierarchical Clustering Analysis of metabolomic data.** The Hierarchical Clustering Analysis of the normalized metabolomic data was performed using gplots R package (Warnes *et al.* 2015), with Euclidean distance and Ward’s aggregation method (Ward 1963). 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 2B, 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, with a significance level of 0.05. These compounds were fed to FELLA algorithm (Picart). The graph over which FELLA was performed consisted of all entries in the KEGG database for *Pseudomonas aeruginosa* strain UCBPP-PA14. We filtered the results table 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.

**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

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

Bru J-L, Rawson B, Trinh C, Whiteson K, Høyland-Kroghsbo NM & Siryaporn A (2019) PQS Produced by the Pseudomonas aeruginosa Stress Response Repels Swarms Away from Bacteriophage and Antibiotics. *J. Bacteriol.* **201:**

Cai Y, Chandrangsu P, Gaballa A & Helmann JD (2017) Lack of formylated methionyl-tRNA has pleiotropic effects on Bacillus subtilis. *Microbiology (Reading, Engl.)* **163:** 185–196

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

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

Hori K, Marsudi S & Unno H (2002) Simultaneous production of polyhydroxyalkanoates and rhamnolipids by Pseudomonas aeruginosa. *Biotechnol. Bioeng.* **78:** 699–707

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

Lee JY, Kim DG, Kim B-G, Yang WS, Hong J, Kang T, Oh YS, Kim KR, Han BW, Hwang BJ, Kang BS, Kang M-S, Kim M-H, Kwon NH & Kim S (2014) Promiscuous methionyl-tRNA synthetase mediates adaptive mistranslation to protect cells against oxidative stress. *J. Cell Sci.* **127:** 4234–4245

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

Liao C, Blanchard AE & Lu T (2017) An integrative circuit-host modelling framework for predicting synthetic gene network behaviours. *Nat. Microbiol.* **2:** 1658–1666

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*

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, Bains M, Brazas MD & Hancock REW (2008) Swarming of Pseudomonas aeruginosa is a complex adaptation leading to increased production of virulence factors and antibiotic resistance. *J. Bacteriol.* **190:** 2671–2679

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

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

Pham TH, Webb JS & Rehm BHA (2004) The role of polyhydroxyalkanoate biosynthesis by Pseudomonas aeruginosa in rhamnolipid and alginate production as well as stress tolerance and biofilm formation. *Microbiology (Reading, Engl.)* **150:** 3405–3413

Piatkov KI, Vu TTM, Hwang C-S & Varshavsky A (2015) Formyl-methionine as a degradation signal at the N-termini of bacterial proteins. *Microb. Cell* **2:** 376–393

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

Reese AT, Pereira FC, Schintlmeister A, Berry D, Wagner M, Hale LP, Wu A, Jiang S, Durand HK, Zhou X, Premont RT, Diehl AM, O’Connell TM, Alberts SC, Kartzinel TR, Pringle RM, Dunn RR, Wright JP & David LA (2018) Microbial nitrogen limitation in the mammalian large intestine. *Nat. Microbiol.* **3:** 1441–1450

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

Wang Q, Frye JG, McClelland M & Harshey RM (2004) Gene expression patterns during swarming in Salmonella typhimurium: genes specific to surface growth and putative new motility and pathogenicity genes. *Mol. Microbiol.* **52:** 169–187

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

Wood TL, Gong T, Zhu L, Miller J, Miller DS, Yin B & Wood TK (2018) Rhamnolipids from Pseudomonas aeruginosa disperse the biofilms of sulfate-reducing bacteria. *npj Biofilms and Microbiomes* **4:** 22

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 are represented by the actual swarming images on the agar plates (left column). The ability of producing rhamnolipids of these strains are indicated by circle sizes (right column). Both phenotypes do not correlate with phylogeny. All swarmers can secrete rhamnolipids but some rhamnolipids producers cannot swarm. **C**. Population density changes of these clinical strains grown in a synthetic minimal medium using glycerol as the sole carbon source. The swarmers tend to have a shorter lag phase, suggesting that the link between metabolic capability and swarming which we observed in the PA14 background (Boyle *et al*, 2017) may be conserved across the phylogenetic tree.

A screenshot of a cell phone

Description automatically generated

**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 rhamnolipids 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

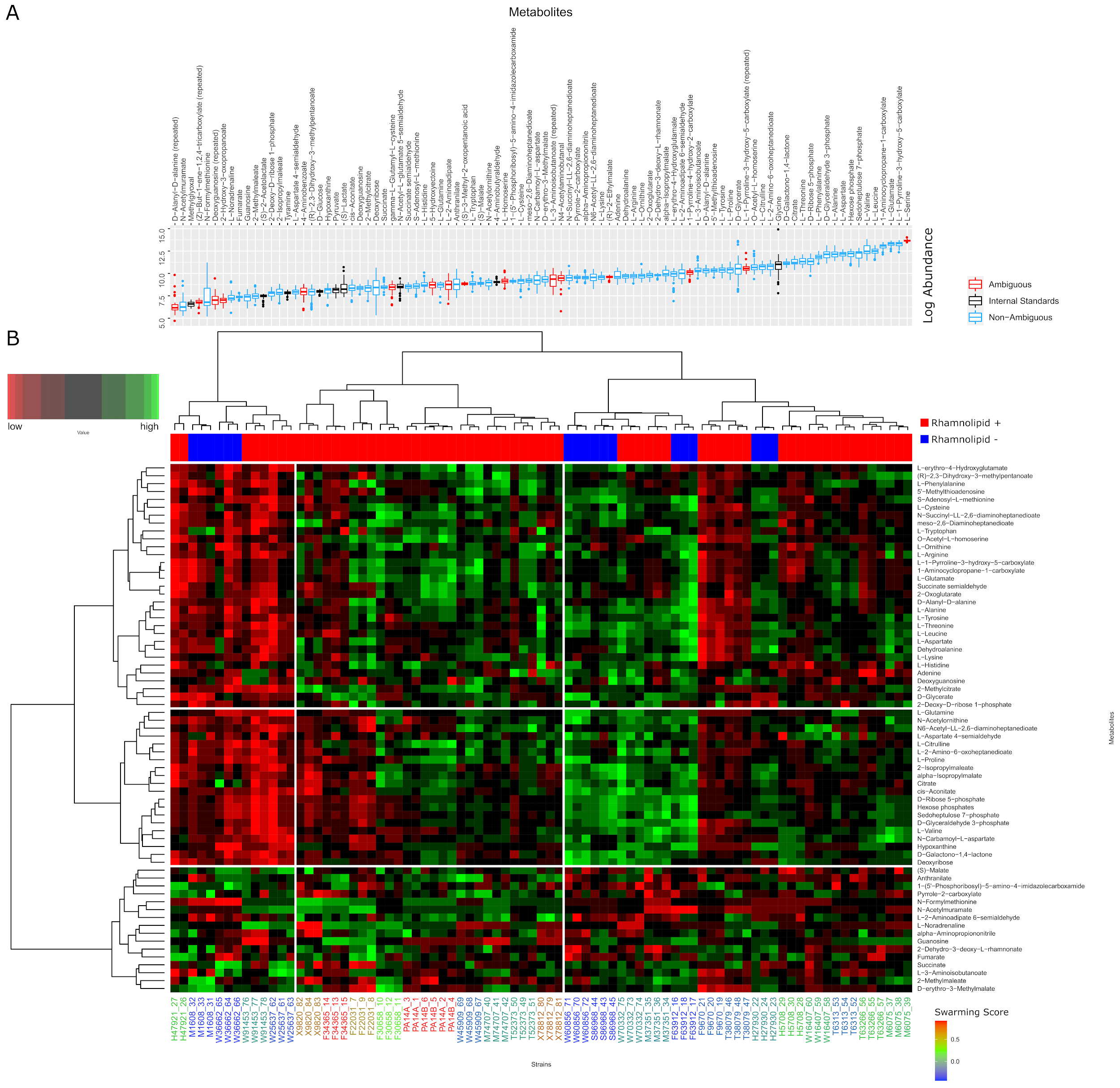
Description automatically generated

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

A close up of a map

Description automatically generated

**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.

A close up of a map

Description automatically generated

**Figure 4. Differential metabolomics between rhamnolipid producers and non-producers.** (**A)** Scores of 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.

A screenshot of a cell phone

Description automatically generated

**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.

A close up of a logo

Description automatically generated

**Figure 7**. Metabolic model of rhamnolipids production in *P. aeruginosa*. A) During aerobic growth, cells rely on the TCA cycle to generate energy molecules such as NADH and invest in biomass. B) When cells reach a certain density and nitrogen becomes limited, rhamnolipids production is turned on by quorum sensing and this step consumes NADH. Enzymes that facilitate cell growth at this condition are synthesized and some proteins are recycled with fMet as a degradation signal. At the same time, TCA cycle is slowed down by redox stress. For the strains that could produce rhamnolipids, the membrane redox stress is partially released and therefore the flux between succinate and fumarate is less reduced (\*), resulting in higher fumarate level than in rhamnolipids nonproducer. red dots: metabolites whose levels are lower in rhamnolipids producers; green dots: metabolites whose levels are higher in rhamnolipids producers.