**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 different patients differ in their abilities to swarm and produce rhamnolipid—a necessary but not sufficient requirement for swarming. Despite the diversity of swarming can be partially explained by the absence of quorum-sensing and flagellar genes, the differentially expressed metabolites between rhamnolipid producing and non-producing strains suggest redox homeostasis as a more fundamental metabolic constraint. The increased oxidative stress in non-producers inhibited Fe-S containing enzymes such as the succinate dehydrogenase and resulted in reduced tricarboxylic acid (TCA) cycle activity and slower growth rate. Genome-scale model simulations and H2O2 detoxication experiments validate our theory on the links between growth, redox homeostasis and carbon overflow. Although the ability to direct excess carbon to the production of rhamnolipids is metabolically constrained at the individual level, it may be key to the evolution of *P. aeruginosa* swarming at the population level as rhamnolipids are public goods that help stabilize microbial cooperation. Our study also adds to the expanding literature on the observation that fast-growing bacteria enter overflow metabolism,

# add rhamnolipids production are prudently controlled

**Significance statement**

**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 and 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 and 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 and 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 and 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 metabolic signals such as the presence of excess carbon in relative to metabolic capacity (Xavier et al. 2011; Boyle et al. 2015; Mellbye and 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 and Morowitz 2004) and a major determinant of social behavior such as swarming (Biro and 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, 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; Yan et al. 2019), which vary widely in their ability to swarm and to produce rhamnolipids. 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 at both phenotypic and genetic levels. 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 quorum-sensing and flegellar genes in the accessory genomes. Using metabolomics and growth curve data acquired in the liquid culture with glycerol as the sole carbon source, we show more fundamental characteristics associated with the non-producers of rhamnolipids such as reduced tricarboxylic acid (TCA) cycle activity and slower growth, which are presumably caused by accumulated reactive oxygen species (ROS) and increased oxidative stress. With mathematical modeling and independent validation experiments, we provide strong evidences that rhamnolipid production is constrained by redox homeostasis with rhamnolipid producers more capable of detoxifying 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, with a half of strains unable to swarm and 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 fully 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 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 and 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 and 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 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 similarity of the entire curves, we found that most of the swarming strains (except for PAO1, F30658 and F23197) clustered together and the strains in the cluster are also strong swarmers with high swarming scores (Supplementary Fig. S3). 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 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 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) visually 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. SX) and defined 7 quantitative features to characterize each growth phase (Fig. 2D). 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 1 (Fig. 2E). Considering the majority of rhamnolipids are produced in phase 2 when an essential non-carbon nutrient becomes limited (Boyle et al. 2015), our finding reveals a strong link across the temporal scale between phase 1 growth rate and phase 2 rhamnolipid production. We hypothesized that the link may be mediated by metabolic homeostasis as its disruption slows down growth and requires more maintenance energy from glycerol as the sole carbon and energy source, thus disfavoring its overflow in forms of rhamnolipids.

**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 (except for M55212 and F23197 because they grew too slow) during the transition between phase 1 and phase 2 when rhamnolipids production begins (see Methods). Hierarchical clustering shows the data consistency across all three replicates for each strain (Fig. 3B), except for one replicate of strain H47921 (sample 25) which was removed from further analysis. The clustering of the metabolomics data yielded three major patterns with one pattern common to 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 association may be founded by specific metabolites carrying out key metabolic functions.

To identify such metabolites, we fitted our metabolomic data to rhamnolipids production phenotype 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-producers 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 compounds to *P. aeruginosa* strain UCBPP-PA14 in 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 (Figure 4B, Figure S4) and found that the most perturbed pathways were related with the central TCA cycle and amino acid metabolism (Fig. 4B and Supplementary Fig. S4). Pyruvate was found to be constant across all the strains and therefore was removed in our analysis. The fact that pyruvate is not altered indicates that the regulation happens in the TCA cycle on itself, not in its upstream of central carbon metabolism. Interestingly, the TCA cycle is divided into two parts, the metabolites in one part are lower in rhamnolipid producers while metabolites in the other part, namely fumarate and malate are higher in rhamnolipids producers. Three out of 7 compounds in this pathway were significantly different between rhamnolipid producers and non producers (fumarate, succinate and citrate). The reaction where the association reverses is the production of fumarate from succinate, the only reaction in the TCA cycle that happens in the cell membrane and couples FADH2 synthesis. Although the level of metabolites does not represent the flux of metabolism, the flip of sign in association tells us that it is highly possible that the reaction rate between succinate and fumarate is different between rhamnolipids producers and nonproducers.

Another tightly associated pathway is amino acids biosynthesis, especially branched chain and sulfur containing amino acids, where all metabolites are in high abundance in rhamnolipids non-producers. This could be linked to an altered proteome turnover. An interesting exception to this is N-Formylmethionine, which was found in high abundance in rhamnolipid producer strains. This could be again associated with protein synthesis or degradation. N-Formylmethionine is the starting residue in bacterial translation. The formyl group is added to methionine by methionyl-tRNA formyltransferase after Met-tRNA is formed and removed after translational initiation. Also, fMet could serve as degradation signal in bacteria (ref).

**Rhamnolipids non-producers may have increased oxidative stress**

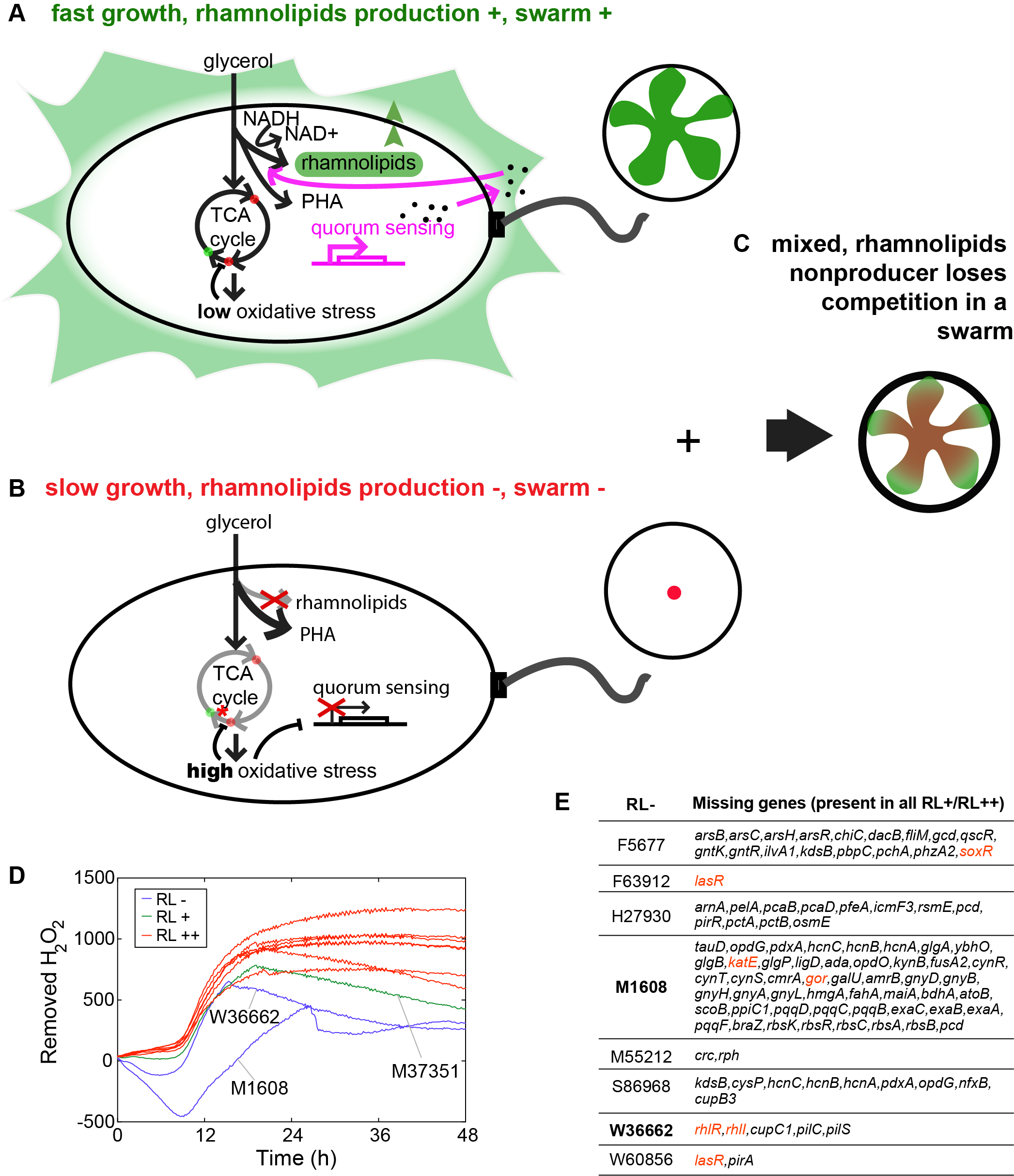
Combining results from growth curve analysis and comparative metabolomics, we found that rhamnolipids producers can grow faster, contain lower amino acids level, and possibly sustain a higher rate of TCA cycle activity compared to non-producers. It is not surprising to see that a faster TCA cycle is linked to faster cell growth. However, this coupling is fragile and hard to maintain: TCA cycle is a main source of Reactive Oxygen Species (ROS) (Mailloux et al. 2016). The five enzymes catalyze the reactions in TCA cycle all contain Fe-S clusters (aconitase A, aconitase B, succinate dehydrogenase subunit B, fumarase A, fumarase C) which makes them prone to ROS stress (Flint et al. 1993). Among the five enzyems, succinate dehydrogenase subunit B (adhB) encodes one component of the complex that catalyzes the reaction between succinate and fumarate, the two compounds that are associated with rhamnolipids non-producers and producers respectively. (Fig. 3C). The opposite correlation of fumarate and succinate loading values indicates that succinate dehydrogenase activity is greatly reduced in rhamnolipid non0producers but not producers. Succinate dehydrogenase B locates in the membrane, which makes it specially prone to suffer from oxidative stress (Stark 2005). These observations suggest that the rhamnolipids producers have lower oxidative stress and therefore maintain higher flux in TCA cycle to sustain faster cell growth (Vemuri et al. 2006).

**Rhamnolipid production requires redox homeostasis in simulations**

To verify our hypothesis that rhamnolipids are not secreted under oxidative stress conditions, we used a genome-scale model of *Pseudomonas* metabolism to simulate the maximum growth rate and rhamnolipid secretion potential under different redox stress levels. 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 while allowing fastest growth rate (Supp. Fig. X). 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 and likely to be the main sources of ROS (ref). 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 (upper panels of A-C, Fig. X). 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 (lower panels of A-C, Fig. X). Importantly, none of these secretion fluxes are mandatory and provides growth benefit (i.e., each individual flux can be as low as zero without compromising 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 status.

**Clinical isolates stop producing rhamnolipids when grown in succinate as sole carbon source**

We monitored the growth curve of the clinical isolates and the production of rhamnolipids in a synthetic medium with succinate as the sole carbon source in an analog way than the one exposed before for glycerol. Interestingly, all the strains that produce rhamnolipids when grown in glycerol become non-producers after growth in succinate. The growth curves showed that the isolates transited directly from phase 1 (exponential) to phase 3 (decay). Rhamnolipid production occurs in phase 2 (stationary). *P. aeruginosa* grows faster in succinate than in glycerol, and its lag phase is shorter [Dolan *et al.* 2020 DOI: [10.1128/mBio.02684-19](https://doi.org/10.1128/mbio.02684-19)]. Succinate enters the carbon metabolism directly through the TCA cycle. This faster growth might be producing an increased concentration of NAD(P)H that would cause the TCA cycle to stop. As succinate cannot be redirected to overflow metabolism neither rhamnolipid production cells transite from phase 1 to 3, as NAD(P)H concentration becomes toxic and cells stop sustaining the growth, producing a decay in the growth rate. Rhamnolipids therefore are not produced. The fact that in rhamnolipid non-producers grown in glycerol, despite the metabolomic data suggesting that TCA cycle is blocked during late phase 1 of growth, a sustained growth is observed and phase 2 is present, suggests that non-producer strains might be obtaining their energy from glycerol using a fermentative pathway. Rhamnolipid production might be a way of keeping the TCA cycle active during phases 1 and 2 of growth.



**Figure 6**. 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.

**Discussion**

Rhamnolipids production is not activated unless there is extra carbon, which resembles overflow metabolism, a mechanism that is generally associated with fast growth of bacteria when there is a high level of glucose as the carbon source.

(Szenk et al. 2017)

Add resemblance to carbon overflow metabolism

In order to keep fast growth, many cells utilize a strategy called overflow metabolism, to redirect their carbon flow and secrete carbon containing molecules such as acetate and ethanol. This phenomenon has been observed repeatedly in *Escherichia coli*, *Saccharomyces cerevisiae* (Crabtree effect (De Deken 1966)) and cancer cells (Warburg effect) when theygrow in high concentration of glucose,, Farmer and Jones 1976 <https://doi.org/10.1111/j.1432-1033.1976.tb10639.x>)(Vander Heiden *et al.* 2009 doi:10.1126/science.1160809). . *P. aeruginosa* produces rhamnolipids when the cells exit the fast growing exponential phase and enter phase II, during which the cells are able to continue to grow even at a lower rate (Xavier et al. 2011, Boyle et al. 2015). This process shares similarities with metabolism overflow where the secretion of carbon molecules is coupled with cell growth when carbon becomes excess . In *E. coli*, overexpression of the NADH oxidase increases the flux of glucose and releases the repression of TCA cycle, (Vemuri *et al.* 2006 doi: 10.1128/AEM.72.5.3653-3661.2006.) indicating that NAD(P)H/NAD(P) homeostasis is important for the overflow of carbon. Would the redox balance also be important for rhamnolipids secretion in *P. aeruginosa*? .

Secondary metabolism such as rhamnolipids production is usually associated with environmental stress such as nutrient limitation. Nitrogen limitation especially is ubiquitous in the microbial world. High C:N ratio is found in the intestine of 30 mammals including humans and it shapes microbiome composition (https://www.nature.com/articles/s41564-018-0267-7). Therefore understanding bacterial metabolism in nutrient imbalance and how it shapes cell-cell interaction is crucial to understand the emerging of social traits. *P. aeruginosa* uses metabolic prudence to prioritize the nutrient investment to biomass production when both nitrogen and carbon are abundant. When nitrogen becomes limited and carbon remains excessive, *P. aeruginosa* secretes rhamnolipids as public goods that enable the population to move cooperatively as a swarm and explore new territories. How cellular metabolic networks regulate individual cell growth and social cooperation, two traits that are usually considered tradeoffs through evolution? How to prevent cheating even in the individual level? How cellular metabolism impacts structured bacterial communities? Our phylogenetic analysis shows that the ancestor of *P. aeruginosa* strains is a rhamnolipids producer and this phenotype remains relatively stable across the phylogenetic tree. Gene sequence analysis confirms that *rhlA* is highly conserved in all of our clinical isolations even in rhamnolipids non producers, further supporting that this pathway, although not essential, is under stringent selection along evolution. In this study, our systematic analysis reveals that rhamnolipids production couples individual cell growth to social cooperative traits.

Previous studies showed that rhamnolipids production is strictly regulated by quorum sensing signal (N-butanoyl-L-homoserine lactone [C4-HSL] binding to RhlR). Here in our clinical isolates, we found that a few rhamnolipids non-producers indeed lose quorum sensing regulators and quorum sensing synthesis enzyme. However, the losing of quorum sensing genes could not explain all the non-producing phenotype: other non-producers lose genes involving central carbon catabolism or energy producing pathways *etc.,* indicating a link between rhamnolipids, carbon metabolism and energy processing. The expression of *rhlAB* operon is not constitutive but activated when cells switch from a fast growing stage into a slow growing stage in the presence of excessive carbon (doi.org/10.1371/journal.pcbi.1004279). The tight temporal regulation prevented the invasion of cheaters in a swarm even though they can free ride rhamnolipids to expand (doi.org/10.1111/j.1365-2958.2010.07436.x). Deleting *rhlA* or C4-HSL did not increase the growth rate of *P. aeruginosa* in liquid glycerol media (doi.org/10.1186/1471-2180-11-140), indicating that the carbon used for rhamnolipids production does not interfere with biomass production. Would this tight coupling of growth and rhamnolipids production impose any advantage? In a swarm, when cell reaches certain density and nitrogen becomes scarce but not completely deprived, the extra carbon is used to support cell growth in a lower rate as well as to produce rhamnolipids that can break the surface tension and enable the population to expand to places with more nutrients that enable cell proliferation. We hypothesize that *P. aeruginosa* rely on redox homeostasis to achieve this beautiful coordination of cellular growth and social translocation.On the other hand, if cells fail to lower the ratio between NAD(P)H and NAD(P),high concentrations of NAD(P)H can act as a source of ROS: they are toxic to the cell and inhibit cell growth (Murphy 2008 doi: [10.1042/BJ20081386](https://dx.doi.org/10.1042%2FBJ20081386), Esterházy *et al*. 2008 doi: 10.1021/bi702243b. Vinogradov and Grivennikova 2016<https://doi.org/10.1016/j.bbabio.2015.11.004>, Larosa and Remacle 2018 doi: [10.1042/BSR20171492](https://dx.doi.org/10.1042%2FBSR20171492)).

How could this mechanism be implemented? Why don’t we see cells keep growing at a high growth rate? The capacity of TCA cycle turns out to be limited by either cell membrane capacity for electron transport chain complexes needed for NADH oxidation (Szenk *et al.* 2017 doi: 10.1016/j.cels.2017.06.005., Andersen and Meyenburg 1980 DOI: [10.1128/JB.144.1.114-123.1980](https://www.researchgate.net/deref/http%3A%2F%2Fdx.doi.org%2F10.1128%2FJB.144.1.114-123.1980?_sg%5B0%5D=4a4VfWx217JL1ZWFRFKETzqKft_yGEr0HnJglpQUxqdqy-tCbo_7etRr3MNCbtgVdzj1kya18LwxnoXMiznzl8VdsA.rKKt_pzBl2pXsLGsmCantLZCsJWeKQQaZceq6ymNy5TXVYGQZ-qmR7SrufEVgbwJJlFpAqhhRkgP_Hjs-VKv7g)), or by proteome efficiency during respiration (Basan *et al.* 2015 doi:10.1038/nature15765.)

We use metabolic flux analysis to demonstrate that high NAD(P)H and GSH flux coupling with high C:N ratio can indeed lead to production of HAA, mono-rhamnolipid, acetate and lactate etc. Another association between rhamnolipids production and NADH flux is RhlG, an enzyme upstream of RhlA that catalyzes the synthesis of β-hydroxy acid moiety of rhamnolipids. RhlG functions as an NADPH-dependent β-ketoacyl-acyl carrier protein (ACP) reductase (DOI: 10.1128/JB.180.17.4442-4451.1998), indicating the the restoration of redox homeostasis could happen before *rhlAB* is activated

Why is it important to secrete the extra metabolites outside the cells? There can be reasons such as detoxification, recycling of nutrients for diauxie, or help with selection for social traits. Another well known polymer produced by *P. aeruginosa* is polyhydroxyalkanoate (PHA, mainly polyhydroxybutyrate [PHB]), which shares common precursors with rhamnolipids (ref) and can also be induced by nitrogen starvation. Unlike rhamnolipids that are secreted outside of bacterial cells and shared by the entire population as ‘public goods’, PHA molecules are stored inside cells as granules and kept as ‘private’ resources that can increase stress tolerance of *P. aeruginosa* cells (DOI 10.1099/mic.0.27357-0). This strategy, although would increase cell resistance to environmental stress, does not contribute to population expansion. The interplay between intracellular and intracellular polymer production would greatly facilitate cell adaptation through evolution.

Overall, our analysis, for the first time, reveals that metabolic constraints impose both individual advantage (growth rate) and social advantage (swarming) for *P. aeruginosa* and rhamnolipids production greatly facilitate this regulation, which may lead to swarming diversity.

**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 and Collins n.d.; 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.

**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.

**Intracellular Metabolites extraction**

**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).

Missing values, if any of the replicates had a non-missing value for the particular compound, were imputed to 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 doi: 10.1021/ac901143w) with NormalizeMets R package (De Livera *et al*. 2018 https://doi.org/10.1007/s11306-018-1347-7). 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, 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-Armada 2018 doi: 10.1186/s12859-018-2487-5]. 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.

**References**

Abdel-Mawgoud, A.M., Lépine, F. and Déziel, E. 2010. Rhamnolipids: diversity of structures, microbial origins and roles. *Applied Microbiology and Biotechnology* 86(5), pp. 1323–1336.

Biro, P.A. and Stamps, J.A. 2010. Do consistent individual differences in metabolic rate promote consistent individual differences in behavior? *Trends in Ecology & Evolution* 25(11), pp. 653–659.

Boyle, K.E., Monaco, H., van Ditmarsch, D., Deforet, M. and Xavier, J.B. 2015. Integration of metabolic and quorum sensing signals governing the decision to cooperate in a bacterial social trait. *PLoS Computational Biology* 11(5), p. e1004279.

Boyle, K.E., Monaco, H.T., Deforet, M., et al. 2017. Metabolism and the evolution of social behavior. *Molecular Biology and Evolution* 34(9), pp. 2367–2379.

Bru, J.-L., Rawson, B., Trinh, C., Whiteson, K., Høyland-Kroghsbo, N.M. and Siryaporn, A. 2019. PQS Produced by the Pseudomonas aeruginosa Stress Response Repels Swarms Away from Bacteriophage and Antibiotics. *Journal of Bacteriology* 201(23).

Chen, C.Y., Baker, S.C. and Darton, R.C. 2007. The application of a high throughput analysis method for the screening of potential biosurfactants from natural sources. *Journal of Microbiological Methods*.

Chong, H. and Li, Q. 2017. Microbial production of rhamnolipids: opportunities, challenges and strategies. *Microbial Cell Factories* 16(1), p. 137.

De Deken, R. 1966. The Crabtree Effect: A Regulatory System in Yeast. *Journal of general microbiology* 44(2), pp. 149–156.

Deforet, M., van Ditmarsch, D., Carmona-Fontaine, C. and Xavier, J.B. 2014. Hyperswarming adaptations in a bacterium improve collective motility without enhancing single cell motility. *Soft matter* 10(14), pp. 2405–2413.

van Ditmarsch, D., Boyle, K.E., Sakhtah, H., et al. 2013. Convergent evolution of hyperswarming leads to impaired biofilm formation in pathogenic bacteria. *Cell reports* 4(4), pp. 697–708.

Inoue, T., Shingaki, R., Hirose, S., Waki, K., Mori, H. and Fukui, K. 2007. Genome-wide screening of genes required for swarming motility in Escherichia coli K-12. *Journal of Bacteriology* 189(3), pp. 950–957.

Jain, D.K. and Collins, D.L. A drop-collapsing test for screening surfactant-producing microorganisms. *Thompson*.

Kamatkar, N.G. and Shrout, J.D. 2011. Surface hardness impairment of quorum sensing and swarming for Pseudomonas aeruginosa. *Plos One* 6(6), p. e20888.

Kearns, D.B. 2010. A field guide to bacterial swarming motility. *Nature Reviews. Microbiology* 8(9), pp. 634–644.

Kim, W. and Surette, M.G. 2004. Metabolic differentiation in actively swarming Salmonella. *Molecular Microbiology* 54(3), pp. 702–714.

Klevens, R.M., Edwards, J.R., Richards, C.L., et al. 2007. Estimating health care-associated infections and deaths in U.S. hospitals, 2002. *Public health reports (Washington, D.C. : 1974)* 122(2), pp. 160–166.

Köhler, T., Curty, L.K., Barja, F., van Delden, C. and Pechère, J.C. 2000. Swarming of *Pseudomonas aeruginosa* is dependent on cell-to-cell signaling and requires flagella and pili. *Journal of Bacteriology* 182(21), pp. 5990–5996.

Kollaran, A.M., Joge, S., Kotian, H.S., et al. 2019. Context-Specific Requirement of Forty-Four Two-Component Loci in Pseudomonas aeruginosa Swarming. *iScience* 13, pp. 305–317.

Luo, Y., Zhao, K., Baker, A.E., et al. 2015. A hierarchical cascade of second messengers regulates *Pseudomonas aeruginosa* surface behaviors. *mBio* 6(1).

Mattingly, A.E., Kamatkar, N.G., Borlee, B.R. and Shrout, J.D. 2018. Multiple Environmental Factors Influence the Importance of the Phosphodiesterase DipA upon *Pseudomonas aeruginosa* Swarming. *Applied and Environmental Microbiology* 84(7).

Medina, G., Juárez, K., Valderrama, B. and Soberón-Chávez, G. 2003. Mechanism of Pseudomonas aeruginosa RhlR transcriptional regulation of the rhlAB promoter. *Journal of Bacteriology* 185(20), pp. 5976–5983.

Mellbye, B. and Schuster, M. 2014. Physiological framework for the regulation of quorum sensing-dependent public goods in Pseudomonas aeruginosa. *Journal of Bacteriology* 196(6), pp. 1155–1164.

Overhage, J., Bains, M., Brazas, M.D. and Hancock, R.E.W. 2008. Swarming of Pseudomonas aeruginosa is a complex adaptation leading to increased production of virulence factors and antibiotic resistance. *Journal of Bacteriology* 190(8), pp. 2671–2679.

Overhage, J., Lewenza, S., Marr, A.K. and Hancock, R.E.W. 2007. Identification of genes involved in swarming motility using a *Pseudomonas aeruginosa* PAO1 mini-Tn5-lux mutant library. *Journal of Bacteriology* 189(5), pp. 2164–2169.

Pessi, G. and 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. *Journal of Bacteriology* 182(24), pp. 6940–6949.

Picart-Armada, S., Fernández-Albert, F., Vinaixa, M., Yanes, O. and Perera-Lluna, A. 2018. FELLA: an R package to enrich metabolomics data. *BMC Bioinformatics* 19(1), p. 538.

Smith, E. and Morowitz, H.J. 2004. Universality in intermediary metabolism. *Proceedings of the National Academy of Sciences of the United States of America* 101(36), pp. 13168–13173.

Stickland, H.G., Davenport, P.W., Lilley, K.S., Griffin, J.L. and Welch, M. 2010. Mutation of nfxB causes global changes in the physiology and metabolism of Pseudomonas aeruginosa. *Journal of Proteome Research* 9(6), pp. 2957–2967.

Szenk, M., Dill, K.A. and de Graff, A.M.R. 2017. Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the Membrane Real Estate Hypothesis. *Cell Systems* 5(2), pp. 95–104.

Thévenot, E.A., Roux, A., Xu, Y., Ezan, E. and 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. *Journal of Proteome Research* 14(8), pp. 3322–3335.

Tremblay, J. and Déziel, E. 2010. Gene expression in Pseudomonas aeruginosa swarming motility. *BMC Genomics* 11, p. 587.

Vemuri, G.N., Altman, E., Sangurdekar, D.P., Khodursky, A.B. and Eiteman, M.A. 2006. Overflow metabolism in Escherichia coli during steady-state growth: transcriptional regulation and effect of the redox ratio. *Applied and Environmental Microbiology* 72(5), pp. 3653–3661.

Wang, Q., Frye, J.G., McClelland, M. and Harshey, R.M. 2004. Gene expression patterns during swarming in Salmonella typhimurium: genes specific to surface growth and putative new motility and pathogenicity genes. *Molecular Microbiology* 52(1), pp. 169–187.

Westbrock-Wadman, S., Sherman, D.R., Hickey, M.J., et al. 1999. Characterization of a Pseudomonas aeruginosa efflux pump contributing to aminoglycoside impermeability. *Antimicrobial Agents and Chemotherapy* 43(12), pp. 2975–2983.

Xavier, J.B., Kim, W. and Foster, K.R. 2011. A molecular mechanism that stabilizes cooperative secretions in *Pseudomonas aeruginosa*. *Molecular Microbiology* 79(1), pp. 166–179.

Yan, J., Deforet, M., Boyle, K.E., et al. 2017. Bow-tie signaling in c-di-GMP: Machine learning in a simple biochemical network. *PLoS Computational Biology* 13(8), p. e1005677.

Yan, J., Estanbouli, H., Liao, C., et al. 2019. Systems-level analysis of NalD mutation, a recurrent driver of rapid drug resistance in acute Pseudomonas aeruginosa infection. *PLoS Computational Biology* 15(12), p. e1007562.

Yeung, A.T.Y., Bains, M. and Hancock, R.E.W. 2011. The sensor kinase CbrA is a global regulator that modulates metabolism, virulence, and antibiotic resistance in *Pseudomonas aeruginosa.* *Journal of Bacteriology* 193(4), pp. 918–931.

Zhu, K. and Rock, C.O. 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.* *Journal of Bacteriology* 190(9), pp. 3147–3154.

**Figures**

A picture containing screenshot

Description automatically generated

**Figure 1.** Diversity of swarming across the *P. aeruginosa* phylogeny. **A.** Phylogeny of clinical isolates obtained from patients with cancer at MSKCC (Yan et al. 2017; Yan et al. 2019) together with reference strains PAO1, PA14 and PA7. The tissue of origin of each patient isolate is also shown. **B**. The ability to swarm (swarming picture shown) does not correlate with phylogeny; however, all swarmers can secrete rhamnolipids. They also tend to have a shorter lag phase in a minimal growth medium using glycerol as the sole carbon source, 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

**Figure S1:** Quantitative analysis of the swarming phenotype. **A**. We imaged swarming colonies grown overnight growth at 37ºC in soft-agar casamino acids plates. Roughly it is possible to identify swarmers by simple observation: the strains in bold). With the aim of getting a quantitative measure of swarming, we tested different morphological features of the colony using computational image analysis: maximum length (larger diagonal of the rectangle fitting the colony), skeleton, perimeter, area of the colony, area percentage (percentage of the plate occupied by the colony), circularity and eccentricity. Many of these features are correlated, so we used a principal component analysis to identify features that explained the majority variation among the strains, shown in the biplot presented here. *Max. length* and *circularity* were the two features that better separated the strains according to swarming. The top 3 strains with largest and shortest colonies (best swarmers and non swarmers), determined by eye, are highlighted in red and blue and the picture of their swarm is shown. Rhamnolipid production is also indicated.

A screenshot of a cell phone

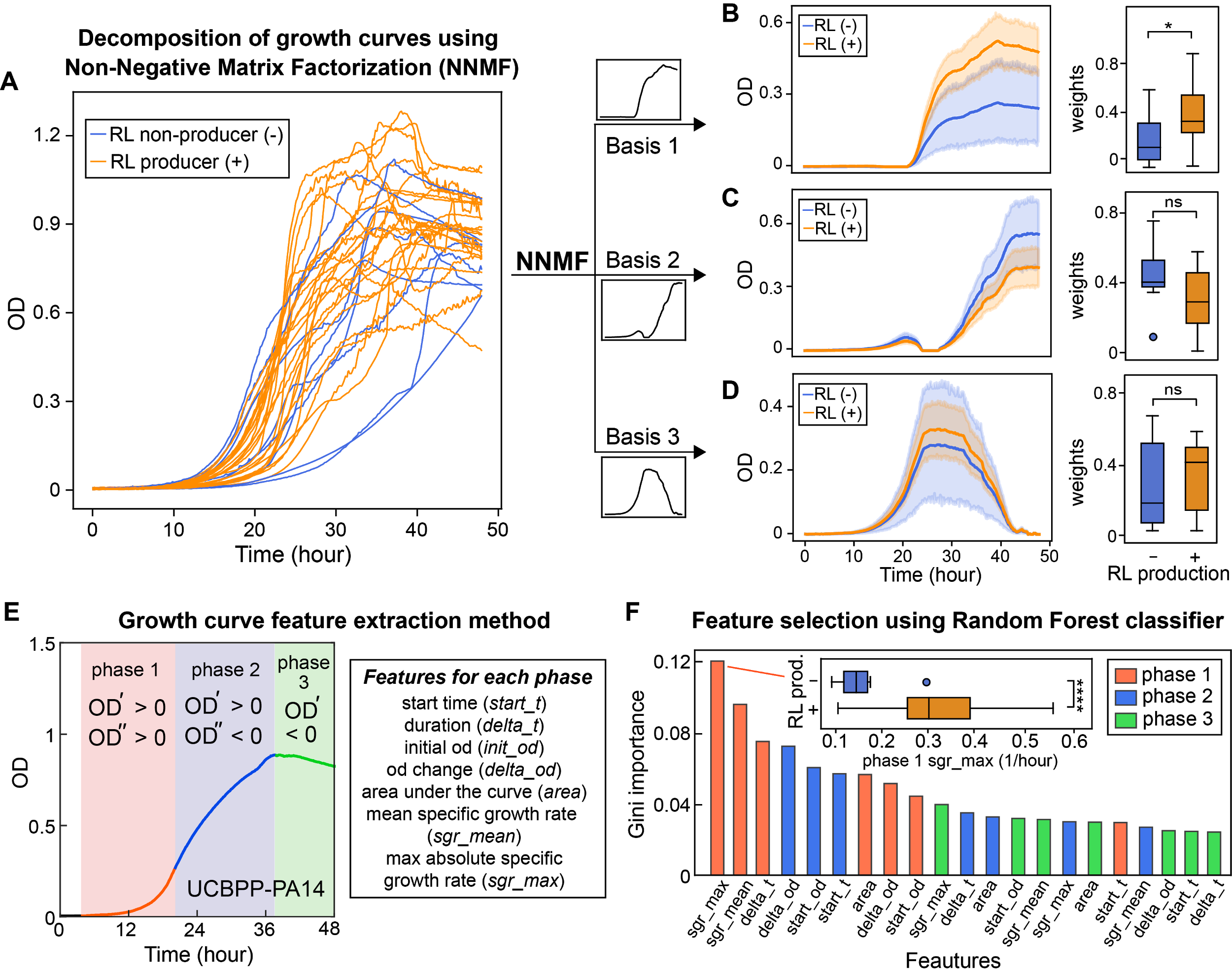
Description automatically generated

**Figure S2** Phylogenetic reconstruction of rhamnolipids production (A) and swarming phenotype (B). The pie chart in the nodes of branches indicate the likelihood of the phenotype for each ancestral state.

A screenshot of a cell phone

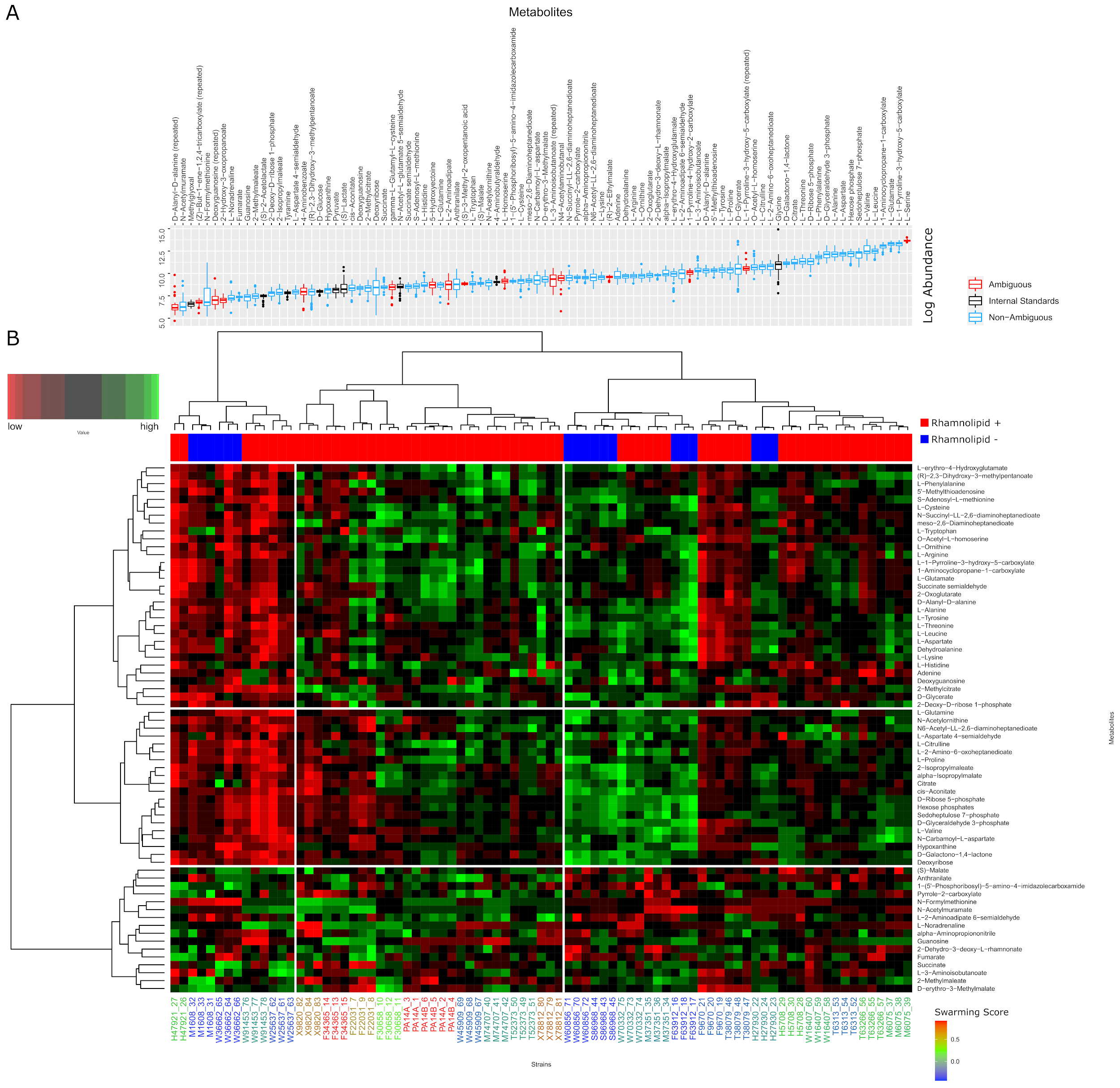
Description automatically generated

**Figure S3** Clustergram of growth curve in glycerol of *P. aeruginosa* isolates. The clinical isolates along with type strains PA14, PAO1 and PA7 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. The growth curves were clustered using Euclidean distance. The strains that by simple observation are swarmers are indicated in bold. Most of the swarmers are aggregated together, with the exception of F30658 and PAO1, which are both mild swarmers, and F23197, which has a longer lag phase in comparison to the rest of the isolates.



**Figure 2** Growth curve features distinguish rhamnolipid producers from non-producers. (**A-D**) 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-D**) 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. (**E**, **F**) Supervised feature selection using Random Forest classifier. (**E**) Feature extraction method. Each growth curve (excluding the initial lag phase) was divided into three phases (left panel) and each phase was described by 7 quantitative features (right panel). (**F**) Ranking of feature importance in classifying rhamnolipid producers. Inset: boxplot of maximum specific growth rate of phase 1 grouped by rhamnolipid production. Welch’s t-test was used in (**B-D**) and (**F**) for significance testing. \*\*\*\*, p-value 0.0001; \*, p-value 0.05; ns, p-value > 0.05.

**3**



**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.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.** A. Score 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 in our metabolomic dataset appear in bold. The colored ones are the ones that are correlated to any of the rhamnolipid production phenotypes. Pyruvate does not appear in B because it was found to be constant, so it was removed in the normalization step.

A screenshot of a cell phone

Description automatically generated

**Figure S4.** Loadings of the predictive component of the OPLS-DA model obtained to classify the strains according to if they produce or not rhamnolipids. 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 pathway for each one of the metabolites.

A close up of a sign

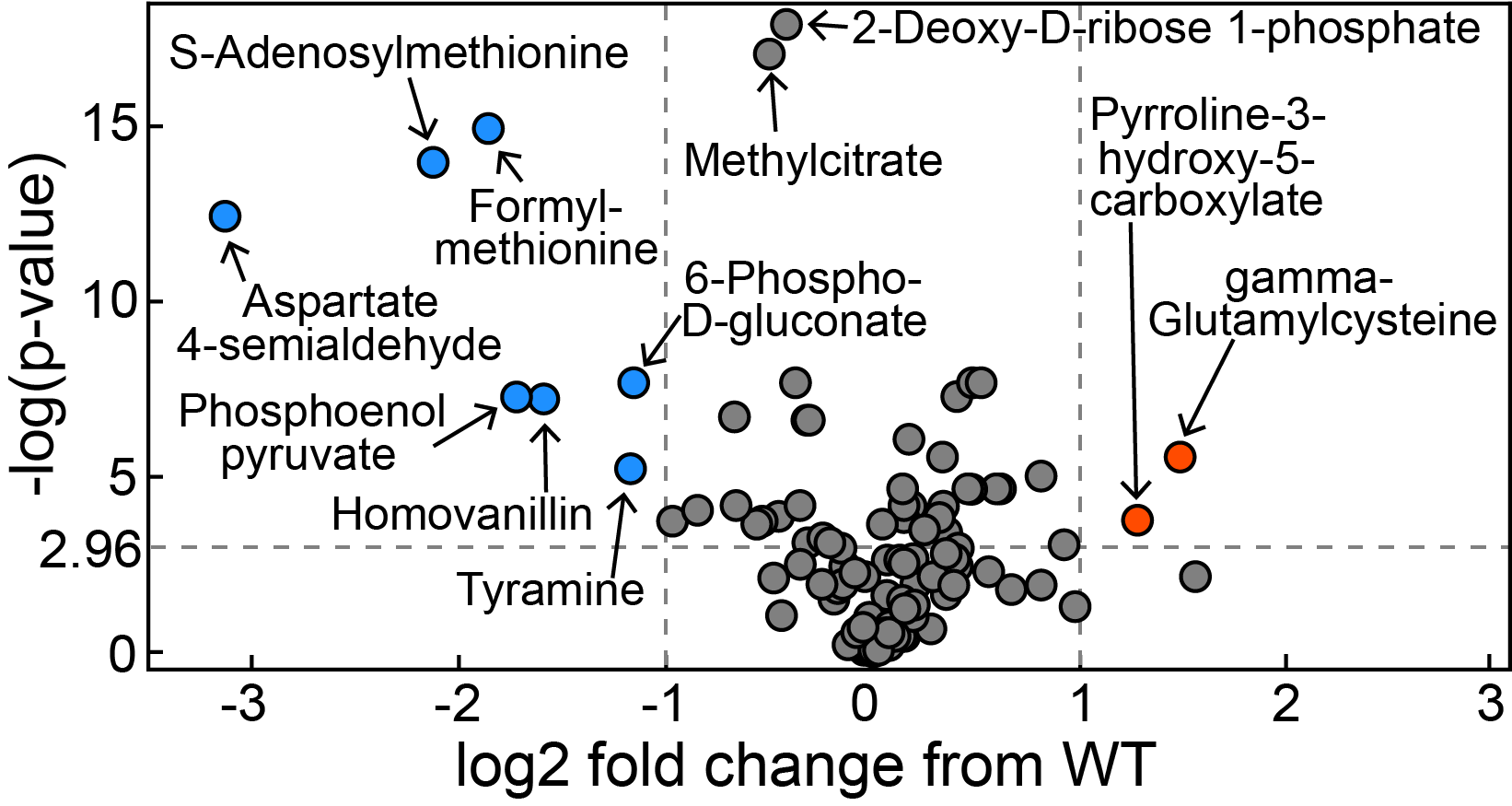
Description automatically generated

**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. dash line:

A picture containing kite, flying, fireworks

Description automatically generated

**Figure S5** 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.



**Figure S6.** Volcano plot of metabolomics data between wild-type *P. aeruginosa* UCBPP-PA14 strain and its *rhlA* mutant (replotted from Boyle et al. 2017).