

ARTICLE

Received 3 Jun 2016 | Accepted 18 Jan 2017 | Published 7 Mar 2017

DOI: 10.1038/ncomms14631

OPEN

Reconstruction of the metabolic network of *Pseudomonas aeruginosa* to interrogate virulence factor synthesis

Jennifer A. Bartell^{1,2,*}, Anna S. Blazier^{2,*}, Phillip Yen², Juliane C. Thøgersen³, Lars Jelsbak³, Joanna B. Goldberg^{4,5} & Jason A. Papin²

Virulence-linked pathways in opportunistic pathogens are putative therapeutic targets that may be associated with less potential for resistance than targets in growth-essential pathways. However, efficacy of virulence-linked targets may be affected by the contribution of virulence-related genes to metabolism. We evaluate the complex interrelationships between growth and virulence-linked pathways using a genome-scale metabolic network reconstruction of *Pseudomonas aeruginosa* strain PA14 and an updated, expanded reconstruction of *P. aeruginosa* strain PAO1. The PA14 reconstruction accounts for the activity of 112 virulence-linked genes and virulence factor synthesis pathways that produce 17 unique compounds. We integrate eight published genome-scale mutant screens to validate gene essentiality predictions in rich media, contextualize intra-screen discrepancies and evaluate virulence-linked gene distribution across essentiality datasets. Computational screening further elucidates interconnectivity between inhibition of virulence factor synthesis and growth. Successful validation of selected gene perturbations using PA14 transposon mutants demonstrates the utility of model-driven screening of therapeutic targets.

¹The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2970 Hørsholm, Denmark. ²Biomedical Engineering, University of Virginia, Charlottesville, Virginia 22908, USA. ³Department of Biotechnology and Biomedicine, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark. ⁴Department of Pediatrics, Division of Pulmonology, Allergy/Immunology, Cystic Fibrosis and Sleep, Children's Healthcare of Atlanta, Atlanta, Georgia 30322, USA. ⁵Emory + Children's Center for Cystic Fibrosis Research, Emory University and Children's Healthcare of Atlanta, Atlanta, Georgia 30322, USA. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.A.P. (email: papin@virginia.edu).

There is a need for new drugs that effectively inhibit microbial infection while avoiding the development of resistance. Traditional antibiotics that inhibit growth of bacteria by targeting growth-essential functions actively select for antibiotic-resistant mutants that overtake the infection. This growth-based selection promotes the rapid development of resistance and consequently exacerbates infections¹, resulting in substantial patient morbidity, mortality and health-care costs². Inhibiting mechanisms of infection by targeting the synthesis of virulence factors (VFs) and virulence-linked genes may be a promising new therapeutic strategy that avoids growth-based target selection, improves patient outcomes and mitigates the spread of resistance^{1,3,4}. However, genes that contribute to growth and genes that contribute to virulence are not necessarily distinct actors in an organism's genetic network; understanding the impact of genes on each pathogen directive (growth versus virulence) is critical to therapy design and prediction of resistance development.

Virulence-linked genes contribute to survival and fitness within a host. Many of these genes encode the synthesis pathways of VFs, pathogen-produced small molecules that are involved in activities such as iron sequestration and bacterial communication that enable adaptation to the host environment and enhance infection potential^{5,6}. In targeting the synthesis of these metabolites, resistance may develop more slowly because of weakened selection pressure versus traditional targets that directly impact growth-essential catabolism of substrates or cell wall construction and repair³. However, our understanding of the role of virulence-linked genes is evolving⁷—significant links between virulence and pathogen metabolism are now emerging. For example, antibiotic pigments called phenazines enable opportunistic bacteria to combat the effects of immune cell oxidative bursts, but these pigments may also induce rewiring of redox-linked pathways within the pathogen⁸. Furthermore, the production of virulence-linked compounds relies on essential components of central metabolism that connect substrate catabolism to VF synthesis pathways. A clear division between therapeutic targets impacting growth and virulence is therefore not straightforward⁷. We need to map the interconnectivity of these systems to identify genes that contribute to either or both systems, determine their function and essentiality in a clinically relevant environment, and estimate the impact of their inhibition on virulence versus growth.

To study the relationship between VF synthesis and growth from a systems level perspective, we used genome-scale metabolic network models (GEMs). Assembled from annotated genomic data, GEMs are mathematical frameworks that incorporate biochemical, genetic and cell phenotypic data and account for hundreds to thousands of gene-protein-reaction (GPR) relationships and reaction stoichiometry and directionality⁹; they have been used to predict novel drug targets that inhibit growth¹⁰ as well as probe the capability of an organism to synthesize various metabolites¹¹, including VFs (ref. 12).

Here, we present a new GEM of *Pseudomonas aeruginosa* strain PA14 (iPau1129) as well as an updated GEM of reference strain *P. aeruginosa* PAO1 (iPae1146). *P. aeruginosa* is a Gram-negative opportunistic pathogen capable of developing multi-drug antibiotic resistance, hospital-acquired infections^{13–15}, and infections in cystic fibrosis patient lungs, burn wounds and immunocompromised individuals. We validate our GEMs using substrate utilization data and gene essentiality screens from transposon mutant libraries and use six previously published transposon sequencing (Tn-seq) screens to evaluate essential virulence-linked genes^{16–18}. To study the relationship between VF production and growth, we compare the effect of *in silico* gene knockouts on synthesis of biomass versus 17 VFs and

identified genes uniquely critical for VF production, genes solely important for the synthesis of biomass, as well as genes involved in both VF production and biomass production. A case study of the VF pyoverdine shows the utility of GEMs in probing network dependencies that offer novel insights into links between virulence and metabolism that may enhance design of cycled or combination drug therapies as well as reduce the development of resistance.

Results

Metabolic network reconstruction of *P. aeruginosa*. Here, we present an updated GEM of *P. aeruginosa* strain PAO1 (iPae1146) as well as a new GEM of *P. aeruginosa* strain PA14 (iPau1129) (for ease of reference in this study, we refer to these reconstructions as mPAO1 and mPA14, respectively). The network reconstruction process began with previous *P. aeruginosa* PAO1 GEMs (refs 19,20). We implemented a more detailed biomass equation, incorporated new biological information, and curated the model against carbon source utilization and gene essentiality data (see below). We also assigned potential roles to 59 and 44 genes annotated as hypothetical proteins in PAO1 and PA14 genome annotations from the *Pseudomonas* Genome Database (PGD), respectively. In conclusion, the new GEM mPA14 accounts for the function of 1,129 genes, 1,495 reactions and 1,286 metabolites while the updated GEM mPAO1 accounts for the function of 1,146 genes, 1,493 reactions and 1,284 metabolites (Fig. 1a). The distribution of genes, metabolites and reactions in mPA14 across a variety of KEGG functional categories is shown in Fig. 1b (for the distribution of mPAO1, see Supplementary Fig. 1).

During curation, we specifically accounted for the synthesis pathways of several small molecule VFs. *P. aeruginosa* produces an array of VFs which can be grouped into several categories including exopolysaccharides, lipopolysaccharides, phenazines, quorum sensing signal molecules, siderophores and surfactants^{21,22}. Table 1 lists the compounds that can be synthesized by mPA14—the six italicized factors are new to mPAO1 and mPA14 compared to previous GEMs. Bolded dihydroaeruginic acid is a recently identified PA14-specific VF included only in mPA14 (ref. 23). Additionally, we evaluated a list of 454 genes linked to virulence of PAO1, PA14 or both in the Virulence Factor Annotations tool from the recently updated PGD to identify model genes that are associated with virulence. Only 123 of these virulence-linked genes were annotated as part of a BRITE metabolic pathway by KEGG, and 49 of the 454 genes were annotated as hypothetical proteins. Using KEGG and PseudoCAP annotations (functional system annotations developed by the PGD) as well as literature on VF synthesis, we focused on accounting for genes relevant to metabolism and virulence-linked synthesis pathways. Ultimately, there are 112 and 108 virulence-linked genes incorporated into mPAO1 and mPA14, respectively²⁴.

Model validation. We used two data types to curate and validate the models: a carbon source utilization data set and a published gene essentiality data set. We generated the substrate utilization data set using BIOLOG phenotype microarrays, which indicated whether PAO1 and PA14 were able to grow on particular carbon sources. We then compared these results to model predictions of biomass production (an approximation of growth) on different minimal media. After extensive transport reaction curation and refinement of metabolic pathways, mPA14 and mPAO1 account for 91 and 93 carbon sources and predict utilization with accuracies of 81% and 80%, respectively (Supplementary Fig. 2).

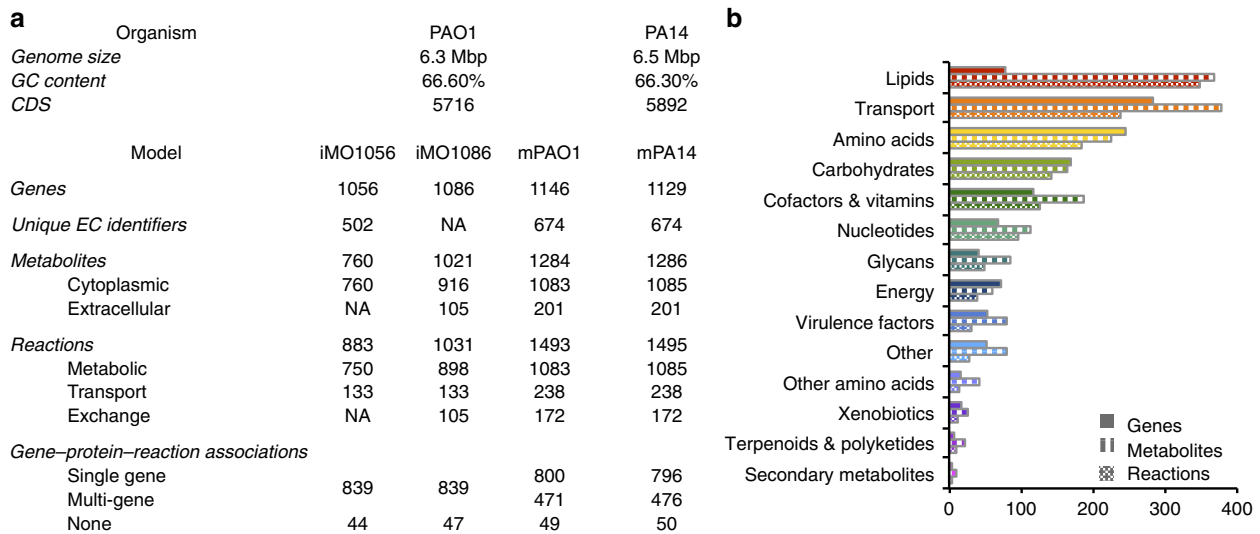


Figure 1 | Network model characteristics. (a) Properties of the updated PAO1 model as compared to previously published GEMs for *P. aeruginosa*, iMO1056 and iMO1096, as well as properties of the new PA14 model. (b) The number of genes, metabolites and reactions in mPA14 grouped into functional categories as defined by KEGG (ref. 74). For the distribution of genes, metabolites and reactions in mPAO1, see Supplementary Fig. 1.

Table 1 | Small molecules associated with virulence accounted for in mPA14 and mPAO1.

Category	Virulence factor	Metabolite ID
Exopolysaccharide	Alginate	cpd17074[c]
Lipopolysaccharide	A-band-O-antigen	cpd17056[c]
	B-band-O-antigen	cpd17057[c]
	Lipid A	cpd17066[c]
Phenazines	Chorismate	cpd00216[c]
	<i>1</i> -Carboxyphenazine	cpd17083[c]
	Pyocyanin	cpd01206[c]
Quorum sensing signal molecules	Acyl-homoserine-lactone (AHL)	cpd17082[c] and cpd08635[c]
	<i>cis</i> -2-Decenoic acid (DSF)	<i>cJB00127</i> [c]
	<i>2</i> -heptyl-4-quinolone (HHQ)	cpd17078[c]
	<i>Pseudomonas</i> quinolone signal (PQS)	cpd17085[c]
Siderophores	Dihydroaeruginic acid (Dha)	cJB00126 [c]
	<i>Pyochelin</i>	cpd08828[c]
	<i>Pyoverdine</i>	cPY00164[c]
	<i>Salicylate</i>	cpd00599[c]
Surfactant	Rhamnolipid	cpd17081[c] and cpd17080[c]

The six italicized factors are new additions to mPA14 and mPAO1 as compared to previous GEMs, and bolded dihydroaeruginic acid is a recently identified PA14-specific VF included only in mPA14 (ref. 23). The metabolite ID is the compound reference ID used in our models.

For the gene essentiality validation, we used a published data set comprised of the overlap between the essential PAO1 genes identified in Jacobs *et al.*²⁵ and the essential PA14 genes identified in Liberati *et al.*²⁶ by creation of transposon insertion mutant libraries in a Luria-Bertani (LB) media background. We used this overlap data set for our curation efforts instead of the individual libraries because there is a higher confidence in which genes are essential since they were not disrupted in either of the PAO1 and PA14 screens and the libraries are validated, publically available and created with established approaches in rich media. To compare this overlap data set with our model predictions, we performed *in silico* single gene knockouts in our models and measured the subsequent effects on biomass production.

Ultimately, both mPA14 and mPAO1 can be used to predict gene essentiality with an accuracy of 91% (Supplementary Fig. 3).

Virulence associations of Tn-seq-based essential genes.

The recent advent of Tn-seq high-throughput screening has enabled the rapid evaluation of the fitness of a transposon insertion mutant in a given condition. These screens contribute important information regarding bacterial survival in specific contexts. Given our interest in studying the relationship between growth and virulence, we sought to determine how many essential genes were also virulence-linked in recent *P. aeruginosa* Tn-seq screens.

Table 2 | Comparison of experimental essentiality screens.

Strain	PA14				PAO1				
	Reference Media	Pier LB	Whiteley Sputum	Manoil LB	Whiteley LB	Manoil Sputum	Whiteley Sputum	Manoil Pyruvate	Whiteley Succinate
No. of essential genes		634	510	201	336	224	445	182	641
No. of VF-linked essential genes		49	25	21	20	30	41	27	54
%VF-linked essential genes		7.73	4.90	10.45	5.95	13.39	9.21	14.84	8.42

The number of essential genes, virulence-linked essential genes (VF-linked essential genes), and the proportion of essential genes that are virulence-linked for all eight of the transposon mutant screens analysed. The data are assembled from studies by the Pier Lab in 2013 (ref. 16) and the Whiteley Lab¹⁷ and Manoil Lab¹⁸ in 2015.

We obtained data from published Tn-seq screens for PAO1 and PA14 in several culturing conditions and identified essential genes for each individual screen. We then compared these individual essential gene lists to a list of either PAO1 or PA14 virulence-linked genes from the PGD to identify virulence-linked essential genes for each screen (Table 2). The number of virulence-linked essential genes identified across the screens ranged from 20 in the Whiteley PAO1 LB data set¹⁷ to 54 in the Whiteley PAO1 succinate data set¹⁷. Furthermore, the proportion of essential genes that were also virulence-linked varied across the screens, ranging from 4.9% in the Whiteley PA14 sputum data set¹⁷ to 14.8% in the Manoil PAO1 pyruvate data set¹⁸. This variability in the percentage of virulence-linked essential genes may stem from the variability in transposon insertion coverage of the individual screens. While some screens identified over 600 essential genes^{16,17}, other screens identified less than 200 essential genes¹⁸ in the same media.

The moderate number of virulence-linked genes present in the Tn-seq screens can partially be explained by the lack of host selection pressure in the generation of the mutant libraries and the imperfect replication of *in vivo* growth conditions in *in vitro* studies. Tn-seq screens performed in infection models have demonstrated that mutants unable to synthesize certain VFs are unable to colonize the infection site²⁷, suggesting that virulence-linked genes may be essential in some contexts, while elsewhere (such as in liquid culture) they are unnecessary for bacterial fitness. However, using only Tn-seq screens that differ by growth media rather than host selection pressure to contextualize virulence-linked gene essentiality still shows that some virulence-linked genes have important, potentially non-virulence related, functions. This analysis indicates that these genes may play a more focused metabolic role in the development of infection or are capable of dual functions linked to both virulence and growth.

To evaluate the potential overlap of virulence-linked genes with growth activity, we first used mPA14 as a framework to compare sets of growth essential genes and virulence-linked genes that have been curated as functionally relevant to metabolic activity using the Whiteley PA14 sputum screen¹⁷. Figure 2 shows the model reactions linked to 205 genes required for growth of PA14 in sputum (blue), and the 108 PA14 virulence-linked genes from the PGD (red). The overlap between reactions associated with required genes and virulence-linked genes, totalling 21 reactions (11 genes) are linked to a broad array of systems and present at high density in central metabolic pathways, amino acids, lipids and nucleotide metabolism (overlap reactions in purple). Intriguingly, many reactions associated only with growth or virulence group together in the same pathways, which may indicate functional connections even if specific genes are not shared between the distinct gene sets. This analysis supports the need for a mechanistic evaluation of virulence-linked genes in the context of growth.

Modelling VF production capabilities. While infection-based Tn-seq screens have demonstrated that mutants incapable of VF synthesis lack the ability to infect, it is challenging to discern whether this occurs due to the inhibited gene's essentiality for the expression of virulence-linked compound(s), essentiality for growth, or essentiality for both^{7,28}. To address this gap in knowledge, we employed genome-scale metabolic network modelling. We implemented a medium that mimics the lung of cystic fibrosis patients (synthetic cystic fibrosis medium, SCFM) in order to more closely model *in vivo* conditions²⁹. The ability of *P. aeruginosa* to maintain decades-long infections in the lungs of cystic fibrosis patients may be due to both its metabolic adaptability and deployment of an array of VFs, such that pathway interconnectivity may proffer unique metabolic benefits as well as enable resistance to treatment³⁰. Using an *in silico* SCFM medium, we performed *in silico* single-gene knockouts and assessed the levels of growth inhibition and VF synthesis inhibition by normalizing the resulting biomass flux and VF flux to wildtype production levels. By repeating this analysis for all 17 VFs in our model, we quantitatively compared the broad effects of simple genetic perturbations on the production of different VFs versus growth.

Core set of growth-essential genes impact VFs. To study the role of genes critical to both growth and VF synthesis, we compared the 116 genes predicted by mPA14 as essential for growth on SCFM to the genes essential for synthesis of VFs and found that 46 of the growth-essential genes are also essential for the production of at least one VF. These 46 genes critical to both biomass production and virulence are listed in Fig. 3 with their PseudoCAP category and function and a heatmap showing the affected VFs. The PseudoCAP category critical for the largest number of VFs is fatty acid and phospholipid metabolism, with seven genes predicted to be essential for the production of at least eight VFs in addition to biomass production. Additionally, several *aro* operon genes contributing to aromatic amino acid synthesis are essential for the production of six VFs in the phenazine and siderophore families, while an array of genes involved in purine metabolism fully inhibit only the production of A-band-O-antigen. Ultimately, this analysis provides a novel list of genes ranked by their impact on virulence pathways in addition to growth inhibition, which may assist the design of therapeutics with broad impact on metabolic processes.

We then expanded our analysis to all genes in our model, plotting inhibition of each VF versus growth (Fig. 4). Each point in the resulting plots indicates the level of growth inhibition (*x* axis) and VF inhibition (*y* axis) relative to wild-type for a given *in silico* knockout. All data points are transparent such that a high density of data points results in an increase in colour intensity. Thus, the colour intensity at the origin of the plots indicates a high number of gene deletions that have no effect on production

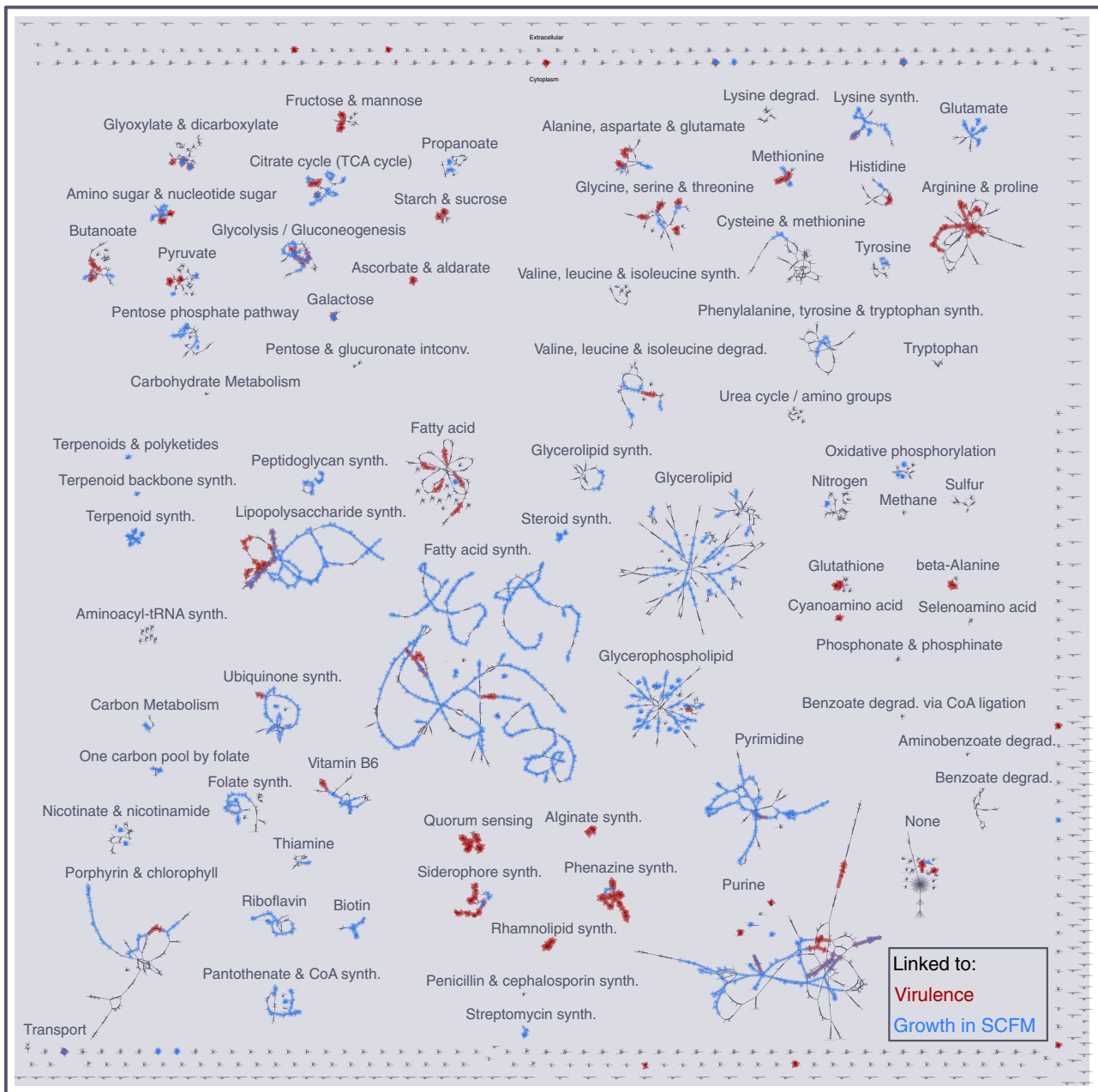


Figure 2 | Visualization of experimental virulence-linked essential genes. Distribution of virulence-linked genes and growth-essential genes from experiments in CF sputum visualized across all mPA14 reactions (grey) using MetDraw. Reactions associated with virulence-linked genes (as defined by the *Pseudomonas* Genome Database) are highlighted in red, and reactions associated with genes essential to growth in synthetic cystic fibrosis sputum are shown in blue. Purple reactions are associated with both virulence and growth essentiality. All reactions and metabolites are labelled with unique identifiers referenced in the model, visible at high magnification and text-searchable.

of either biomass or the indicated VF. Data points in the upper right corner of each plot represent genes essential to both VF production and growth, while data points arrayed between axes indicate the degree of biased impact on growth versus VF production by a given knockout.

This analysis enabled the identification of non-obvious relationships between growth and VF production. Unsurprisingly, most gene knockouts resulted in marginal or no growth defects, as indicated by data point clusters near the origin along the *x* axis. This result was mirrored for VF synthesis, with most gene knockouts also resulting in marginal or no VF production defects. We hypothesized that VF synthesis would be less robust to perturbation as compared to growth because these compounds

rely on the catabolism of growth substrates prior to VF anabolism. We instead see that for several VFs, many genes essential for growth only partially inhibit synthesis when disrupted. The number of genes essential solely for the production of a given VF varies considerably, and is not always correlated with the complexity of the synthesis pathway or final compound. These results highlight critical differences in the degree of interconnectivity of VF synthesis and biomass production across the VFs, which we can evaluate mechanistically through the use of our computational model.

VFs that are less sensitive to genetic perturbations than biomass production may have a high degree of redundancy in their synthesis pathways. For example, relatively few genes impact

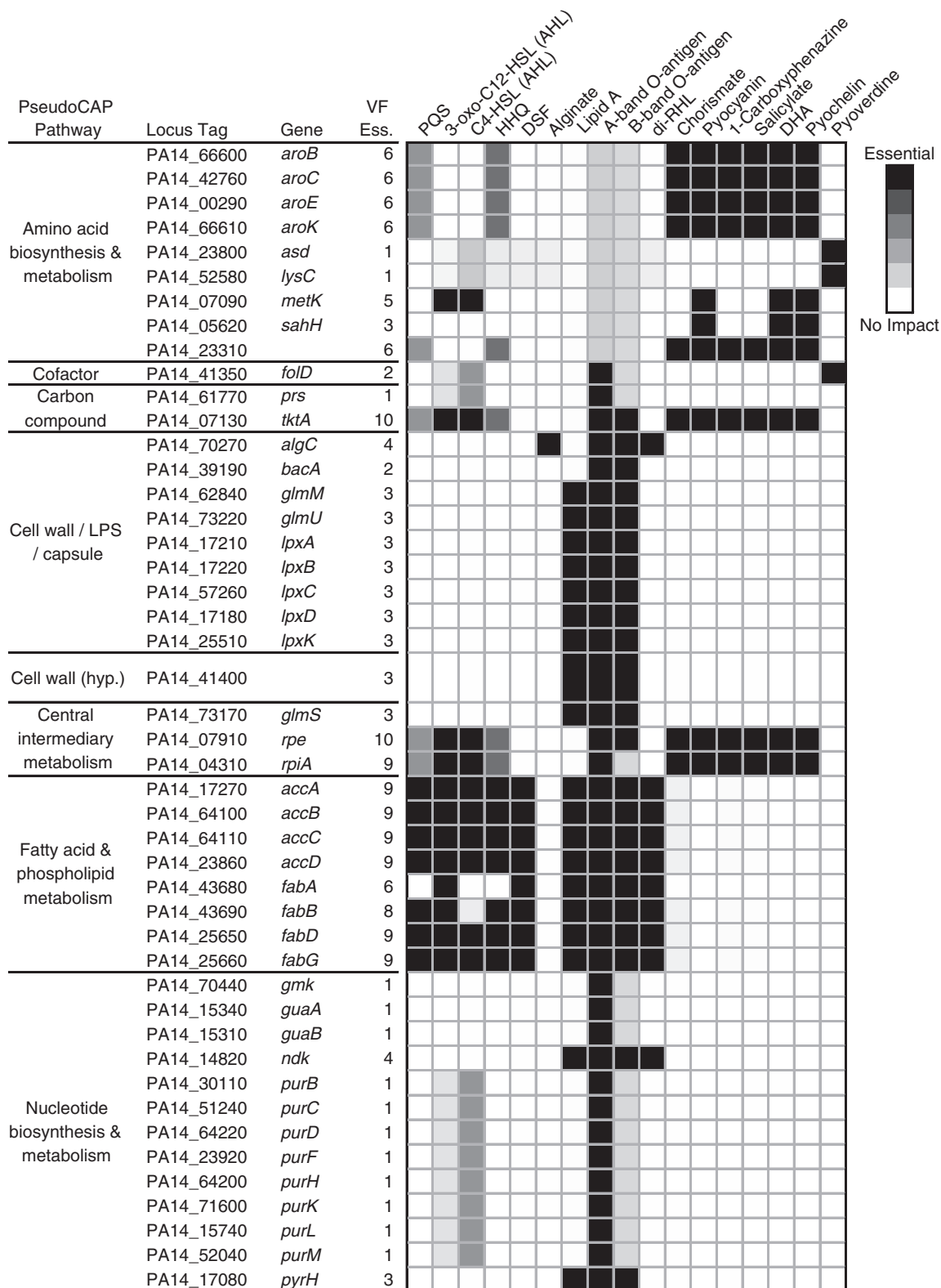
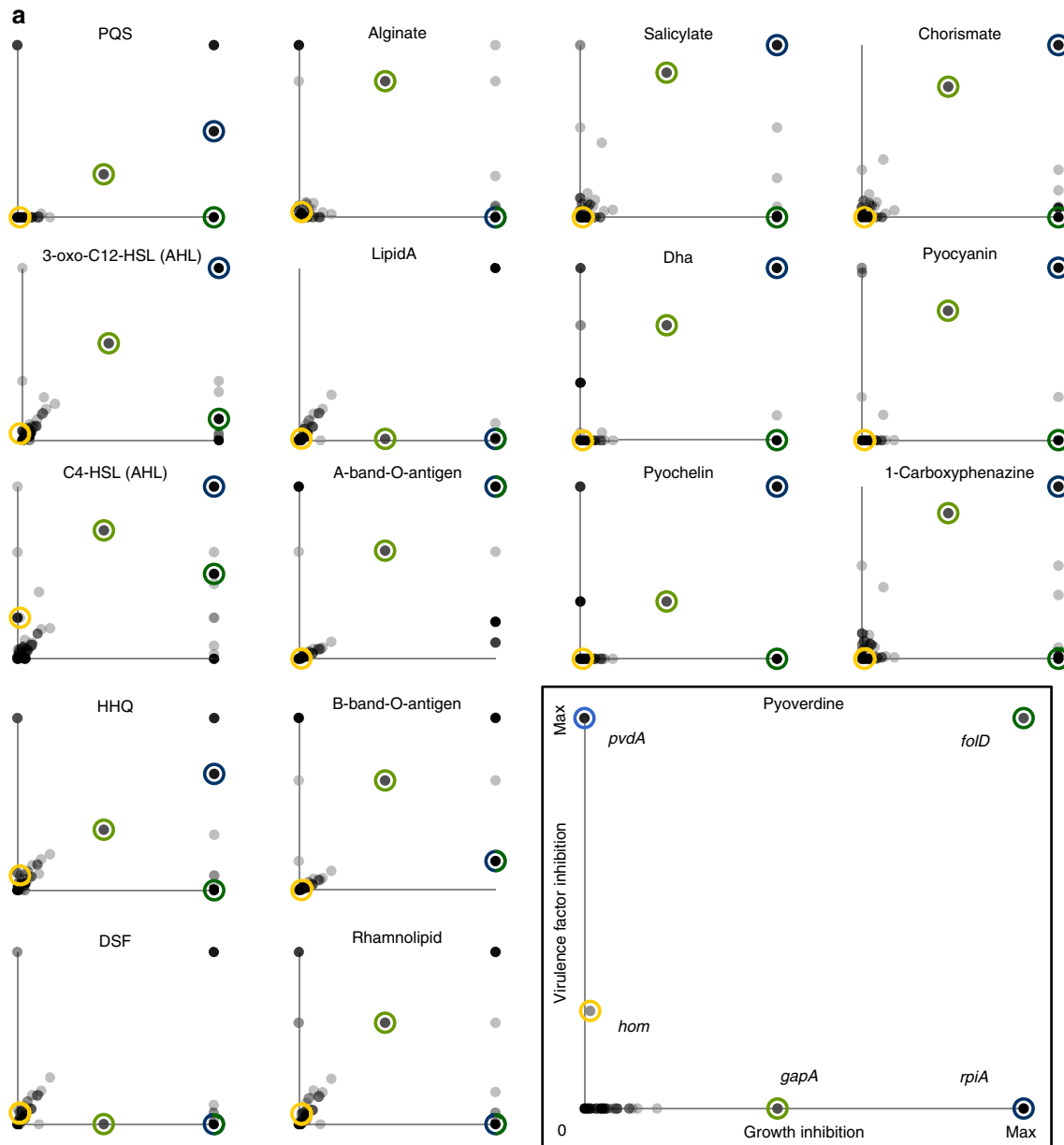


Figure 3 | Genes essential for VF synthesis versus growth in SCFM. The table lists the 46 genes essential for growth and production of at least one VF. Pathway assignment via PseudoCAP annotation and tabulated count of VFs for which the gene is essential are also included. Impact of a given gene's deletion is shown as white indicating 0% inhibition and black indicating 100% inhibition.

lipid A, chorismate and 1-carboxyphenazine production without also impacting growth and no gene is essential solely for the production of the respective VF. Instead, genes that are essential for VF production are also essential for growth, thus indicating

the high level of integration of VF synthesis with the overall metabolism of *P. aeruginosa*. While this integration of VF and biomass synthesis is expected for lipid A given its presence in the biomass reaction in the model as an essential component, this



b

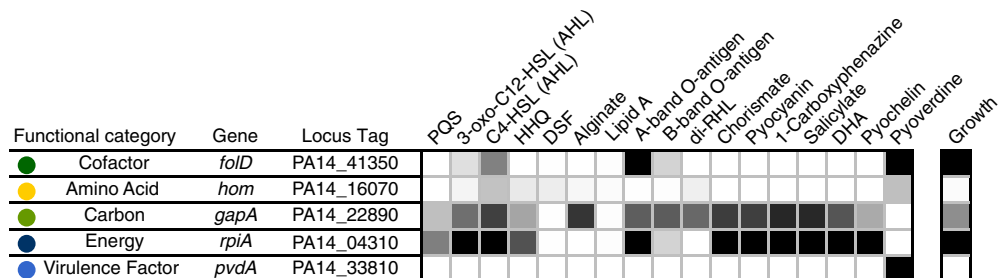


Figure 4 | VF synthesis and growth interconnectivity. (a) *In silico* gene knockouts were performed and the subsequent levels of biomass and VF production for each VF in the model were measured. The amount of growth inhibition was calculated by normalizing the knockout biomass production to the wild-type biomass production. Likewise, the amount of VF synthesis inhibition was calculated by normalizing the mutant level of VF production to the wild-type level of VF production. Each point indicates the growth inhibition (x axis) and VF inhibition (y axis) relative to wild-type for a given *in silico* knockout. All data points are transparent such that a high density of data points results in an increase in colour intensity. Coloured circles are used to indicate genes of interest as labelled in the pyoverdine example with yellow representing genes involved in amino acid metabolism, green carbohydrate metabolism, dark blue energy metabolism and light blue VF metabolism. (b) We highlight genes representing unique subtypes of impact on pyoverdine synthesis versus growth in a quantitative way that enables easy comparison of the activity of these genes across all VFs, with white indicating 0% inhibition and black indicating 100% inhibition.

was a surprising result for chorismate and 1-carboxyphenazine as we do not consider these essential components. Upon closer network inspection, we find alternative pathways for the production of chorismate and multiple isozymes for the synthesis of 1-carboxyphenazine. Both instances highlight redundancies in the network that reduce the occurrence of predicted essential genes unique to these two VFs.

Interestingly, B-band-O-antigen and A-band-O-antigen demonstrate the other extreme—all growth essential genes also impact the synthesis of both O-antigens to some extent. This case highlights the dependency of the production of these O-antigens on some of the biomass components themselves, namely lipid A. Since the O-antigens rely on the production of lipid A, all the genes that inhibit the synthesis of lipid A (and, thus, biomass) also inhibit the synthesis of the O-antigens. This result indicates the importance of biomass function formulation; here, we retain a standard list of components for consistency with other models, but more targeted analyses may improve upon addition and removal of less integral biomass components.

In contrast to these VFs, there were several VFs that displayed much less interconnectivity with the rest of the network. For example, relatively few genes maximally inhibit both growth and alginate production compared to the other VFs in the model. Most of the genes that are essential for alginate production have no impact on growth when the associated function is removed from the model, thus indicating that the synthesis pathway for alginate is less highly integrated into the metabolism of *P. aeruginosa*. VFs like alginate may be more peripheral to the general metabolic function of *P. aeruginosa* due to specialization. While *in vivo* studies have highlighted the importance of these metabolites in maintaining infections^{31–34}, here we systematically demonstrate pathway independence from essential metabolic function.

Interconnectivity of synthesis of pyoverdine and biomass. In addition to studying the interconnectivity of individual VF synthesis pathways, this analysis also enabled the investigation of the role of individual genes. While the disruption of some gene functions results in a similar response across the VFs, other gene function disruptions produce a highly varied response, suggesting that these genes play a unique role in the synthesis of each VF (example genes circled in Fig. 4). Using the VF pyoverdine as a reference, when the function of the gene *hom*, encoding for homoserine dehydrogenase, is removed, there is a very slight impact on growth and only marginal impact on each of the VFs, with pyoverdine synthesis demonstrating the most inhibition as a result of a *hom* knockout. Similarly, while a functional disruption of *pvdA*, which encodes for L-ornithine N5-oxygenase, maximally inhibits pyoverdine production, it has no impact on synthesis of the other VFs accounted for in mPA14.

In contrast, other simulated gene knockouts have more varied impact on VF inhibition. Functional disruption of *gapA*, which encodes for glyceraldehyde-3-phosphate dehydrogenase, has varied impact on the synthesis of VFs in the network, illustrated by preservation of pyoverdine production but near maximal impact on salicylate production. Similarly, while functional disruption of *rpiA*, which encodes for ribose-5-phosphate isomerase, again does not inhibit pyoverdine production, it does inhibit production of the AHLs incorporated into mPA14. Likewise, functionally disrupting *folD*, which encodes for 5,10-methylene-tetrahydrofolate dehydrogenase, maximally inhibits pyoverdine production and has no impact on PQS production. Thus, we can tease out the role of different genes on the synthesis of different VFs, with some simulated gene knockouts demonstrating consistent levels of inhibition

across all of the VFs and others demonstrating varied levels of inhibition.

Experimental evaluation of pyoverdine mutants. We chose to extend our investigation of the inhibition of pyoverdine synthesis because of the important role it has in iron scavenging and the tractability of experiments measuring pyoverdine production. In fluorescent *Pseudomonads*, pyoverdine is the main siderophore, a molecule that solubilizes iron for use by essential metabolic processes. It has been implicated in bacterial interactions in biofilms, it is essential for burn wound colonization, and it is upregulated in initial CF lung colonization^{35–37}. Pyoverdine is also considered a ‘public good’ compound that is produced by select members of a community to benefit the whole. Thus, the inhibition of pyoverdine synthesis within the small group of producer cells may affect the whole community while reducing the possibility of acquisition and spread of resistance genes^{38,39}.

To interrogate the relationship between growth and pyoverdine synthesis, we identified gene function disruptions with varied impact on pyoverdine synthesis and growth as shown by the circled points of Fig. 4. We chose *pvdA* because it was predicted to be essential for pyoverdine production but not growth. Conversely, *rpiA* was chosen because it was predicted to be essential for growth but not for pyoverdine production. We chose *folD* because our model predicted it to be essential for both growth and pyoverdine production, and *hom* and *gapA* because of their predicted sub-inhibitory effects on pyoverdine production and growth, respectively. We then investigated the accuracy of these predictions with literature and experiments using available transposon mutants.

Mutants for both *folD* and *rpiA* were not present in the PA14 genome-wide transposon mutant library²⁶, suggesting that both these genes are indeed essential for growth of *P. aeruginosa*. Involved in the folate biosynthetic pathway, *folD* plays a critical upstream role in the synthesis of several compounds such as thymidine, purines and various amino acids. Studies have investigated *folD* as a potential therapeutic target to kill a variety of pathogens including *P. aeruginosa*^{40–42}. Also important for purine synthesis, *rpiA* plays a critical role in the pentose phosphate pathway, converting D-Ribulose-5 to D-Ribose-5. Due to their growth essentiality, it is not feasible to study their role in VF synthesis experimentally—we instead use our computational model to offer unique insight. While mPA14 predicts that *rpiA* is not important in pyoverdine synthesis via a simulated knockout, it does predict that *folD* plays a crucial role, as evidenced by a simulated knockout resulting in total inhibition of pyoverdine synthesis. An analysis of pyoverdine synthesis precursors that cannot be produced after an *in silico folD* knockout in mPA14 highlights N5-formyl-N5-hydroxy-L-ornithine as the missing metabolite. This metabolite is not included in the much longer list of missing metabolites including purines that prevent biomass formation by the model. Thus, while *folD* may be essential for growth because of its role in purine synthesis, it appears to be essential for pyoverdine synthesis because of its role in amino acid metabolism. Understanding the metabolic interconnectivity of these genes provides insight into their potential impact on multiple systems if targeted therapeutically and we are able to determine the role of growth essential genes in VF synthesis which would otherwise be intractable.

Using transposon mutants of *pvdA*, *hom* and *gapA* from the PA14 genome-wide transposon mutant library²⁶, we performed absorbance-based assays of pyoverdine production and growth in SCFM as described in the methods. The extent of growth and pyoverdine production (normalized to growth) for wild-type

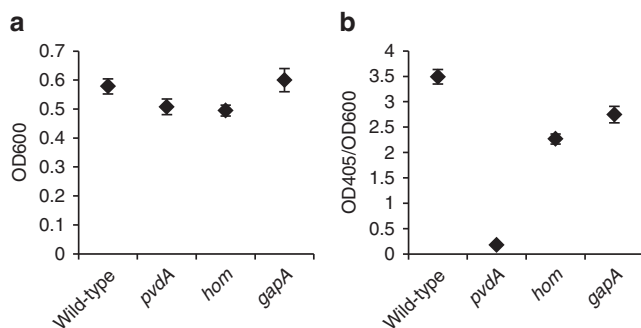


Figure 5 | Pyoverdine synthesis capabilities *in vitro* on SCFM. PA14 wild-type and *pvdA*, *hom* and *gapA* PA14 mutants were grown to stationary phase in SCFM and growth was measured using OD600 (a). Subsequently, the supernatants were isolated and the OD405 of each condition's supernatant was measured as a proxy for pyoverdine levels. The OD405 was divided by the OD600 of the culture in order to normalize pyoverdine production to growth (b). Error bars indicate s.d. among five biological replicates.

PA14 and each mutant strain are shown in Fig. 5a and b, respectively. As expected, the *pvdA* mutant showed markedly decreased pyoverdine production compared to wild-type, as it is an established pyoverdine assay control directly connected to the pyoverdine synthesis pathway. Interestingly, the *pvdA* mutant resulted in a minor growth defect relative to wild-type. This result could indicate that the lack of pyoverdine, and thus lack of access to iron, somewhat hindered the growth of the *pvdA* mutant. The *hom* mutant displayed a marginal growth defect and an approximately 1.5-fold decrease in pyoverdine production relative to wild-type. These results are consistent with model predictions that a *hom* knockout would result in slight growth inhibition and significant, but not total, inhibition of pyoverdine synthesis. Homoserine dehydrogenase, the gene product of *hom*, catalyses the reaction converting L-homoserine to L-aspartate 4-semialdehyde which then gets converted to L-2,4-diaminobutyrate, an important precursor of pyoverdine⁴³. Thus, a mutated *hom* limits the production of L-aspartate 4-semialdehyde, creating a bottleneck in pyoverdine synthesis. Because, L-aspartate 4-semialdehyde is not a growth-essential metabolite, both *in vitro* and *in silico*, targeting it may specifically prevent pyoverdine production without strong growth-based resistance selection. Unlike the *pvdA* and *hom* mutants, the *gapA* mutant did not exhibit a growth defect, disagreeing with our model prediction of an approximately 50% reduction in growth. Additionally, while we predicted that a *gapA* knockout would not impact pyoverdine production, we observed that the *gapA* mutant did indeed reduce pyoverdine synthesis, albeit to a lesser extent than the *pvdA* and *hom* mutants. *gapA* encodes for the enzyme glyceraldehyde-3-phosphate dehydrogenase, which catalyses the conversion of glyceraldehyde-3-phosphate to 1,3-bisphospho-D-glycerate, a central reaction in glycolysis. The discrepancy between our model predictions and experimental results for both growth and pyoverdine synthesis identifies a 'gap' in our knowledge regarding the function of glyceraldehyde-3-phosphate dehydrogenase in the overall metabolic network. Thus, we experimentally tested model predictions regarding genes in both growth and VF synthesis to tease out the role of genes upstream of pyoverdine synthesis and identify 'gaps' in our current understanding of *P. aeruginosa* metabolism.

Discussion

We utilized a novel approach to systematically evaluate the contribution of metabolic genes to the synthesis of factors critical

to virulence as well as growth at genome-scale using metabolic models. A new GEM for *P. aeruginosa* PA14 and an updated GEM for strain PAO1 were curated using single transposon insertion mutant data, virulence-linked gene databases, substrate utilization data, updated genome annotations and recent literature. Using our PA14 model, we contextualized the PGD database of virulence-linked genes that were identified as essential for growth in Tn-seq screens, and then identified a core set of metabolic genes that were necessary for both growth and the synthesis of at least one VF. Subsequent analyses mapped the metabolic interconnections between growth and the synthesis of individual VFs, using a case study of pyoverdine to demonstrate model utility in teasing out the role of individual genes with regards to both growth and VF production. Our work enhances understanding of relationships between VF synthesis and growth, which is challenging to elucidate with experimental approaches. By quantifying the impact of genetic targets on growth versus virulence using a mechanistic model, we contribute novel insights for the design of therapeutic strategies that account for potential resistance development.

The rapid fitness screening enabled by transposon insertion mutants have produced valuable insights into gene function in different environments⁴⁴, but the genome-scale data sets can be difficult to interpret. Signature-tagged mutagenesis screens in infection models have identified virulence-linked genes, but require a highly accurate replication of growth conditions *in vitro* for a baseline of effective comparison with infection^{45,46}. Recently, Tn-seq has been used to measure *in vivo* gene fitness⁴⁴, but transposon insertion coverage, interpretation of Tn-seq results and essential gene identification are difficult to replicate across studies. Our model provides important nuance when considering the true definition of an 'essential' gene and survival fitness in varied environmental conditions; genes can be classified in a quantitative manner instead of a binary 'essential/nonessential' categorization. Thus, the high degree of variability in the number and roles of genes identified as critical for fitness even in different Tn-seq studies of the same growth environment can be elucidated when paired with mechanistic modelling. While we focused our efforts on mapping the distribution of virulence-linked genes in the data sets in an effort to understand their impact without the pressure of survival in a host, there is a rich opportunity to expand this comparison to the distribution of all metabolic genes in future work.

Our analysis provides an important expansion of genes to consider when studying VF synthesis during adaptation. We identify 46 genes as critical for the production of up to 10 of the 17 assessed VFs as well as biomass, which represent a novel core set of metabolic functions integral to the development of infection by *P. aeruginosa*. Long-term adaptation may result in altered virulence capabilities due to accumulated mutations in these genes as well as regulatory genes and genes linked directly to product synthesis⁴⁷. When we expand our analysis to all model genes, we can then group genes by their functional impact on growth, VF synthesis or a combination of roles highlighting a higher degree of connection. This novel analysis provides testable hypotheses regarding the contribution of a gene to a given synthesis task. While we focus on experimental assessment of uniquely categorized genes in our pyoverdine analysis, this approach can be expanded to the other VFs. The analysis also demonstrates the flexibility of *P. aeruginosa* metabolism in a substrate-rich environment; the impact of competition for resources and the role of auxotrophy in evolving strains can be compared using predictions of differing optimal growth and production levels. Thus, further mapping of gene function may be enhanced by repeating the study in less complex and varied growth conditions.

Therapeutically targeting virulence-related pathways is an approach attracting much attention from a field struggling to find effective treatment for drug-resistant pathogens^{1,48}. Quorum sensing inhibitors have been investigated through small molecule screening for a range of pathogens including *P. aeruginosa*^{49–51} partially as a means of broad-spectrum anti-virulence treatment. A recent study showed that gallium-based quenching of extracellular siderophore activity successfully inhibited infection of caterpillars by *P. aeruginosa* while avoiding the development of resistance⁵². While inhibiting siderophore synthesis may increase resistance incidence in comparison to quenching, it will also enable pathogen-specific targeting of iron sequestration and avoid other systemic side effects (for example, radiation with respect to gallium). While these VFs are regulated by quorum sensing molecules, related signalling networks are complex; more direct routes of inhibition provide an efficient avenue for precision treatment. This study provides curated sets of potential targets for diminishing or preventing the production of a large array of VFs. Reducing experimental costs and time to identify targets while simultaneously elucidating the underlying mechanisms by which targets inhibit infection are major contributions of our models to effective development of new therapies.

Our quantitative analysis of metabolic gene contribution to both growth and virulence is the most comprehensive genome-scale computational screen to date of virulence-related metabolism. Concerns regarding resistance to growth-targeting antibiotics in the context of multi-drug treatments may benefit from incorporation of new therapeutics that target VF synthesis^{53–55}. However, new proposals regarding sequential cycling of drugs with different mechanisms of action in an attempt to avoid drug resistance may favour drugs that inhibit VF production and growth simultaneously to maximize impact^{56,57}. Our mechanistic modelling approach allows us to predict the graded contribution of a given target gene to growth versus virulence systems to aid in these treatment designs. Ultimately, our updated models are valuable tools for quantitatively assessing relationships that would be challenging to interrogate experimentally at genome-scale. Our experimental validation of model predictions indicates that our approach provides testable hypotheses of gene function that can be used to elucidate critical interactions that may inform development of ‘resistance-resistant’ therapeutics.

Methods

Metabolic network reconstruction. Previously published iterations of *P. aeruginosa* PAO1 GEMs iMO1056 (ref. 19) and iMO1086 (ref. 20) were both used as resources during reconstruction efforts. iMO1056 was created using field-standard syntax consistent with many models in the BiGG database, while iMO1086 was built using the ToBiN platform which is not currently available^{19,20}. Since these original models were published, the modelSEED has become a favoured draft reconstruction resource, and offers a comprehensive database of balanced reactions and metabolites referenced from KEGG and MetaCyc from which hundreds of draft models have been created for use within the modelling community⁵⁸. In light of this, we used a draft conversion of iMO1056 to model SEED format as the starting point for our GEM update to enable consistency with our past *P. aeruginosa* models, improve annotation of model reactions and metabolites (KEGG IDs, E.C. numbers, pathway assignments) and enable easy comparison with a large collection of models created and curated by other groups^{12,59–61}. Because the conversion was an automated step performed by the modelSEED in an early iteration of the SEED database, manual curation was performed to add additional species-specific reactions that did not successfully convert from the original iMO1056 model or were present in iMO1086 as well as to correct conversion errors in reaction stoichiometry, directionality and GPR assignments. Further updates to SEED reactions and metabolite names using the modelSEED database were implemented to ensure consistency, and a KEGG subsystem assignment was added to each reaction when possible⁶².

The genomic contents of *P. aeruginosa* PAO1 and *P. aeruginosa* PA14 and two closely related pathogens from the *Burkholderia cepacia* complex were compared to assist development of new, reconciled GEMs for each strain from previously built models. *P. aeruginosa* PA14 is a primary clinical isolate that is used as a model strain due to its substantial virulence in a variety of hosts, while *P. aeruginosa* PAO1, a wound isolate, is the main reference strain of this species⁴⁸. We used

Burkholderia species specifically because of their similarities to *Pseudomonas* as opportunistic Gram-negatives that also chronically infect cystic fibrosis patients and share similar virulence mechanisms. We also previously built and extensively curated GEMs for these species in modelSEED syntax as described further below, making them useful resources. *P. aeruginosa* PA14, *P. aeruginosa* PAO1, *Burkholderia cenocepacia* J2315 and *Burkholderia multivorans* ATCC17616 were compared using genome-scale reciprocal BLASTP with an E-value cutoff of 0.01 with no low-complexity filter using CLC Main Workbench (CLC bio, Aarhus, Denmark). Hits with E-values below 1E-40 were considered high confidence hits and automatically matched. Genes with hits that received a higher E-value score were manually evaluated based on predicted function, gene descriptions and PseudoCAP category (custom system/pathway annotations) on the PGD (ref. 24) before inclusion in the models in the few cases they were employed. There is a distinction, however, between confident gene matches between the organisms used, and utilization of genes annotated as hypothetical proteins in all species. We propose functions for a number of hypothetical proteins in the models, some of which are missing a specific functional annotation only in PA14 versus PAO1, and some of which are hypothetical proteins in both strains which we have utilized in the models based on functional domain associations and other predictions provided by PGD annotations, literature and manual curation based on BLAST results against other species. Many of these hypothetical proteins are implemented in transport reactions, fatty acid and lipid pathways, and VF pathways. A table of these low-confidence gene assignments and hypothetical proteins to which we assigned functions in the model is provided in Supplementary Data 1.

Updated, strain-specific biomass formulas were created using a field standard approach that approximates the biomass composition by accounting for DNA, RNA, protein, cell wall components, lipids and organism-specific compounds whose production is required for growth^{9,12}. This effort expanded the number of components considered necessary for growth according to an improved biomass formulation and an updated search of literature pertaining to *Pseudomonas* species. Additional *Pseudomonas*-specific requirements, such as preference for ubiquinone-9 versus ubiquinone-8 as a key cofactor in respiration⁶³, were implemented. More specific lipids were implemented using recent studies from literature and as enabled by the expanded lipid reactions used in modelSEED draft reconstructions. Specifically, while iMO1056 and iMO1086 accounted for simple representations of cardiolipin, phosphatidylethanolamine, phosphatidylglycerol and phosphatidylserine, we implemented the specific saturated, unsaturated and cyclopropane fatty acids making up the full lipid profile of *P. aeruginosa* as described in the literature^{64–67}. We provide further details of the new biomass formulations at the following website in conjunction with the model files in spreadsheet and SBML format (<http://bme.virginia.edu/csbl/Downloads1-pseudomonas.html>).

To fill gaps and improve predictions, additional model components were first derived from iMO1086 and recently published GEMs of *Burkholderia* species. We built on prior curation efforts while maintaining consistent modelSEED syntax to enable future cross-species comparisons and community modelling. If SEED reactions in *Burkholderia* models were not present in the new *Pseudomonas* SEED model, the high confidence BLASTP results were used in conjunction with the PGD and *Burkholderia* Genome Database⁶⁸ and literature to evaluate addition of these reactions. Many of the new reactions were added to increase the number of Biolog carbon sources accounted for in the *Pseudomonas* models (from only the PM1 substrate set to both PM1 and PM2a substrate sets); this effort was guided by previous work we performed for the highly catabolically flexible *Burkholderia*¹². We also implemented new VF synthesis pathways using similar *Burkholderia* pathways as a guide. Other new reactions were added to expand lipid metabolism pathways using literature regarding *Pseudomonas*-specific lipid composition and the increased specificity of SEED reactions in this subsystem. Reactions implemented in the other well-curated SEED model available during our build work, *B. subtilis* iBsu1103, as well as reactions included in the MetaCyc and MetRxn databases were also used as resources^{59,69,70}. PAO1 and PA14 genes categorized as linked to virulence via data from experimental studies incorporated into the PGD v3 (ref. 24) were specifically evaluated for inclusion in the models to expand clinically relevant functional prediction ability (Supplementary Data 2 and Methods—Screen and Database Assembly).

Model validation. Models were validated using new, comprehensive assessments of experimental data from genome-scale transposon libraries and carbon utilization screening. Similar data had been used with prior models, but unexpected discrepancies identified in comparisons between PAO1 and PA14 measurements motivated careful re-assessment of data sets and experimental confirmation of results.

Gene essentiality predictions were performed by *in silico* deletions of single genes while optimizing for production of biomass using flux balance analysis via the COBRA Toolbox⁷¹. Predicted essential genes were compared with a list of genes that were not successfully targeted by transposon insertions in both genome-scale transposon insertion libraries of *P. aeruginosa* PAO1 (ref. 25) and *P. aeruginosa* PA14 (ref. 26). By using genes lacking transposon insertions in both studies, which used different transposon systems and resulted in differing levels of insertion rate and genome coverage, we increased our confidence that these genes were truly essential for growth in rich media for *P. aeruginosa* strains. Curation

with essentiality data resulted in improved prediction accuracy of gene essentiality via curated GPR relationships as well as the addition of new components to the biomass formula.

Single carbon source catabolic ability of the strains was predicted by providing a single carbon source and salts to the model via exchange constraints and optimizing for biomass production using flux balance analysis¹². Carbon utilization data were compiled from literature for both PAO1 and PA14, but discrepancies between studies motivated us to perform our own growth screens for both strains using Biolog phenotype arrays PM1 and PM2a. Growth curve screens were performed in triplicate using a microplate reader with shaking at 37 °C for 48 h. Curves were evaluated to identify substrates enabling growth versus no growth¹². Results guided specific curation of catabolic pathways and expansion of transport systems included in the model to improve prediction accuracy.

Screen and database assembly. Information on virulence-linked genes was compiled from the PGD (ref. 24) (current as of February 2016) using the Annotations by Category tool that provides Virulence Factor Annotation lists for several strains. We used the lists for PAO1 and PA14, which provided 427 and 208 genes, respectively, which were culled by the PGD from experimental screens in many different infection models, the Virulence Factor Database, and the Victors database as indicated in Supplementary Data 2. The bias towards PAO1 is partly due to more screens and studies performed for PAO1 versus PA14 in the literature; however, 419 of these genes are present in both genomes. We assumed that many of the genes identified as virulence-linked in PAO1 could also be virulence-linked in PA14; however, virulence-linked genes truly active in only one strain would be of interest to track in future work; these genes must then have alternate functions in addition to a role in virulence. Nevertheless, building on the above assumption, we created a combined list of genes associated with virulence that included any genes noted in either list which were present in both genomes to which we then added strain-specific virulence genes. The resulting lists included 432 and 441 plausible virulence-linked genes for PAO1 and PA14, respectively; the strain in which each gene was originally classified as a VF is also indicated as well as its presence in each model in Supplementary Data 2.

For the Tn-seq-based essential gene analysis, we obtained gene essentiality data from eight recently published Tn-seq screens for PA14 and PAO1 in a variety of culturing conditions. These screens are listed in Table 2 and are identified by the name of the paper's senior author (Pier¹⁶, Whiteley¹⁷ and Manoil¹⁸), strain, and media condition. For the PA14 Pier dataset¹⁶, we used the essential genes identified in Supplementary Table 1 of the original manuscript. Similarly, for the PAO1 and PA14 Whiteley data sets¹⁷, we used the essential genes identified in Data set S1 and Data set S3 of the original manuscript, respectively. For the PAO1 Manoil data sets¹⁸, we curated the 'General essential genes' identified in Data set S1 of the original manuscript to determine essential genes for each of the three media conditions studied: LB, sputum and pyruvate. Specifically, we applied a cutoff such that if a mutant for a particular gene failed to be generated in at least one of the independent transposon mutant pools for a particular media condition, that gene was deemed essential for that media condition. This approach does not take into account the location of the transposon insertion and, thus, may miss some essential genes. In the end, we obtained eight unique lists of essential genes for either PAO1 or PA14 in different media conditions based on the Pier, Whiteley and Manoil data sets. Once we obtained these lists of the essential genes identified in each screen, we compared them individually to the list of virulence-linked genes from the PGD database for either PA14 or PAO1 as appropriate. Genes that were in both a particular screen's essential list and the virulence-linked list were categorized as virulence-linked essential genes for that particular screen.

Prediction of virulence-related production versus growth. VF production capacity was first evaluated by optimizing the flux through an artificial 'demand' reaction for each virulence-related metabolite. Single gene deletions were implemented by identifying reactions for which a given gene was essential via the model's Boolean relationships and then constraining the flux through each of these reactions to zero. The effect of each of these deletions was evaluated by predicting production levels of each VF and biomass separately; resulting production levels lower than 0.001 were categorized as completely inhibitory (that is, the deleted gene is essential for production of that component). Production levels were normalized by maximum possible production of a component under wild-type conditions for comparison within VFs.

Network visualization. mPA14 was visualized using a command line implementation of MetDraw⁷² that enables colour overlay which was then edited in Inkscape (<https://inkscape.org/en/>).

Strains and growth conditions. Wild-type strains of *P. aeruginosa* PAO1 and PA14 and PA14 single gene knock-out mutants from the PA14 non-redundant genome-scale transposon library²⁶ were grown in LB media supplemented with 15 µg ml⁻¹ gentamycin as necessary at 37 °C with aeration for liquid cultures.

Pyoverdine assay. To measure pyoverdine production, strains were grown in synthetic cystic fibrosis media²⁹ for 24 h in 50 ml flasks and the absorbance of culture supernatants was measured at 405 nm according to a previously published protocol⁷³. All measurements were normalized to culture density as determined by the absorbance of the bacterial culture at 600 nm.

Data availability. The new metabolic network reconstructions for *P. aeruginosa* PAO1 and PA14, iPae1146 and iPau1129, respectively, are provided in spreadsheet format (Supplementary Data 3 and 4) that includes curation notes and SBML file format at our lab website (<http://bme.virginia.edu/csbl/Downloads1-pseudomonas.html>). The lab website also includes files detailing the development of strain-specific biomass formulas. Experimental data sets are available by request to the corresponding author.

References

- Allen, R. C., Popat, R., Diggle, S. P. & Brown, S. P. Targeting virulence: can we make evolution-proof drugs? *Nat. Rev. Microbiol.* **12**, 300–308 (2014).
- Centers for Disease Control and Prevention (US) *et al.* *Antibiotic resistance threats in the United States*. <https://stacks.cdc.gov/view/cdc/20705> (2013).
- Clatworthy, A. E., Pierson, E. & Hung, D. T. Targeting virulence: a new paradigm for antimicrobial therapy. *Nat. Chem. Biol.* **3**, 541–548 (2007).
- Rasko, D. A. & Sperandio, V. Anti-virulence strategies to combat bacteria-mediated disease. *Nat. Rev. Drug Discov.* **9**, 117–128 (2010).
- Konings, A. F. *et al.* *Pseudomonas aeruginosa* uses multiple pathways to acquire iron during chronic infection in cystic fibrosis lungs. *Infect. Immun.* **81**, 2697–2704 (2013).
- Wilder, C. N., Diggle, S. P. & Schuster, M. Cooperation and cheating in *Pseudomonas aeruginosa*: the roles of the las, rhl and pqs quorum-sensing systems. *ISME J.* **5**, 1332–1343 (2011).
- de Lorenzo, V. *Pseudomonas aeruginosa*: the making of a pathogen. *Environ. Microbiol.* **17**, 1–3 (2015).
- Price-Whelan, A., Dietrich, L. E. P. & Newman, D. K. Pyocyanin alters redox homeostasis and carbon flux through central metabolic pathways in *Pseudomonas aeruginosa* PA14. *J. Bacteriol.* **189**, 6372–6381 (2007).
- Thiele, I. & Palsson, B. Ø. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* **5**, 93–121 (2010).
- Chavali, A. K., D'Auria, K. M., Hewlett, E. L., Pearson, R. D. & Papin, J. A. A metabolic network approach for the identification and prioritization of antimicrobial drug targets. *Trends Microbiol.* **20**, 113–123 (2012).
- Schellenberger, J. *et al.* Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc.* **6**, 1290–1307 (2011).
- Bartell, J. A., Yen, P., Varga, J. J., Goldberg, J. B. & Papin, J. A. Comparative metabolic systems analysis of pathogenic *Burkholderia*. *J. Bacteriol.* **196**, 210–226 (2014).
- Rodríguez-Rojas, A., Oliver, A. & Blázquez, J. Intrinsic and environmental mutagenesis drive diversification and persistence of *Pseudomonas aeruginosa* in chronic lung infections. *J. Infect. Dis.* **205**, 121–127 (2012).
- Curcio, D. Multidrug-resistant gram-negative bacterial infections: are you ready for the challenge? *Curr. Clin. Pharmacol.* **9**, 27–38 (2014).
- Falkinham, J. O., Hilborn, E. D., Arduino, M. J., Pruden, A. & Edwards, M. A. Epidemiology and ecology of opportunistic premise plumbing pathogens: *Legionella pneumophila*, *Mycobacterium avium*, and *Pseudomonas aeruginosa*. *Environ. Health Perspect.* **123**, 749–758 (2015).
- Skurnik, D. *et al.* A comprehensive analysis of *in vitro* and *in vivo* genetic fitness of *Pseudomonas aeruginosa* using high-throughput sequencing of transposon libraries. *PLoS Pathog.* **9**, e1003582 (2013).
- Turner, K. H., Wessel, A. K., Palmer, G. C., Murray, J. L. & Whiteley, M. Essential genome of *Pseudomonas aeruginosa* in cystic fibrosis sputum. *Proc. Natl Acad. Sci. USA* **112**, 4110–4115 (2015).
- Lee, S. A. *et al.* General and condition-specific essential functions of *Pseudomonas aeruginosa*. *Proc. Natl Acad. Sci. USA* **112**, 5189–5194 (2015).
- Oberhardt, M. A., Puchałka, J., Fryer, K. E., Martins dos Santos, V. A. P. & Papin, J. A. Genome-scale metabolic network analysis of the opportunistic pathogen *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* **190**, 2790–2803 (2008).
- Oberhardt, M. A., Puchałka, J., Martins dos Santos, V. A. P. & Papin, J. A. Reconciliation of genome-scale metabolic reconstructions for comparative systems analysis. *PLoS Comput. Biol.* **7**, e1001116 (2011).
- Ballok, A. E. & O'Toole, G. A. Pouring salt on a wound: *Pseudomonas aeruginosa* virulence factors alter Na⁺ and Cl⁻ flux in the lung. *J. Bacteriol.* **195**, 4013–4019 (2013).
- Williams, P. & Cámara, M. Quorum sensing and environmental adaptation in *Pseudomonas aeruginosa*: a tale of regulatory networks and multifunctional signal molecules. *Curr. Opin. Microbiol.* **12**, 182–191 (2009).

23. Maspoli, A., Wenner, N., Mislin, G. L. A. & Reimmann, C. Functional analysis of pyochelin-/enantiopyochelin-related genes from a pathogenicity island of *Pseudomonas aeruginosa* strain PA14. *BioMetals* **27**, 559–573 (2014).
24. Winsor, G. L. *et al.* Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Res.* **44**, gkv1227 (2016).
25. Jacobs, M. A. *et al.* Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *Proc. Natl Acad. Sci. USA* **100**, 14339–14344 (2003).
26. Liberati, N. T. *et al.* An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. *Proc. Natl Acad. Sci. USA* **103**, 2833–2838 (2006).
27. Feinbaum, R. L. *et al.* Genome-wide identification of *Pseudomonas aeruginosa* virulence-related genes using a *Caenorhabditis elegans* infection model. *PLoS Pathog.* **8**, e1002813 (2012).
28. Hilker, R. *et al.* Interclonal gradient of virulence in the *Pseudomonas aeruginosa* pangenome from disease and environment. *Environ. Microbiol.* **17**, 29–46 (2014).
29. Palmer, K. L., Aye, L. M. & Whiteley, M. Nutritional cues control *Pseudomonas aeruginosa* multicellular behavior in cystic fibrosis sputum. *J. Bacteriol.* **189**, 8079–8087 (2007).
30. Folkesson, A. *et al.* Adaptation of *Pseudomonas aeruginosa* to the cystic fibrosis airway: an evolutionary perspective. *Nat. Rev. Microbiol.* **10**, 841–851 (2012).
31. Uehlinger, S. *et al.* Identification of specific and universal virulence factors in *Burkholderia cenocepacia* strains by using multiple infection hosts. *Infect. Immun.* **77**, 4102–4110 (2009).
32. Byrd, M. S. *et al.* Direct evaluation of *Pseudomonas aeruginosa* biofilm mediators in a chronic infection model. *Infect. Immun.* **79**, 3087–3095 (2011).
33. Turner, K. H., Everett, J., Trivedi, U., Rumbaugh, K. P. & Whiteley, M. Requirements for *Pseudomonas aeruginosa* acute burn and chronic surgical wound infection. *PLoS Genet.* **10**, e1004518 (2014).
34. Hunter, R. C. *et al.* Phenazine content in the cystic fibrosis respiratory tract negatively correlates with lung function and microbial complexity. *Am. J. Respir. Cell Mol. Biol.* **47**, 738–745 (2012).
35. Yeterian, E. *et al.* Synthesis of the siderophore pyoverdine in *Pseudomonas aeruginosa* involves a periplasmic maturation. *Amino Acids* **38**, 1447–1459 (2010).
36. Yang, L., Nilsson, M., Gjermansen, M., Givskov, M. & Tolker-Nielsen, T. Pyoverdine and PQS mediated subpopulation interactions involved in *Pseudomonas aeruginosa* biofilm formation. *Mol. Microbiol.* **74**, 1380–1392 (2009).
37. Meyer, J. M. *et al.* Pyoverdine is essential for virulence of *Pseudomonas aeruginosa*. These include: pyoverdine is essential for virulence of *Pseudomonas aeruginosa*. *Infect. Immun.* **64**, 518–523 (1996).
38. Pepper, J. W. Defeating pathogen drug resistance: guidance from evolutionary theory. *Evolution (NY)* **62**, 3185–3191 (2008).
39. Pepper, J. W. Drugs that target pathogen public goods are robust against evolved drug resistance. *Evol. Appl.* **5**, 757–761 (2012).
40. Eadsforth, T. C. *et al.* Assessment of *Pseudomonas aeruginosa* N5,N10-methylenetetrahydrofolate dehydrogenase—cyclohydrolase as a potential antibacterial drug target. *PLoS ONE* **7**, e35973 (2012).
41. Eadsforth, T. C. *et al.* Characterization of 2,4-diamino-6-oxo-1,6-dihydro-pyrimidin-5-yl ureido based inhibitors of trypanosoma brucei FdD and testing for antiparasitic activity. *J. Med. Chem.* **58**, 7938–7948 (2015).
42. Eadsforth, T. C., Maluf, F. V. & Hunter, W. N. *Acinetobacter baumannii* FdD ligand complexes—potent inhibitors of folate metabolism and a re-evaluation of the structure of LY374571. *FEBS J.* **279**, 4350–4360 (2012).
43. Vandenberg, C. S., Vlasschaert, M. & Seah, S. Y. K. Functional characterization of an aminotransferase required for pyoverdine siderophore biosynthesis in *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* **186**, 5596–5602 (2004).
44. van Opijnen, T. & Camilli, A. Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. *Nat. Rev. Microbiol.* **11**, 435–442 (2013).
45. Saenz, H. L. & Dehio, C. Signature-tagged mutagenesis: technical advances in a negative selection method for virulence gene identification. *Curr. Opin. Microbiol.* **8**, 612–619 (2005).
46. Mazurkiewicz, P., Tang, C. M., Boone, C. & Holden, D. W. Signature-tagged mutagenesis: barcoding mutants for genome-wide screens. *Nat. Rev. Genet.* **7**, 929–939 (2006).
47. Andersen, S. B., Marvig, R. L., Molin, S., Krogh Johansen, H. & Griffin, A. S. Long-term social dynamics drive loss of function in pathogenic bacteria. *Proc. Natl Acad. Sci. USA* **112**, 10756–10761 (2015).
48. Lee, D. G. *et al.* Genomic analysis reveals that *Pseudomonas aeruginosa* virulence is combinatorial. *Genome Biol.* **7**, R90 (2006).
49. Fletcher, M. P., Diggle, S. P., Cámara, M. & Williams, P. Biosensor-based assays for PQS, HHQ and related 2-alkyl-4-quinolone quorum sensing signal molecules. *Nat. Protoc.* **2**, 1254–1262 (2007).
50. Bjarnsholt, T. *et al.* *In vitro* screens for quorum sensing inhibitors and *in vivo* confirmation of their effect. *Nat. Protoc.* **5**, 282–293 (2010).
51. Valentine, C. D. *et al.* Small molecule screen yields inhibitors of *Pseudomonas* homoserine lactone-induced host responses. *Cell. Microbiol.* **16**, 1–14 (2013).
52. Ross-Gillespie, A., Weigert, M., Brown, S. P. & Kümmerli, R. Gallium-mediated siderophore quenching as an evolutionarily robust antibacterial treatment. *Evol. Med. Public Heal.* **2014**, 18–29 (2014).
53. Chait, R., Craney, A. & Kishony, R. Antibiotic interactions that select against resistance. *Nature* **446**, 668–671 (2007).
54. Torella, J. P., Chait, R. & Kishony, R. Optimal drug synergy in antimicrobial treatments. *PLoS Comput. Biol.* **6**, e1000796 (2010).
55. Ejim, L. *et al.* Combinations of antibiotics and nonantibiotic drugs enhance antimicrobial efficacy. *Nat. Chem. Biol.* **7**, 348–350 (2011).
56. Imamovic, L. & Sommer, M. O. A. Use of collateral sensitivity networks to design drug cycling protocols that avoid resistance development. *Sci. Transl. Med.* **5**, 204ra132 (2013).
57. Kim, S., Lieberman, T. D. & Kishony, R. Alternating antibiotic treatments constrain evolutionary paths to multidrug resistance. *Proc. Natl Acad. Sci. USA* **111**, 14494–14499 (2014).
58. Henry, C. S. *et al.* High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* **28**, 977–982 (2010).
59. Henry, C. S., Zinner, J. F., Cohoon, M. P. & Stevens, R. L. iBsu1103: a new genome-scale metabolic model of *Bacillus subtilis* based on SEED annotations. *Genome Biol.* **10**, R69 (2009).
60. Freilich, S. *et al.* Competitive and cooperative metabolic interactions in bacterial communities. *Nat. Commun.* **2**, 589 (2011).
61. Devoid, S. *et al.* Automated genome annotation and metabolic model reconstruction in the SEED and Model SEED. *Methods Mol. Biol.* **985**, 17–45 (2013).
62. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40**, D109–D114 (2012).
63. Yamada, Y., Takinami-nakamura, H., Tahara, Y., Oyaizu, H. & Komagata, K. The ubiquinone systems in the strains of *Pseudomonas* species. *J. Gen. Appl. Microbiol.* **28**, 7–12 (1982).
64. Zhu, K., Choi, K.-H., Schweizer, H. P., Rock, C. O. & Zhang, Y.-M. Two aerobic pathways for the formation of unsaturated fatty acids in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **60**, 260–273 (2006).
65. Hoffmann, N., Steinbüchel, A. & Rehm, B. H. The *Pseudomonas aeruginosa* phaG gene product is involved in the synthesis of polyhydroxyalkanoic acid consisting of medium-chain-length constituents from non-related carbon sources. *FEMS Microbiol. Lett.* **184**, 253–259 (2000).
66. Chao, J., Wolfaardt, G. M. & Arts, M. T. Characterization of *Pseudomonas aeruginosa* fatty acid profiles in biofilms and batch planktonic cultures. *Can. J. Microbiol.* **56**, 1028–1039 (2010).
67. Bayssse, C. & O’Gara, F. in *Pseudomonas*. Vol. 5 (eds Ramos, J.L. & Filloux, A.) 193–224 (Springer, 2007).
68. Winsor, G. L. *et al.* The Burkholderia Genome Database: facilitating flexible queries and comparative analyses. *Bioinformatics* **24**, 2803–2804 (2008).
69. Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res.* **42**, D471–D480 (2014).
70. Kumar, A., Suthers, P. F. & Maranas, C. D. MetRxn: a knowledgebase of metabolites and reactions spanning metabolic models and databases. *BMC Bioinform.* **13**, 6 (2012).
71. Becker, S. A. *et al.* Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat. Protoc.* **2**, 727–738 (2007).
72. Jensen, P. A. & Papin, J. A. MetDraw: automated visualization of genome-scale metabolic network reconstructions and associated high-throughput data. *Bioinformatics* **30**, 1–2 (2014).
73. Imperi, F., Tiburzi, F. & Visca, P. Molecular basis of pyoverdine siderophore recycling in *Pseudomonas aeruginosa*. *Proc. Natl Acad. Sci. USA* **106**, 20440–20445 (2009).
74. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40**, 109–114 (2011).

Acknowledgements

We thank Heath Damron, Mariette Barbier and Glynis Kolling for discussions about virulence factor assays as well as Edik Blais for discussions regarding modelling. Support for this project was provided by the National Institute of General Medical Sciences (GM088244 and GM108501) and the Whitaker Foundation.

Author contributions

J.A.B., A.S.B. and J.A.P. conceived and designed the study. J.A.B. and A.S.B. reconstructed and curated the models with assistance from P.Y. and J.C.T. J.A.B. and A.S.B. conducted all simulations and analyses. A.S.B. performed the experiments. J.B.G. and L.J. provided commentary and revision for the manuscript written by J.A.B., A.S.B., and J.A.P.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Bartell, J. A. *et al.* Reconstruction of the metabolic network of *Pseudomonas aeruginosa* to interrogate virulence factor synthesis. *Nat. Commun.* **8**, 14631 doi: 10.1038/ncomms14631 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017