Synergy and antagonism in a genome-scale model of metabolic hijacking by bacteriophage

Jordan C. Rozum[1†], William Sineath[1,2†], Doo Nam Kim[1], Pavlo Bohutskyi[1], Connah Johnson[3], James Evans[4,5], David Pollock[6,1], Wei-Jun Qian[1], Margaret S. Cheung[4,7], Ruonan Wu[1*], Song Feng[1*]

[1]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA.

[2]Department of Biological Engineering, Utah State University, Logan, UT, USA.

[3]Advanced Computing, Mathematics, and Data Division, Pacific Northwest National Laboratory, Richland, WA, USA.

[4]Environmental Molecular Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA.

[5]School of Biological Sciences, Washington State University, Pullman, WA, USA.

[6]Department of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, Aurora, CO, USA.

[7]Department of Physics, University of Washington, Seattle, WA, USA.

[*]Corresponding author. Email: song.feng@pnnl.gov (SF); ruonan.wu@pnnl.gov (RW)

[†]These authors contributed equally to this work.

ABSTRACT

Bacteriophage auxiliary metabolic genes (AMGs) alter host metabolism by hijacking reactions. Previous studies used functional annotations to infer AMG impacts, but how these propagate and interact to affect global metabolism and phage production remains understudied. We demonstrate the first integration of AMGs and phage assembly into a genome-scale metabolic model, using P-HM2 infection

of *Prochloroccocus marinus* MED4 as a test system. Our model predicts that 17 directly hijacked reactions substantially impact over 30% of the metabolism, including carbon fixation, photosynthesis, and nucleotide synthesis. These indirect impacts are synergistically and antagonistically coupled. We find that hijacked reactions are either phage-aligned–shifting feasible reaction velocities in accordance with maximal phage production–or phage-antialigned. Phage-alignment predicts changes to the host-phage biomass Pareto front: phage-aligned reactions limit host growth; phage-antialigned reactions do not. We provide a systems-level understanding of AMG perturbations, revealing nontrivial cascading effects and highlighting their role in shaping microbial functions.

Introduction

The diverse genotypes of co-evolved hosts and phages make predicting their interactions challenging. Understanding the dynamics of host-phage interactions is crucial for predicting and controlling host response and viral replication, because viruses, including bacteriophages, rely heavily on their host organisms for replication. During replication, viruses hijack the host's cellular machinery to synthesize the structural proteins and genomic materials needed to assemble new viral particles. A key aspect of viral replication is acquiring necessary building blocks, such as amino acids and nucleotides (*1, 2*). Additionally, viral replication has significant energy and redox demands, including molecules like ATP and NAD(P)H (*3*). Shifts in host metabolism are needed to meet energy, redox, and material requirements critical for the successful proliferation of viruses. These changes are not merely theoretical but are supported by experimental evidence (*4–6*). For instance, infected cells often exhibit increased uptake and usage of nutrients to meet the demands of viral replication (*7, 8*). This can significantly disrupt host metabolism, leading to observable changes in not only the infected cell's physiology (*9*) but also in the chemistry of the host's surrounding environment (*10, 11*).

Previous research has demonstrated that phages require specific nutrient ratios to maximize their replication efficiency (*12*). Thus, the uptake rate of specific nutrients in the host cells can be a limiting factor in phage replication, particularly when phages infect their hosts in environments that

are optimal for the hosts. In such settings, the competition for essential nutrients intensifies. Marine cyanophages provide an illustrative example of this nutrient limitation. These phages carry the *phoH* gene, which enhances phosphate uptake. This adaptation is beneficial for the phage because its biomass contains a significantly higher phosphorous ratio than the cyanobacterial host biomass (*12*). The conservation of the *phoH* phage gene suggests that phosphate availability is critical for successful cyanophage replication, highlighting a specific nutrient demand that is less pronounced in their hosts. These findings underscore the adaptive strategies phages employ to redirect host metabolism so that the phage can thrive in nutrient-limited environments. Understanding how this redirection occurs can provide deeper insights into phage-host interactions and the evolutionary pressures that shape them in various ecological contexts.

Interactions between host and phage are shaped by evolutionary pressures on both host and virus (*13*). This co-evolution results in a dynamic landscape in which ecosystem-scale effects are coupled to molecular interactions and where understanding global (i.e., organism-level) effects of specific molecular interactions is paramount. The interaction between viral nonstructural proteins (i.e., those that are not structural components of virus particles) and host components is one venue for such co-evolution. Nonstructural proteins can modulate or reprogram the host's innate immune responses (*14*), stress responses (*15*), and metabolic functions (*16*). In bacteriophage, genes that encode the nonstructural proteins thought to hijack metabolic reactions (e.g., by catalysis or sequestration) are called *auxiliary metabolic genes* (AMGs).

AMGs are widely studied, especially in cyanophage-cyanobacteria systems. Understanding the metabolic hijacking process in these systems has significant biotechnology implications, including the development of novel strategies for controlling harmful cyanobacterial blooms (*17*), developing phage therapy against antibiotic-resistant bacteria (*18*), and industrial bioproduction (*19*). Recent work has shown that bioproduction by *S. elongatus* PCC 7942 is improved by altering its metabolism through expression of cyanophage-derived sigma factors, demonstrating the potential of phage-inspired approaches for metabolic engineering (*19*).

Marine viral AMGs are generally classified into two groups: Class I AMGs encode enzymes involved in core cellular metabolic processes, including photosynthesis, carbon and amino acid metabolism; while Class II AMGs encode proteins with peripheral functionalities, including transport, iron-cluster assembly, cofactor synthesis, and stress response (e.g., chaperones) (*20–22*).

To ensure sufficient energy for replication, phage-encoded AMGs modulate host energy production by remodeling photosynthetic light-harvesting and electron flow components through Class I AMGs such as the genes *psbA* and *psbD*, which encode the core photosystem II reaction center proteins D1 and D2 (*23*). Additional genes contributing to optimized photosynthetic energy production include *cpeT*, *ho1*, and *pebS*, which are involved in the synthesis of phycobilisome light-harvesting complexes (*24*). High light-inducible proteins (hlip) protect photosynthetic complexes and membranes from photo-damage by dispersing excessive light energy (*23*). Furthermore, phage-encoded Class I AMGs *cp12* and *talC* redirect energy and carbon metabolism toward dNTP production by suppressing CO2 fixation (Calvin cycle) and activating the pentose phosphate pathway (PPP) (*25*). Adjustment of nutrient acquisition and metabolism is essential to ensure sufficient nitrogen and phosphorus for phage replication in nutrient-scarce environments. These adaptations involve Class II AMGs including the phosphate starvation-inducible protein phoH, as well as several AMGs that rebalance nitrogen compounds toward nucleotide biosynthesis, such as ribonucleoside-diphosphate reductase (*nrdA*), nucleotide pyrophosphohydrolase (*mazG*), and cobalamin synthase (*cobS*) (*26–28*). While these processes involving AMGs are well understood at a high level, there is no clear theoretical framework for exploring the effect of hijacking individual reactions or for how these effects propagate and couple to alter the host metabolic system.

Genome scale metabolic models (*29,30*) are developed for investigating metabolism-wide effects of perturbations to genes and reaction rates. Flux balance analysis (FBA) and related techniques provide a parameter-robust way to quantify the impact on the quasi-steady-state distributions of reaction velocities (fluxes) that are optimal for biomass production. Flux balance methods are especially useful in determining which reactions are essential for cell growth (*31*) and for assessing how metabolic flux redistributes in response to gene knockout-induced reaction stoppage (*29, 32*). Incorporating the effects of exogenous gene expression, however, remains a challenge. Some progress has been made by incorporating enzymatic constrains (*33*), but the necessary parameters can be difficult to measure or estimate, and even state-of-the-art machine learning methods require additional post-inference model-specific adjustment (*34*). To our knowledge, no widely accepted kinetics-free method exists to systematically probe the impact of exogenous expression of metabolic genes, such as can occur during phage infection. Additionally, using existing genome scale metabolic models (*35*) to study viral infection requires specification of the virus' metabolic requirements (*36*).

We are not aware of any genome scale metabolic model of host-bacteriophage interaction that incorporates both exogenous expression of metabolic genes and phage metabolic demands.

To address the knowledge gap, we developed a strategy to explore how cyanophages hijack *Prochloroccocus marinus* metabolism by expressing known AMGs. For this study, we used a previously published iJC568v7 flux balance model (*37*), which is a refinement of an earlier model (*38*). From the literature, we curated AMGs observed in the cyanophage P-HM2 (*39*) and extracted their predicted or measured effects on specific reactions. The identified reactions are catalyzed by enzymes that are either encoded by AMGs or sequestered by AMG-encoded proteins. We call such reactions *AMG-hijacked* and aim to understand their global impact on the host's quasi-steady-state metabolic flux distribution and viral replication using flux balance analysis (FBA) and related techniques. To achieve this, we constructed a phage biomass function from the P-HM2 genome and the stoichiometry of its structural proteins by adapting the procedure of (*36*). We integrated the P-HM2 biomass function into the iJC568v7 model of (*37*). Using this enhanced model, we examined which AMG-induced metabolic alterations may favor phage replication over host growth and identified several unique combinatorial effects of AMG-hijacked reaction modulation.

Our results highlight the complexity of host-phage interactions. At a macro-level, these interactions are driven by a web of biological requirements and evolutionary adaptations; at a micro-level, even setting aside transcriptional and post-translational effects, AMGs directly hijack multiple interdependent metabolic fluxes, which leads to global cascading effects on the host metabolism. In this study, our focus is on how AMGs affect the quasi-steady state distribution of metabolic reaction fluxes, both individually and combinatorially. Understanding the metabolic impacts of combinations of AMGs is a necessary step to probe the whole-cell response to viral infection and provide a more detailed characterization of phage infection dynamics.

1 Results

1.1 Integration of AMGs and P-HM2 phage biomass formation into a genome-scale model

From the literature (Table 1), we curated 14 AMGs present in the P-HM2 phage (*40*) and extracted the 17 reactions within the iJC568v7 model that they hijack using annotations from the KEGG database (*41*). The reaction expressions and the subsystems to which they belong are listed in Supplementary Table S1 and additional background information is provided in Supplementary Text. We also constructed a phage biomass reaction to capture the metabolic demands of phage replication by accounting for the nucleotide synthesis required to replicate the phage genome and the stoichiometric ratios of amino acids in phage structural proteins (see Methods). We computed that the DNA to structure protein mass ratio for the P-HM2 phage to be 0.66, and that the mRNA to structure protein mass ratio to be 0.022 (we assume the mRNA is not packed in the phage heads and is recycled by the host metabolism). When unconstrained by nutrient abundance, the maximum phage growth in this model is a linear rate of 0.151 grams per gDW host per hour. The maximum host growth rate in this model under the same conditions is an exponential growth rate of approximately 0.098 grams per gDW host per hour (in agreement with the model of (*38*) on which it is based). We compared the demands of phage biomass to those of host biomass in Supplementary Table S3. We determined which genes and reactions are required for host or phage biomass production in our model in the standard way using flux variability analysis (FVA; see Methods) (*42*): non-essential reactions are those that contain a flux value of zero within their FVA range when biomass is constrained to be greater than 1% of its maximum value. Consistent with (*37, 38*), we found that 372 of 595 genes (376 of 994 reactions) are essential for host growth [1]. In contrast, only 172 metabolic genes (101 reactions plus the phage biomass reaction itself) are essential for phage biomass production in this model. All phage-essential genes and reactions we identified are also essential for host growth (with the exception of the phage biomass reaction). This smaller number of genes and reactions that are essential to phage compared to host is partly due to the lack of lipids in the P-HM2 phage structure, meaning that much of the host's essential lipid metabolism

[1] Because a reaction can be mediated by several genes, and a gene can mediate several reactions, the number of essential reactions need not be equal to the number of essential genes.

is not needed for assembling P-HM2 phage particles. Of the 17 curated AMG-hijacked reactions, nine (PSIIabs[2], FAKEOrthophosphateEX[3], R01523, R02017, R02019, R02024, R05223, R05817, R05818) are essential for host biomass production, and six (PSIIabs, FAKEOrthophosphateEX, R02017, R02019, R02024, R01523) are essential for phage production.

Next, we simulated how metabolic flux redistributes to accommodate phage particle production and used this to determine target ranges for hijacked reaction fluxes. We conducted FVA on the iJC568v7 MED4 model (*37, 38*) in a nutrient-rich environment so that RuBisCO efficiency, rather than nutrient uptake, was the growth-limiting factor. We considered two independent scenarios: a pre-infection scenario in which host biomass is optimized, and a post-infection scenario in which phage biomass is optimized. We compared the ranges, for each AMG-hijacked reaction, that allow for at least 50% optimal host or phage biomass production (50% is the highest fraction of maximum growth simultaneously achievable by host and phage in a nutrient-rich environment). We identified one of the endpoints of each range, or zero, to be a target point for each hijacked reaction in accordance with the literature (Table 1). To permit a degree of flexibility when combinatorially hijacking reactions, we allow for a small (10% of the FVA range) tolerance in the target flux value. In subsequent results, we constrained individual AMG-hijacked reactions to fall within these target ranges (depicted visually in Figure 1). A majority (13/17) of the AMG-hijacked reaction FVA ranges are largely unaffected (greater than 50% overlap of intervals) by the choice to optimize for host or phage biomass. Indeed, one reaction, R01063 (hijacked by *cp12*) has a minimal impact on either host or phage biomass and its feasible range covers the entire region of the FVA search ($\pm$1000mmol gDW^{-1} h^{-1}). Exceptions include two reactions that are hijacked by *nrdA* and are involved in nucleotide metabolism. The responses of two reactions, R05818 (hijacked by *cpeT/cpcT*) and R05817 (hijacked by *pebS*) are counterintuitive. These reactions are both thought to have increased flux when their hijacking genes are expressed (*44, 48*); however, optimal flux through the phage biomass function instead requires lower flux through these two reactions. In these two instances, we have selected the high-flux portion of the phage-optimal range as the target range.

[2]PSIIabs represent photon absorption in photosystem II.

[3]FAKEOrthophosphateEX is the phosphate exchange reaction; the FAKE prefix indicates that it was added through gap filling by (*38*).

AMG name	Reactions	Direction	Reference	Note
cobS	R05223	Increase	(*43*)	Biosynthesis of cobalamin (vitamin B12), important for nucleotide metabolism.
cp12	R01063, R01523	Absolute Decrease	(*25*)	Calvin Benson cycle inhibitor by sequestering PRK and GAPDH enzymes.
cpeT/*cpcT*	R05818	Increase	(*44*)	Facilitates fast phycobiliprotein assembly during phage infection.
hlip	N/A	N/A	(*45*)	Host survival under moderately high light intensities (not used).
hsp20	N/A	N/A	(*46*)	Heat shock response protein (not used).
ho1	R00311	Increase	(*47*)	Phycobilin synthesis.
mazG	R00426, R00662, R01855, R02100	Increase	(*28*)	Proposed to be involved into host genome degradation to reuse building blocks for phage genome replication.
nrdA	R02017, R02018, R02019, R02024	R02018 Decrease; All others Increase	(*25*)	Codes for ribonucleotide reductase to promote dNTP production.
pebS	R05817	Increase	(*44*)	Phage gene *pebS* integrates functions of *pebA* and *pebB* involved in phycoerythrobilin biosynthesis.
phoH	FAKEOrthophsphateEX	Decrease	(*12*)	Promotes phosphate uptake (modeled implicitly in nutrient scenarios).
psbA/*psbD*	PSIIabs	Increase	(*48*)	The genes *psbA* and *psbD* code for D1 and D2 photosystem II proteins.
talC	R01827	Increase	(*25*)	Facilitates enhanced flux through pentose phosphate pathway for augmented production of NADPH and ribose 5-phosphate.

Table 1: **P-HM2 AMGs identified in the literature.** The AMG's name is given in the first column. The affected reaction ID is given in the Reaction column. In two cases (FAKEOrthophophateEX and PSIIabs), the genes impact meta-reactions in the model (related to phosphate exchange and photosystem II, respectively)–all other reactions are indicated by KEGG ID. The reported effect of the AMG on the affected reaction is noted in the Direction column.

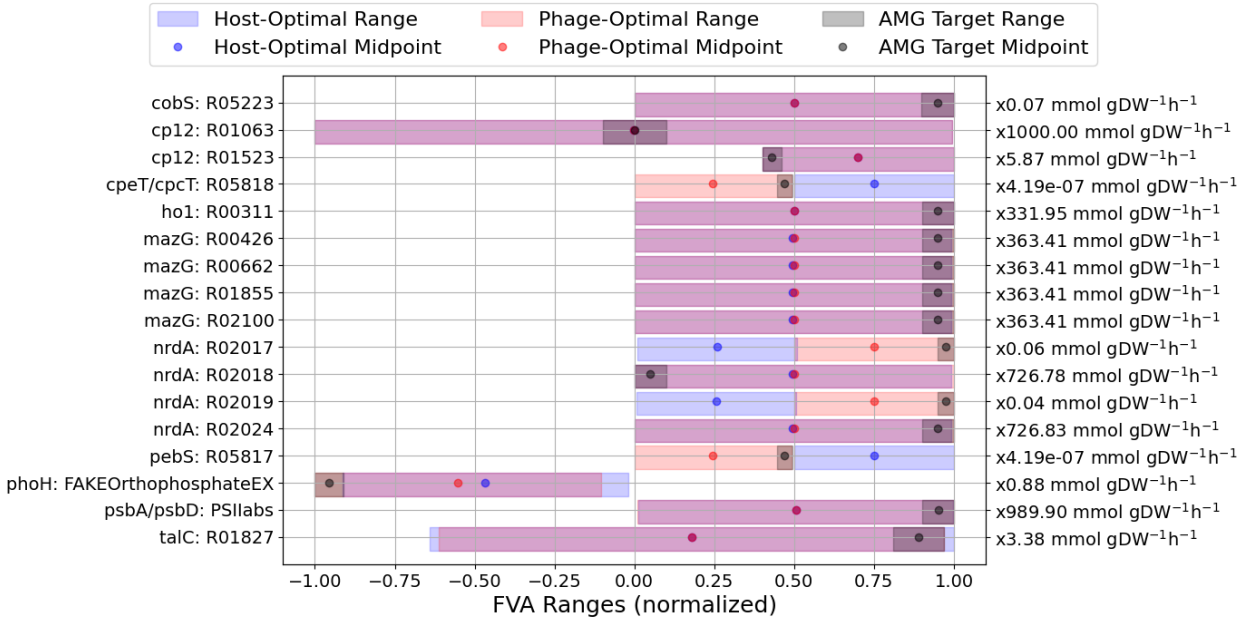

Figure 1: **FVA ranges for AMG-hijacked reactions.** Blue ranges indicate flux variability when maximizing host biomass while red ranges are for the variability associated with maximizing the phage biomass; any overlap is shown in purple. A decrease in optimal biomass production of up to 50% is permitted. Ranges are normalized by the most extreme feasible flux across the two scenarios (normalization factors depicted along the right-hand side of the figure). Reaction IDs and subsystems are given along the left-hand side. Using the predicted direction of each reaction flux change as reported in the literature (Table 1), we select 10% of the phage-optimal FVA range to designate as an AMG target range (black).

1.2 Global impact of hijacking a reaction either aligns or antialigns with phage production

We compared the FVA-informed impact of maximizing host or phage biomass on global host metabolism. We compared the FVA ranges for each reaction by computing the shift in FVA midpoint. The largest 30 FVA range shifts, normalized by the largest feasible flux magnitude for each reaction, are shown in Supplementary Figure S1. Affected reactions belong to many metabolic subsystems including transport and exchange, pyruvate synthesis, glycolysis/gluconeogenesis, carbon fixation, the citrate (TCA) cycle, and various amino acid, nucleotide, and lipid metabolic pathways, consistent with prior literature on the wide-ranging metabolic impacts of P-HM2 infection in MED4 (*25*). We found that ammonia and orthophosphate uptakes (AmmoniaTRANS, AmmoniaEX, OrthophosphateTRANS, and FAKEOrthophosphateEX reactions) are among the most sensitive to changes in host / phage biomass optimization: Optimal phage production required approximately double the rate of ammonia uptake and six times the rate of phosphate uptake required for optimal host growth.

To assess the interplay between phage biomass production and our set of curated AMGs, we fixed the flux of each AMG-hijacked reaction in turn to observe its downstream effects on other reactions when maximizing host growth. We set each AMG-hijacked reaction (one at a time) to the midpoint of its target range (see Figure 1, top panel). The FVA midpoints of 321 out of 995 reactions were shifted by more than 0.25mmol gDW^{-1} h^{-1} in at least one instance due to the fixed AMG-hijacked reaction flux. We compared the host-optimal FVA shifts due to AMG hijacking with the shifts induced by optimizing phage biomass in the absence of AMGs (Figure 2). Principal component analysis (PCA) identified three distinct categories of AMG-hijacked reactions, based on the loadings of the first principal component, $PC0$, which explains 78.7% of the variance among the 17 phage scenarios and the maximum phage biomass scenario. We have named these categories according to how similar their downstream impacts are to the maximum phage biomass scenario ($PC0 = 0.26$): phage-aligned ($PC0 > 0.1$; R05223, R01523, R05818, R02017, R02019, R05817, FAKEOrthophasphateEX, R01827), phage-antialigned ($PC0 < -0.1$; R00311, R00426, R00662, R01855, R02100, R012018, R02024, PSIIabs), and minimal-impact ($-0.1 \leq PC0 \leq 0.1$; R01063).

Hijacking a phage-aligned reaction tends to increase flux through other reactions in the model

when maximizing host biomass. In contrast, hijacking phage-antialigned reactions tend to have the opposite effect. Though this global trend holds in most subsystems of the model, we found that carbon fixation is an exception: hijacking phage-aligned reactions tend to decrease (rather than increase) flux through carbon fixation reactions, which are largely unaffected by phage-antialigned hijacking. Consistent with previous work (*25*), we found that *cp12* hijacking of R01523 redirects metabolic flux from carbon fixation and to the pentose phosphate pathway and nucleotide synthesis (purine and pyrimidine metabolism), and also that *talC* hijacking of R01827 promoted metabolic flux through the pentose phosphate pathway.

1.3 The trade-off between host and phage biomass is modulated by phage-aligned AMGs

In our model, phage-antialigned reactions tend to shift the range of impacted reaction fluxes to lower (or more negative) values, in contrast to phage production, which favors flux ranges with higher midpoints. However, there is no mathematical requirement that the resulting optimal ranges not overlap, in which case maximum phage growth is still be possible. By the same token, there is no mathematical requirement that phage-aligned reactions not inhibit phage replication. To assess whether these mathematical possibilities are realized, we examined the impacts of hijacking phage-aligned and phage-antialigned reactions on the host-phage biomass trade-off. First, we constructed the base-line Pareto front for the multi-objective maximization of the phage and host biomass under various nutrient stresses (e.g., N and P availability) and without hijacking any reactions (left panel of Figure 3).

Based on the insights from previous studies (*12, 37, 38*) and the large change increase in N and P uptake rates we observed when maximizing phage biomass (Figure S1), we selected N and P availability as the two nutrient stressors. Consistent with these results, both phage and host growth rates are compromised in the model when N or P is limited (compare solid and dashed or dotted lines in the leftmost panel of Figure 3). However, maximum phage growth is more severely impacted in the model: in the moderate N limit scenario (90% ammonia uptake compared to host-optimal) maximum phage growth is 52% lower, while maximum host growth is only 9% lower. Similarly, in the moderate P limit scenario (90% phosphate uptake compared to host-optimal) maximum phage

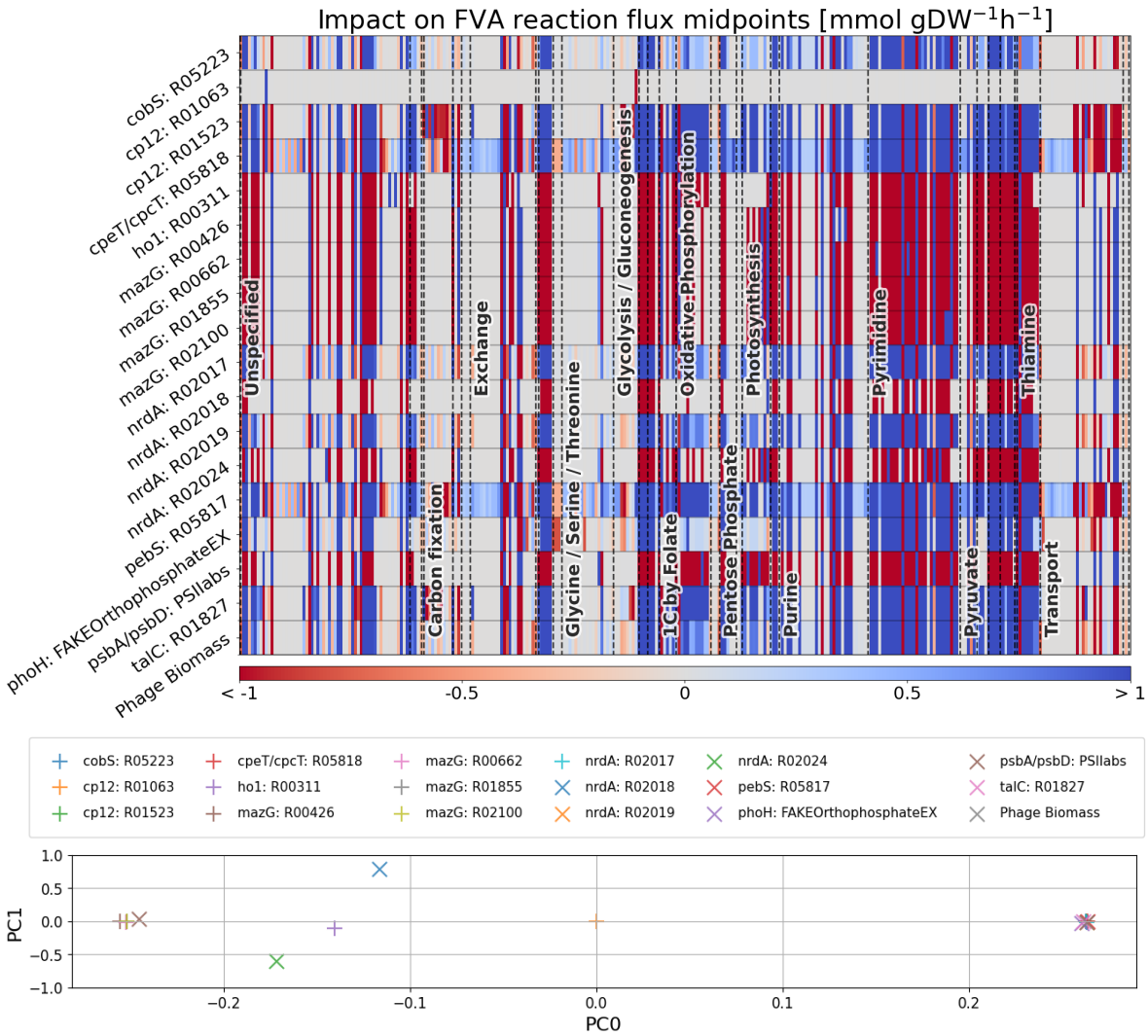

Figure 2: Indirect AMG impact on reactions. (Top) the AMG-hijacked reaction in each row is set to the midpoint of its target range (see Figure 1); in the last row, the phage biomass is instead constrained to its maximum value. Except in this last row, host biomass is kept fixed for a more direct comparison to healthy cells. The impact on each system reaction (columns) is measured from the midpoint of the FVA range and shown in color (red for decreased flux, blue for increased flux). Affected reactions (columns) are only shown if affected by more than 0.25mmol gDW^{-1} h^{-1} in at least one instance. They are arranged by metabolic subsystem, and subsystems with more than five affected reactions are labeled. (Bottom) PCA loadings of AMG impacts in the top panel (including impacts on reactions that are not shown because they fall below the 0.25mmol gDW^{-1} h^{-1} cutoff).

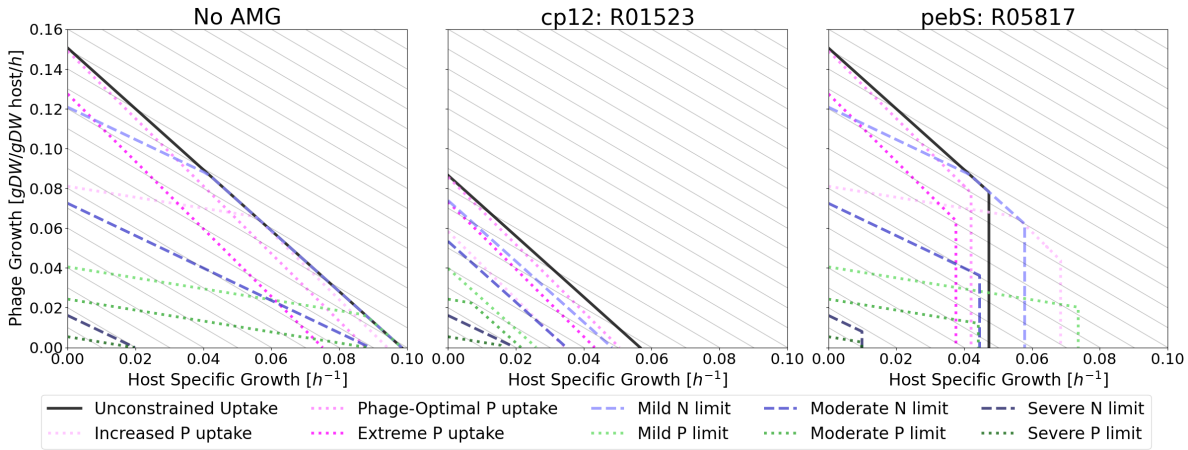

Figure 3: **Simultaneous maximization of phage production and host biomass flux.** Each panel depicts the instantaneous relationships between maximum (exponential) host specific growth rate and maximum (linear) phage growth rate for a either no AMG-induced hijacking (left), or for two AMG-hijacked reactions (center, right), as indicated in the figure titles. We show Pareto fronts for several nutrient-limiting scenarios: unconstrained uptake, and mild/moderate/severe limiting of N or P uptake. Mild, moderate, and severe limits correspond to 150%, 90%, and 20% of the values required for maximal host biomass production, respectively. When growth is not constrained by nutrient uptake, it is instead limited by RuBisCO efficiency. In addition, because the gene *phoH* is thought to promote increased P uptake (modeled here as hijacking the FAKEOrthophosphateEX exchange reaction), we also considered two scenarios in which the cell is forced to uptake more phosphate than is optimal (Increased P uptake: 3x host-optimal uptake; Phage-optimal P uptake: 6x host-optimal uptake; Extreme P uptake: 12x host-optimal uptake). We provide the remaining 14 AMG-impacted trade-off curves in Supplementary Figure S3.

growth is 84% lower, while maximum host growth is 10% lower. Additionally, in all nutrient stress scenarios, the slope of the Pareto front is less negative than in the unconstrained uptake scenario, meaning that an increase in host growth rate has a smaller impact on phage production under nutrient stress in the model.

Motivated by prior observations of increased P uptake during infection by cyanophage (*12*), we considered the effects of P uptake rates above the host-optimal level. Fixing P uptake at 300% the host-optimal level (increased P uptake scenario) results in slightly decreased optimal host growth rate (4%); but because this uptake rate is approximately half the phage-optimal value, maximum phage growth is reduced by 46%. At a phage-optimal P uptake rate, host growth is 11% lower. Beyond these levels (double the phage-optimal rate) maximum phage and host growth rates are 15% and 24% lower, respectively. Thus, host and phage have non-overlapping ranges of optimal P uptake rates, but both are have a wide range (spanning more than an order of magnitude in the case of the host) of suboptimal rates that permit above 50% of optimal growth.

To study how AMGs might affect the trade-off between host growth and phage production, we constrained each AMG-hijacked reaction to lie within a target range (recomputed as in Figure 1 under each nutrient scenario) and reconstructed the host-phage biomass trade-off curves. All phage-antialigned reactions (Figure 2) preserve the Pareto front. Despite the poor alignment between their indirect impacts on reaction FVA ranges and the requirements of phage biomass production, they do not prevent optimal phage production in any scenario.

All phage-aligned reactions substantially alter the phage-host biomass trade-off in the model. Reactions R05223 (*cobS*) R01523 (*cp12*), and R01827 (*talC*) shift the Pareto front inwards toward the origin, reducing both phage and host maximum biomass. Additionally, R01523 and R01827 make the Pareto front steeper in some nutrient-limiting scenarios, favoring phage growth (middle panel of Figure 3). The remaining phage-aligned reactions leave the phage-favored upper-left part of the Pareto front unchanged, while rendering host-favored lower-right portion inaccessible (right panel of Figure 3). In these cases of phage-aligned AMG-hijacked reactions, *the optimal host growth is no longer possible, whereas optimal phage growth remains possible*.

1.4 Published AMGs have both synergistic and antagonistic interactions

Because altering individual AMG-hijacked reactions can shift the midpoint of feasible ranges for other reactions in opposite directions, it is important to understand how these competing effects are reconciled within the host metabolism. To investigate this, we fixed biomass production at no less than half its maximal value and maximized the total flux impact on AMG-hijacked reactions while simultaneously minimizing the total absolute flux through other gene-regulated reactions by incorporating the standard parsimonious FBA objective as a constraint on enzyme usage. By introducing a linear *AMG reaction emphasis factor*, denoted γ, to weight these two flux objectives (see Methods), we simulated the trade-off between prioritizing AMG flux impacts and reducing overall flux. Higher values of γ favor AMG impacts, while lower values prioritize reducing total absolute flux (as in parsimonious FBA). We show the resulting change in reaction fluxes, compared to a baseline parsimonious FBA distribution, for the 20 most highly affected reactions in Figure 4, alongside the change in AMG-hijacked reaction fluxes that were obtained during this procedure.

We also considered the impacts on specific mass fluxes between metabolic subsystems (see Methods). We analyzed the changes in specific mass flux through each subsystem using mass flux difference (MFD) graphs for the most affected subsystems. These graphs highlight the change in mass flux through and between metabolic subsystems, and are defined in Methods. MFD graphs for four values of γ are depicted in Figure 5. In Supplementary Figure S2, we plotted the change in mass flux through the 20 most affected subsystems as a function of γ.

For small values of the γ, only the photosynthetic reaction PSIIabs (*psbA*/*psbB*) is directly affected (bottom panel of Figure 4). In our numerical experiments, increasing the flux through this reaction decreased mass flux through photosynthetic pathways and increased mass flux through the oxidative phosphorylation subsystem (top left panel of Figure 5). For $\gamma > 1$, suppression of R01063 by *cp12* (increased negative flux through this reversible reaction) becomes favorable (top right panel of Figure 5). Setting $\gamma = 2$ permits R01855 hijacking by *mazG* and a further increase in PSIIabs flux (bottom left panel of Figure 5). For this value of γ, the mass flux through both photosynthetic and oxidative phosphorylation pathways increased relative to the baseline, with more flux diverted to the biomass and carbon fixation subsystems and away from glycolysis and gluconeogenesis. Beyond $\gamma = 2$, R01855 was hijacked by *mazG*, leading to a large (1.95h^{-1}) increase in specific mass

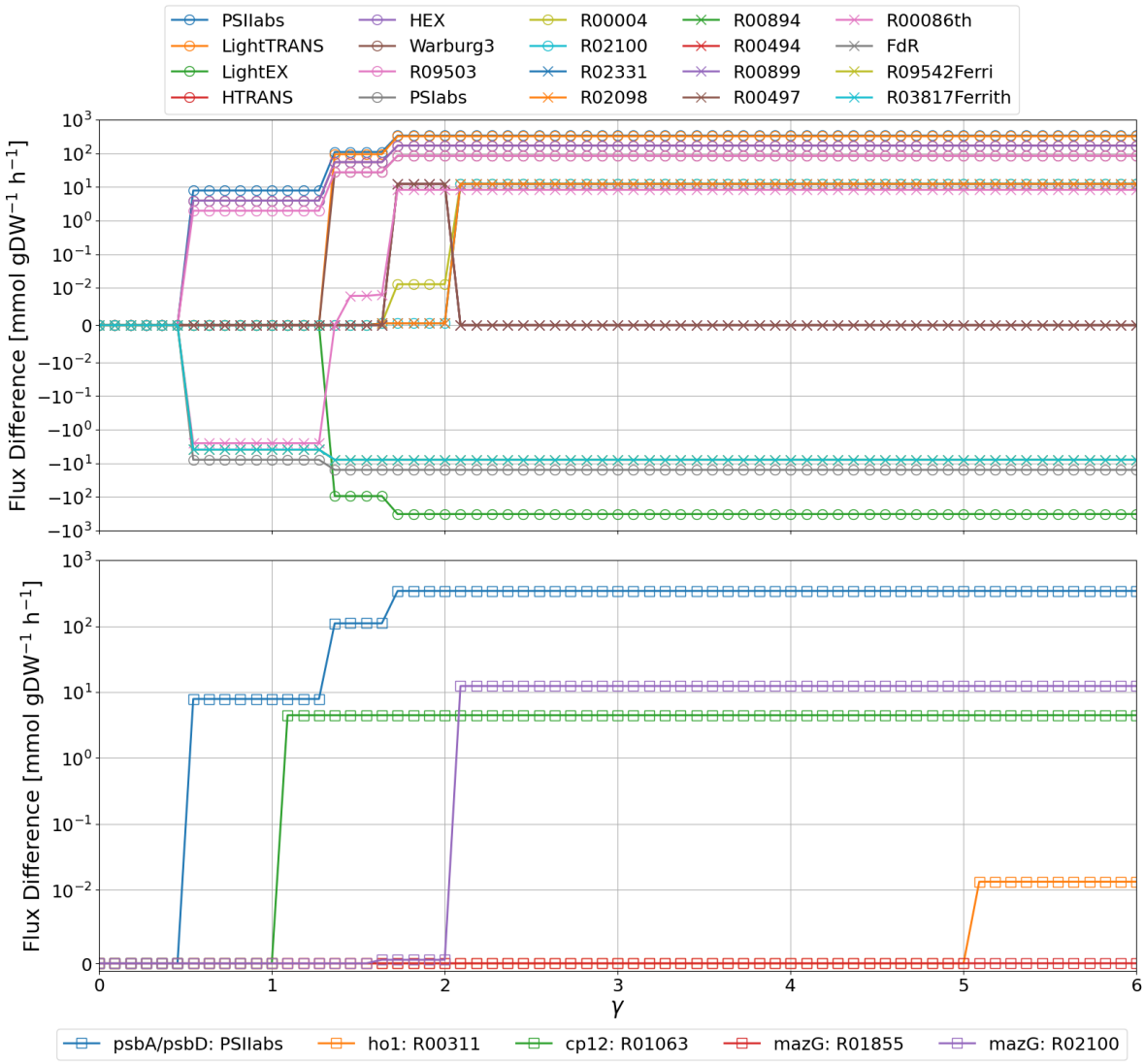

Figure 4: **Changes in reaction fluxes relative to a parsimonious FBA baseline.** Flux change is shown as a function of AMG reaction emphasis factor, γ. Increasing γ more heavily weights changing the flux of AMG-hijacked reactions relative to minimizing the flux through other gene-regulated reactions (see Methods). In the top panel, the 20 reactions with the largest responses are shown. The ten with the largest responses are marked with circles, the remaining ten use an X symbol. In the bottom panel, AMG-hijacked reactions with non-zero flux change are shown. The vertical axes use a symlog scale, so that the region between $\pm 10^{-2}$mmol gDW^{-1} h^{-1} is linearly scaled, while the regions outside this range use a logarithmic scale.

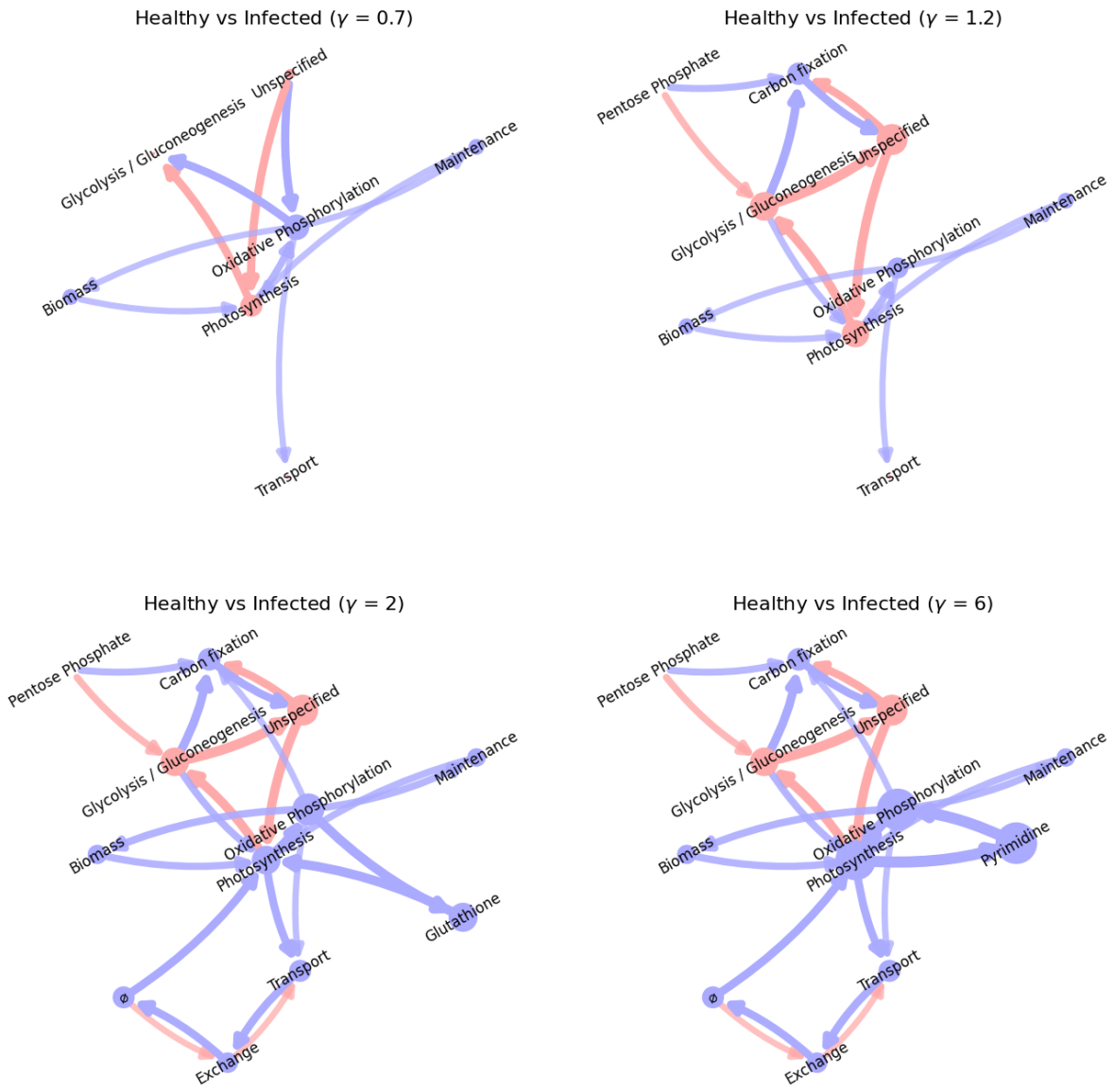

Figure 5: **Changes to subsystem-specific mass flux graphs.** Flux changes are represented as MFD graphs (see Methods), at various values of the AMG reaction emphasis factor, γ. Node and sizes and edge thickness and opacity are scaled according to absolute change in specific mass flux, with blue indicating an increase and red indicating a decrease in the infected system fluxes relative to the uninfected fluxes. Node positions are fixed across the four graphs. Edges with specific mass flux changes below 0.01h^{-1} are not shown, nor are nodes that lack incoming or outgoing edges above this threshold.

flux through purine metabolism (bottom right panel of Figure 5). For $\gamma \geq 5$, R00311 was hijacked by *ho1* in our simulations, but this had a negligible impact on subsystem mass flux (Supplementary Figure S2). All of the increased reactions belong to the phage-antialigned group identified in Figure 2. This may be due to the tendency of this group of reactions to decrease flux globally, which is favored when minimizing enzyme usage.

We further explored the synergistic and antagonistic effects of AMG-hijacked reactions by fixing pairs of AMG-hijacked reaction fluxes to their target ranges (Figure 1 and Supplementary Figure S4). To assess the interdependence of hijacked reaction pairs, we compared the edge-wise sum of MFD graphs computed from each hijacked reaction individually to the MFD graph obtained when hijacking both reactions simultaneously. We quantified this comparison using the edge-wise root mean square difference (RMSD) between these two graphs (Figure 6). In some cases, the two target ranges for the AMG-hijacked reactions are incompatible, resulting in infeasible solutions. Even reactions hijacked by the same gene can have this property, such as R00426 and R00662, both hijacked by the gene *mazG*.

In some scenarios, we observed antagonistic effects. The inset of Figure 6 highlights how *pebS*, via R05817, moderates the impact of *mazG* via R00662 in the model. At the system level, R05817 has only a small impact on mass flux on its own: only three subsystems have incoming or outgoing flux changes larger than 0.01h^{-1}. In contrast, R00662 impacts 20 subsystems above this threshold, with the largest impacts on the mass flux through the Pyrimidine, Photosynthesis, and Oxidative Phosphorylation subsystems. Flux is increased in all three subsystems when R00662 is hijacked. The largest inhibitory impact of R00662 hijacking in the model is on Glycolysis / Gluconeogenesis. Because the impact of R00662 is much larger than the impact of R05817, their combined effect would be very similar to that of R00662 alone if they acted independently–however, we instead observed that the inhibitory impacts of R00662 on Glycolysis / Gluconeogenesis and several of its neighboring subsystems are substantially decreased by R05817 hijacking (see inset of Figure 6). In contrast, the promotive impacts in the combined R00662-R05817 hijacking scenario on Pyrimidine, Photosynthesis, and Oxidative Phosphorylation subsystem mass fluxes remain high. We also noted a $> 0.01\text{h}^{-1}$ increase in mass flow from the Oxidative Phosphorylation subsystem to Biomass reactions in the combined R00662-R05817 hijacking scenario that does not appear when hijacking either reaction alone.

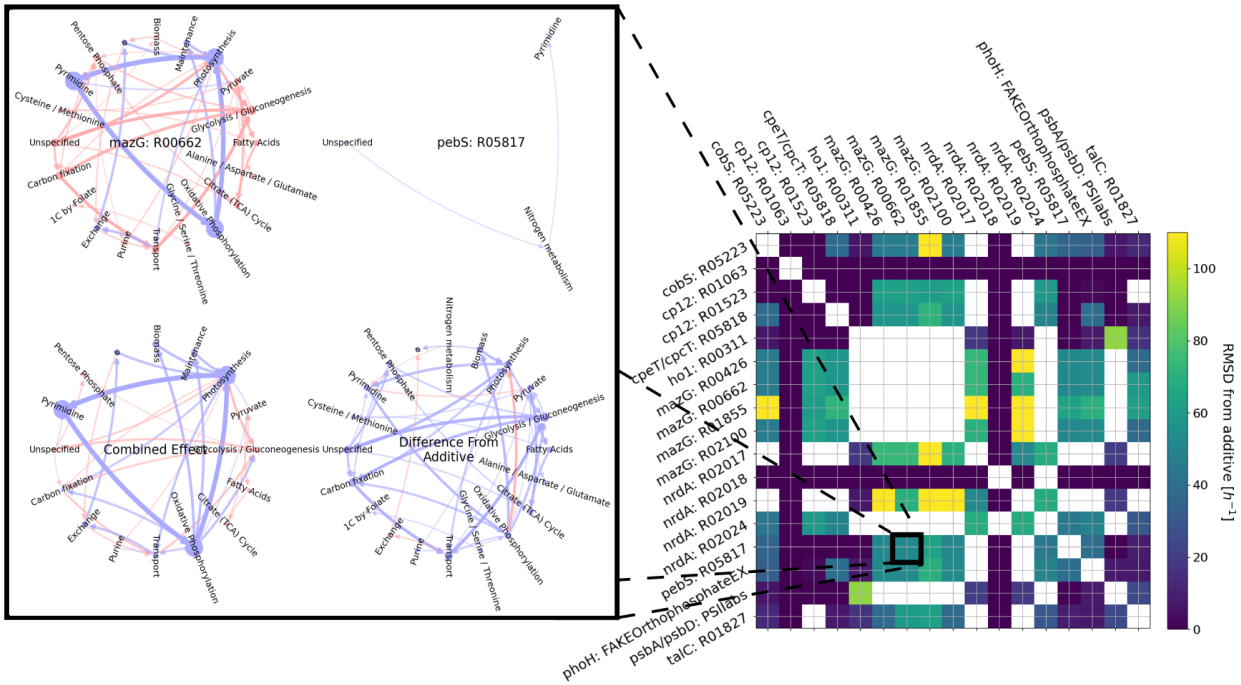

Figure 6: **Interdependence of subsystem mass flux response to AMG-hijacked reactions.** Each entry of the heat map matrix corresponds to a pair of AMG-hijacked reactions, indicated by the row and column labels. For each pair, two combined MFD graphs were computed: the true MFD graph obtained from constraining both hijacked reactions to their target ranges (see Figure 1), and the additive MFD graph obtained by adding (edge-wise) the MFD graphs obtained by individually hijacking each reaction. The non-additivity is the edge-wise RMSD of the additive and true MFD graphs. Off-diagonal white cells indicate that the combined hijacking of the two AMGs did not yield a feasible solution. An example with moderate RMSD comparing the MFD graph for individual and combined effects of hijacking two reactions is shown in the inset. In all graph plots, edges with specific mass flux changes below 0.01h^{-1} are not shown, nor are nodes that lack incoming or outgoing edges above this threshold.

The most extreme deviations from purely additive effects occur between reactions hijacked by *nrdA* and *mazG*, both of which modulate nucleotide metabolism (Figure 6). Apart from these, the largest deviation from a purely additive effect occurs between PSIIabs and R00311, visualized in Supplementary Figure S5, alongside a summary of the subsystem fluxes that most substantially deviate from an additive response. In the example of PSIIabs and R00311, the largest deviations from a purely additive effect involve Photosynthesis, Oxidative Phosphorylation, Pyrimidine, and Porphyrin / Chlorophyll metabolism. In particular, in the combined hijacking scenario, mass flux through the Photosynthesis subsystem is synergistically altered so that it favors diverting mass flux from Oxidative Phosphorylation to Porphyrin / Chlorophyll and then to Pyrimidine metabolism.

2 Discussion

Interactions between bacteriophages and their hosts are complex: they involve ecosystem-scale exchange of metabolic genes over evolutionary time scales, coupled to minute-by-minute changes in intracellular reaction fluxes. A critical step in unraveling this complexity is to characterize how exchanged AMGs hijack metabolic reactions to shape host metabolism and promote viral propagation. Genome-scale metabolic models can accelerate these biotechnology applications by predicting optimal combinations of phage-derived AMGs for redirecting metabolic fluxes toward desired products while maintaining cell viability. Such systems-level understanding of metabolic hijacking mechanisms can inform the rational design of synthetic metabolic circuits that mimic successful viral strategies for metabolic reprogramming.

This is a challenging task in part because dynamical parameters are poorly constrained. Fortunately, static methods, such as flux analysis, can uncover important parameter-robust insights. Our approach represents a framework for studying possible AMG impact scenarios using flux balance analysis and other closely related techniques. In analyzing the *P. marinus* infection by the cyanophage P-HM2, we uncover the systems-level impacts brought about by viral hijacking of individual metabolic reactions.

Our model predicts that, under the same carbon fixation restrictions, optimal phage biomass production requires substantially more nitrogen and phosphorous uptake than optimal host growth. Indeed, it has been previously reported (*12*) that infected host cells exhibit increased nutrient uptake.

This does not necessarily imply that typical phage replication rates require more of these nutrients. This may reflect the phage's insensitivity to other limiting factors (e.g., the energetic demands of lipid formation), allowing phage to better leverage additional available nutrients. Despite the six-fold difference in predicted optimal phosphate uptake, we predict that both host and phage can tolerate wide and overlapping ranges of sub-optimal phosphate uptake rates while maintaining 50% of their optimal biomass accumulation rates (Figure 1 and Figure 3).

Our analyses offer several new predictions–such as that AMG-hijacked reactions fall into phage-aligned and phage-antialigned categories. We predict that phage-aligned hijacking more readily alters the phage-host biomass trade-off in favor of the phage, while phage-antialigned hijacking has no or minimal impact on this trade-off. We also predict that phage-antialigned hijacking more easily shifts host metabolism, as it permits lower reaction fluxes globally, meaning less (metabolism-wide) enzyme production is required for their function. We discover that AMGs can have nontrivial synergistic and antagonistic impacts on host metabolism, and our analysis of AMG-emphasis scenarios shows that subsystem flux can respond non-monotonically to alterations in the flux of reactions hijacked by AMGs (such as in the case of mass flux through the photosynthesis subsystem as flux through PSIIabs increases).

Though our analysis describes the global effects of combinatorial reaction hijacking, not all AMG-hijacked reactions changed flux in our numerical experiments–in fact, most did not. In some cases, this is because increasing flux in some of these reactions requires a large increase in the total flux along the pathway of the AMG-hijacked reaction, thus requiring a substantial increase in enzyme production. Even disregarding enzymatic constraints, it is not stoichiometrically feasible to simultaneously maximize the impact of all observed AMGs. Thus we have considered scenarios in which the individual impacts of AMGs are not maximized. It is possible, for instance, that the produced AMG enzyme is not efficiently used by the host metabolism, and previous work has demonstrated time-dependent transcriptional regulation of AMGs (*27*). Additionally, a host cell may have branch points in its metabolism that are seldom used in tandem. If a phage benefits from hijacking one branch, the cell may divert metabolic flux toward the alternate branch to compensate. In such cases, the phage may benefit from hijacking reactions along both branches, even when only one branch or the other can carry flux. In some cases, however, AMGs have antagonistic, competing impacts on metabolic fluxes. We conjecture that these antagonistic interactions, when coupled to

time-varying expression of AMGs in the host organism (such as differentiated by the day/night cycle (*27, 49*)), may play a role in synchronizing lysis for optimal burst size. This latter explanation is supported by the extent to which hijacking-induced changes to host-optimal flux ranges either align or anti-align with those induced by optimizing phage biomass.

Very few of our system-level predictions can be obtained straightforwardly using purely local methods. Instead, we must consider the global metabolic response to AMG perturbations to uncover phage-aligned or phage-antialigned flux changes across multiple metabolic subsystems. Moreover, the cascading effects that give rise to a global response are not necessarily mediated by traditionally recognized core pathways. When plotting metabolic networks, reactions that share only common metabolites, such as ATP, are not connected to avoid excessive visual complexity–but the metabolism is inherently complex. Frequently disregarded connections can become important in characterizing the global response to a phage hijacking nucleotide metabolism, for example, encouraging ATP (and dATP) production at the cost of lipid metabolism or other normally essential processes. Similarly, competition between subsystems for common metabolites may play an important role in the alteration of host metabolism; predicting the winners and losers of these competitions requires engaging with the metabolic network as a whole. The complexity only increases when we must untangle the collision of perturbation cascades induced by AMGs hijacking reactions from widely separated portions of the metabolic network. This work has begun to address the complexity of the quasi-stationary response to AMG hijacking, but yet another layer of complexity remains to be incorporated in future work: the temporal heterogeneity of AMG expression (as described by (*27*)). The characterization of the metabolic response we have presented here relies on a computationally efficient and parameter-robust method, laying the groundwork for more resource-intensive in-depth investigations of AMG hijacking response during bacteriophage infection.

3 Methods

3.1 Building P-HM2 phage biomass and AMGs into a genome scale model

In this study, we used as a starting point the genome-scale iJC568v7 model of *Prochloroccocus marinus* MED4 metabolism, which was released by (*37*) (retrieved from `https://github.com/`

`segrelab/Prochlorococcus_Model`) as a refinement of the model developed by (*38*), and which has received several small updates since. In addition to various updates by the original creators of iJC568v7 , we have made a small correction to the RNA synthesis reaction so that dNMP (rather than dNTP) polymers are constructed from dNTP, releasing diphosphate in the process. Though in principle this could affect global phosphate balance in the model, we did not observe any significant change in the system-level behavior as a result of this adjustment. Reactions are assigned one of 55 human-annotated subsystems (including one Unspecified subsystem). Metabolites are assigned one of five compartments (extracellular, cytoplasm, carboxysome, thylakoid, or periplasm). This model contains 994 metabolic reactions among 802 metabolites (including duplicates from different compartments). Reaction activity is governed by 595 genes.

To add the phage biomass function to the iJC568v7 model, we identified the stoichiometric requirements for phage growth due to structural proteins and genetic material. We created a dictionary of structural proteins in the P-HM2 phage using previously published protein counts from the homologous and canonical Myoviridae member T4 bacteriophage (*50*). We used the structural proteins in the T4 bacteriophage particle in a BLAST search against the UniProt P-HM2 proteome. We identified each protein with the highest score in the BLAST results as the homologous protein in the P-HM2 virion particle, and used the corresponding stoichiometry of T4 structural proteins as the stoichiometry of their homologous structural proteins in P-HM2. All the stoichiometry matching between T4 and P-HM2 and BLAST results can be found in Supplementary Data S1.

Using the stoichiometry dictionary of structural phage proteins as input, we adapted the procedure of (*36*)–designed for computing biomass of SARS CoV2, a single-stranded RNA virus–to the generate the biomass function of P-HM2, a double-stranded DNA virus. First, we extracted P-HM2 coding sequences from its genome and used the stoichiometry of structural proteins to calculate the number of each type of amino acid and nucleotide required to form a single virion particle. For mRNA requirements, we assume an average of 540 protein copies per mRNA transcript, in line with previous measurements in bacteria (*51*); in practice, so long as it remains large, the precise value of this ratio has a negligible impact on maximal phage production rate. We estimated ATP consumption for amino acid and nucleotide polymerization in each phage using the method of (*36*). From the total molar weight of each component consumed, we computed the molar weight of one P-HM2 phage particle to normalize the biomass expression so that the stoichiometric coefficients

describe the number, in mmol, of each metabolite required to produce one gram of virus. We measure production rates of these metabolites in mmol gDW^{-1} h^{-1}. Thus, our model outputs the virus biomass production rate in grams of virus produced per hour per gram dry weight of host. The detailed composition of the biomass function is given in Supplementary Table S2.

This model assumes a full and proper packing of phage heads, and that the constituent phage structure proteins are only created in the correct stoichiometric ratios. P-HM2 is known to mis-package host DNA inside its capsid, but the frequency of this occurrence is low and unlikely to significantly impact the model (*52*).

To integrate AMGs into the model, we used previously identified AMGs from (*40*) to conduct a BLAST search against the P-HM2 genome. We then identified the reactions in iJC568v7 that are indicated in the KEGG database (*41*) as most directly impacted by each AMG. To determine the direction (increase or decrease) of the AMG's impact on the associated reaction, we conducted a literature search of each AMG (the results are summarized in Table 1). Only one AMG we identified, *cp12*, caused an inhibition of reactions. Its associated protein, CP12, inhibits two reactions that govern the Calvin-Benson cycle (R01063 and R01523) by sequestering the enzymes that catalyze these reactions, as discussed in (*25*). Additional details about individual AMGs are provided in Supplementary Text.

3.2 Assessing AMG impacts using FBA

Here, following a parsimonious approach, we maximized the collective flux change in AMG reactions (taking care to enforce the correct direction of change), while simultaneously minimizing the flux through non-AMG reactions, subject to the constraint that the biomass flux remain near its optimal value. We considered a linear trade-off between the two objectives, parametrized by an *AMG reaction emphasis factor*, γ. Higher values of γ favor maximizing AMG-promoted reaction flux changes, and lower values favoring minimization of non-AMG flux. We sorted each AMG-regulated reaction r depending on the direct, unmediated impact of the AMG perturbations on its flux before (v_r) and after (v'_r) perturbation, as shown in Table 2. We also considered the set $\mathcal{R}_g$ of all gene-regulated reactions (including those regulated by AMGs). After determining the optimal host flux distributions using the primary objective (Equations S1-S3), we then minimized the secondary

Category	Reaction Set	Criterion
Increase	$\mathcal{R}_{ri}$	v'_r should be larger than v_r
Absolute Increase	$\mathcal{R}_{ai}$	$\lvert v'_r \rvert$ should be larger than $\lvert v_r \rvert$
Decrease	$\mathcal{R}_{rd}$	v'_r should be smaller than v_r
Absolute Decrease	$\mathcal{R}_{ad}$	$\lvert v'_r \rvert$ should be smaller than $\lvert v_r \rvert$

Table 2: **AMG-regulated reaction sets.** Each reaction set is denoted $\mathcal{R}_{xx}$, where xx denotes the respective category reactions (ri,ai,rd,ad), can be categorized based on the change in flux through the reactions after infection perturbation.

objective $F_{sec}(v_r)$,

$$F_{sec}(v_r) = \sum_{r \in \mathcal{R}_g} |v'_r| + \gamma \left(\sum_{r \in \mathcal{R}_{rd}} v'_r + \sum_{r \in \mathcal{R}_{ad}} |v'_r| - \sum_{r \in \mathcal{R}_{ri}} v'_r - \sum_{r \in \mathcal{R}_{ai}} |v'_r| \right), \tag{1}$$

subject to the constraints i) that nutrient uptake does not increase too much (we use a factor of 2 as an upper limit, but also test with a factor of 5), and ii) the host biomass flux does not decrease below a specified threshold fraction (we use 50%) from the primary objective. We also retained any additional constraints used in determining the unperturbed flux distributions.

To evaluate how phage production limits host growth in various nutrient scenarios, we swept the full range of possible host growth rates for given constraints on phosphate and ammonia uptake. We considered unconstrained uptake, and uptake-limited scenarios where no more than 150%, 90%, or 20% of the uptake rates that yield optimal host growth can be achieved. Motivated by the reported impact of the AMG *phoH* (*12*), we also considered two scenarios in which phosphate uptake is 3x or 30x greater than optimal for host growth. In each scenario, we considered 50 fixed values of host growth, evenly spaced between optimal growth in the scenario and near zero (0.0001%) growth. At each fixed value of host growth, we computed the maximum attainable phage replication rate, thereby constructing a Pareto front for the trade-off between host and phage growth. We repeated this process, restricting each AMG-hijacked reaction (except for phosphate exchange) to its target range to construct AMG-modulated trade-off curves.

3.3 Subsystem mass flux graphs

We constructed subsystem mass flux graphs to elucidate the impacts of AMGs at a coarse-grained subsystem level. In these graphs, each node represents an annotated metabolic subsystem (for example, biomass formation, photosynthesis, or oxidative phosphorylation). Reactions in the iJC568v7 model were annotated with a subsystem, though a plurality was not assigned a specific subsystem (here, we assigned these the *Unspecified* subsystem label). For a given reaction flux distribution, we computed the mass flux through a subsystem by first finding the net production or consumption rate of each metabolite by reactions in that subsystem. We accounted for metabolites that enter or exit the system (e.g., via exchange reactions or biomass secretion) by introducing an additional reaction subsystem, denoted $\emptyset$, corresponding to the environment which closes the system. Mass conservation requires that the mass-weighted sum of the positive metabolic production rates be equal in magnitude to the mass-weighted sum of the negative metabolic production rates (i.e., the consumption rates) in each reaction subsystem. We took this value to be the total flux through the subsystem. Mass flux between subsystems was calculated for each metabolite by determining the difference in mass production rates between two systems with directed edge weights being assigned the positive direction for mass flux.

To compare two subsystem mass flux graphs (e.g., for a healthy vs an infected cell), we constructed mass flux difference (MFD) and mass flux sum (MFS) graphs. These graphs are constructed with the difference or sum of weighted adjacency matrices for two subsystem mass flux graphs. Specifically, for two interventions, we computed G_{add} as the MFS graph of the two individual MFD graphs relative to an intervention-free baseline, and then computed the MFD graph G_{true} of the simultaneous interventions, relative to the same baseline. The deviation from independent impacts is captured by the MFD of G_{true} relative to G_{add}. As a simple summary of this graph, we computed its root mean square edge-weight.

References and Notes

1. J. R. Waldbauer, *et al.*, Nitrogen Sourcing during Viral Infection of Marine Cyanobacteria. *Proceedings of the National Academy of Sciences* **116** (31), 15590–15595 (2019), doi:10.1073/pnas.1901856116.

2. V. Berzin, G. Rosenthal, E. J. Gren, Cellular Macromolecule Synthesis in *Escherichia Coli* Infected with Bacteriophage MS2. *European Journal of Biochemistry* **45** (1), 233–242 (1974), doi:10.1111/j.1432-1033.1974.tb03547.x.

3. G. Mahmoudabadi, R. Milo, R. Phillips, Energetic Cost of Building a Virus. *Proceedings of the National Academy of Sciences* **114** (22) (2017), doi:10.1073/pnas.1701670114.

4. Y. Yu, A. J. Clippinger, J. C. Alwine, Viral Effects on Metabolism: Changes in Glucose and Glutamine Utilization during Human Cytomegalovirus Infection. *Trends in Microbiology* **19** (7), 360–367 (2011), doi:10.1016/j.tim.2011.04.002.

5. N. D. Maynard, M. V. Gutschow, E. W. Birch, M. W. Covert, The Virus as Metabolic Engineer. *Biotechnology Journal* **5** (7), 686–694 (2010), doi:10.1002/biot.201000080.

6. R. J. Puxty, A. D. Millard, D. J. Evans, D. J. Scanlan, Viruses Inhibit CO 2 Fixation in the Most Abundant Phototrophs on Earth. *Current Biology* **26** (12), 1585–1589 (2016), doi:10.1016/j.cub.2016.04.036.

7. T. El-Bacha, M. M. Menezes, M. C. Azevedo E Silva, M. Sola-Penna, A. T. Da Poian, Mayaro Virus Infection Alters Glucose Metabolism in Cultured Cells through Activation of the Enzyme 6-Phosphofructo 1-Kinase. *Molecular and Cellular Biochemistry* **266** (1/2), 191–198 (2004), doi:10.1023/B:MCBI.0000049154.17866.00.

8. J. Munger, S. U. Bajad, H. A. Coller, T. Shenk, J. D. Rabinowitz, Dynamics of the Cellular Metabolome during Human Cytomegalovirus Infection. *PLoS Pathogens* **2** (12), e132 (2006), doi:10.1371/journal.ppat.0020132.

9. E. T. Sieradzki, J. C. Ignacio-Espinoza, D. M. Needham, E. B. Fichot, J. A. Fuhrman, Dynamic Marine Viral Infections and Major Contribution to Photosynthetic Processes Shown by Spa-

tiotemporal Picoplankton Metatranscriptomes. *Nature Communications* **10** (1), 1169 (2019), doi:10.1038/s41467-019-09106-z.

10. M. B. N. Albright, *et al.*, Experimental Evidence for the Impact of Soil Viruses on Carbon Cycling during Surface Plant Litter Decomposition. *ISME Communications* **2** (1), 24 (2022), doi:10.1038/s43705-022-00109-4.

11. L. Bi, *et al.*, Diversity and Potential Biogeochemical Impacts of Viruses in Bulk and Rhizosphere Soils. *Environmental Microbiology* **23** (2), 588–599 (2021), doi:10.1111/1462-2920.15010.

12. D. B. Goldsmith, *et al.*, Development of *phoH* as a Novel Signature Gene for Assessing Marine Phage Diversity. *Applied and Environmental Microbiology* **77** (21), 7730–7739 (2011), doi:10.1128/AEM.05531-11.

13. M. A. Fortuna, *et al.*, Coevolutionary Dynamics Shape the Structure of Bacteria-phage Infection Networks. *Evolution* **73** (5), 1001–1011 (2019), doi:10.1111/evo.13731.

14. K. D. Seed, D. W. Lazinski, S. B. Calderwood, A. Camilli, A Bacteriophage Encodes Its Own CRISPR/Cas Adaptive Response to Evade Host Innate Immunity. *Nature* **494** (7438), 489–491 (2013), doi:10.1038/nature11927.

15. X.-Q. Luo, *et al.*, Viral Community-Wide Auxiliary Metabolic Genes Differ by Lifestyles, Habitats, and Hosts. *Microbiome* **10** (1), 190 (2022), doi:10.1186/s40168-022-01384-y.

16. M. J. Szul, S. P. Dearth, S. R. Campagna, E. R. Zinser, Carbon Fate and Flux in *Prochlorococcus* under Nitrogen Limitation. *mSystems* **4** (1), e00254–18 (2019), doi:10.1128/mSystems.00254-18.

17. P. Bhatt, B. A. Engel, M. Reuhs, H. Simsek, Cyanophage technology in removal of cyanobacteria mediated harmful algal blooms: A novel and eco-friendly method. *Chemosphere* **315**, 137769 (2023), doi:10.1016/j.chemosphere.2023.137769, `https://www.sciencedirect.com/science/article/pii/S0045653523000358`.

18. S. A. Gómez-Ochoa, *et al.*, Efficacy of phage therapy in preclinical models of bacterial infection: a systematic review and meta-analysis. *The Lancet Microbe* **3** (12), e956–e968 (2022), publisher: Elsevier, doi:10.1016/S2666-5247(22)00288-9, https://www.thelancet.com/journals/lanmic/article/PIIS2666-5247(22)00288-9/fulltext.

19. N. Sawa, *et al.*, Modification of carbon metabolism in *Synechococcus elongatus* PCC 7942 by cyanophage-derived sigma factors for bioproduction improvement. *Journal of Bioscience and Bioengineering* **127** (2), 256–264 (2019), doi:10.1016/j.jbiosc.2018.07.019, https://www.sciencedirect.com/science/article/pii/S1389172318302482.

20. B. L. Hurwitz, J. R. Brum, M. B. Sullivan, Depth-stratified functional and taxonomic niche specialization in the 'core' and 'flexible' Pacific Ocean Virome. *The ISME Journal* **9** (2), 472–484 (2015), doi:10.1038/ismej.2014.143, https://doi.org/10.1038/ismej.2014.143.

21. J. R. Brum, M. B. Sullivan, Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nature Reviews Microbiology* **13** (3), 147–159 (2015), publisher: Nature Publishing Group, doi:10.1038/nrmicro3404, https://www.nature.com/articles/nrmicro3404.

22. B. L. Hurwitz, J. M. U'Ren, Viral Metabolic Reprogramming in Marine Ecosystems. *Current Opinion in Microbiology* **31**, 161–168 (2016), doi:10.1016/j.mib.2016.04.002.

23. D. Lindell, J. D. Jaffe, Z. I. Johnson, G. M. Church, S. W. Chisholm, Photosynthesis genes in marine viruses yield proteins during host infection. *Nature* **438** (7064), 86–89 (2005), publisher: Nature Publishing Group, doi:10.1038/nature04111, https://www.nature.com/articles/nature04111.

24. J. Wiethaus, A. W. U. Busch, T. Dammeyer, N. Frankenberg-Dinkel, Phycobiliproteins in *Prochlorococcus marinus*: Biosynthesis of pigments and their assembly into proteins. *European Journal of Cell Biology* **89** (12), 1005–1010 (2010), doi:10.1016/j.ejcb.2010.06.017, https://www.sciencedirect.com/science/article/pii/S0171933510001317.

25. L. R. Thompson, *et al.*, Phage Auxiliary Metabolic Genes and the Redirection of Cyanobacterial Host Carbon Metabolism. *Proceedings of the National Academy of Sciences* **108** (39) (2011), doi:10.1073/pnas.1102164108.

26. L. Kelly, H. Ding, K. H. Huang, M. S. Osburne, S. W. Chisholm, Genetic diversity in cultured and wild marine cyanomyoviruses reveals phosphorus stress as a strong selective agent. *The ISME Journal* **7** (9), 1827–1841 (2013), doi:10.1038/ismej.2013.58, `https://doi.org/10.1038/ismej.2013.58`.

27. L. R. Thompson, Q. Zeng, S. W. Chisholm, Gene Expression Patterns during Light and Dark Infection of Prochlorococcus by Cyanophage. *PloS One* **11** (10), e0165375 (2016), doi:10.1371/journal.pone.0165375.

28. B. Rihtman, *et al.*, Cyanophage MazG is a pyrophosphohydrolase but unable to hydrolyse magic spot nucleotides. *Environmental microbiology reports* **11** (3), 448–455 (2019).

29. D. Segrè, D. Vitkup, G. M. Church, Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences* **99** (23), 15112–15117 (2002), publisher: Proceedings of the National Academy of Sciences, doi:10.1073/pnas.232349399, `https://www.pnas.org/doi/abs/10.1073/pnas.232349399`.

30. K. Raman, N. Chandra, Flux balance analysis of biological systems: applications and challenges. *Briefings in Bioinformatics* **10** (4), 435–449 (2009), doi:10.1093/bib/bbp011, `https://doi.org/10.1093/bib/bbp011`.

31. M. Breuer, *et al.*, Essential metabolism for a minimal cell. *eLife* **8**, e36842 (2019), publisher: eLife Sciences Publications, Ltd, doi:10.7554/eLife.36842, `https://doi.org/10.7554/eLife.36842`.

32. T. Shlomi, O. Berkman, E. Ruppin, Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proceedings of the National Academy of Sciences* **102** (21), 7695–7700 (2005), publisher: Proceedings of the National Academy of Sciences, doi:10.1073/pnas.0406346102, `https://www.pnas.org/doi/abs/10.1073/pnas.0406346102`.

33. B. J. Sánchez, *et al.*, Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Molecular Systems Biology* **13** (8), 935 (2017), publisher: John Wiley & Sons, Ltd, doi:10.15252/msb.20167411, https://www.embopress.org/doi/full/10.15252/msb.20167411.

34. F. Li, *et al.*, Deep learning-based kcat prediction enables improved enzyme-constrained model reconstruction. *Nature Catalysis* **5** (8), 662–672 (2022), publisher: Nature Publishing Group, doi:10.1038/s41929-022-00798-z, https://www.nature.com/articles/s41929-022-00798-z.

35. Z. A. King, *et al.*, BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Research* **44** (D1), D515–D522 (2016), doi:10.1093/nar/gkv1049, https://doi.org/10.1093/nar/gkv1049.

36. H. Delattre, K. Sasidharan, O. S. Soyer, Inhibiting the Reproduction of SARS-CoV-2 through Perturbations in Human Lung Cell Metabolic Network. *Life Science Alliance* **4** (1), e202000869 (2021), doi:10.26508/lsa.202000869.

37. S. Ofaim, S. Sulheim, E. Almaas, D. Sher, D. Segrè, Dynamic Allocation of Carbon Storage and Nutrient-Dependent Exudation in a Revised Genome-Scale Model of Prochlorococcus. *Frontiers in Genetics* **12**, 586293 (2021), doi:10.3389/fgene.2021.586293.

38. J. R. Casey, A. Mardinoglu, J. Nielsen, D. M. Karl, Adaptive Evolution of Phosphorus Metabolism in *Prochlorococcus*. *mSystems* **1** (6), e00065–16 (2016), doi:10.1128/mSystems.00065-16.

39. M. B. Sullivan, *et al.*, Genomic Analysis of Oceanic Cyanobacterial Myoviruses Compared with T4-like Myoviruses from Diverse Hosts and Environments. *Environmental Microbiology* **12** (11), 3035–3056 (2010), doi:10.1111/j.1462-2920.2010.02280.x.

40. L. T. Crummett, R. J. Puxty, C. Weihe, M. F. Marston, J. B. Martiny, The Genomic Content and Context of Auxiliary Metabolic Genes in Marine Cyanomyoviruses. *Virology* **499**, 219–229 (2016), doi:10.1016/j.virol.2016.09.016.

41. M. Kanehisa, Y. Sato, M. Kawashima, M. Furumichi, M. Tanabe, KEGG as a reference resource for gene and protein annotation. *Nucleic acids research* **44** (D1), D457–D462 (2016).

42. A. Ebrahim, J. A. Lerman, B. O. Palsson, D. R. Hyduke, COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Systems Biology* **7** (1), 1–6 (2013), number: 1 Publisher: BioMed Central, doi:10.1186/1752-0509-7-74, `https://bmcsystbiol.biomedcentral.com/articles/10.1186/1752-0509-7-74`.

43. A. Srivastava, N. Gupta, A. K. Mishra, Cyanophages: interacting mechanism and evolutionary significance, in *Cyanobacteria* (Elsevier), pp. 255–282 (2024).

44. R. Gasper, *et al.*, Distinct features of cyanophage-encoded T-type phycobiliprotein lyase ΦCpeT: The role of auxiliary metabolic genes. *Journal of Biological Chemistry* **292** (8), 3089–3098 (2017).

45. M. M. Konert, A. Wysocka, P. Koník, R. Sobotka, High-light-inducible proteins HliA and HliB: pigment binding and protein–protein interactions. *Photosynthesis Research* **152** (3), 317–332 (2022).

46. H. Maaroufi, R. M. Tanguay, Analysis and Phylogeny of Small Heat Shock Proteins from Marine Viruses and Their Cyanobacteria Host. *PLOS ONE* **8** (11), e81207 (2013), publisher: Public Library of Science, doi:10.1371/journal.pone.0081207, `https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0081207`.

47. Z. Ran, *et al.*, Identification of ac-type heme oxygenase and its function during acclimation of cyanobacteria to nitrogen fluctuations. *Communications Biology* **6** (1), 944 (2023).

48. V. M. Johnson, H. B. Pakrasi, Advances in the Understanding of the Lifecycle of Photosystem II. *Microorganisms* **10** (5), 836 (2022).

49. R. Liu, Y. Liu, Y. Chen, Y. Zhan, Q. Zeng, Cyanobacterial viruses exhibit diurnal rhythms during infection. *Proceedings of the National Academy of Sciences* **116** (28), 14077–14082 (2019), publisher: Proceedings of the National Academy of Sciences, doi:10.1073/pnas.1819689116, `https://www.pnas.org/doi/abs/10.1073/pnas.1819689116`.

50. M. L. Yap, M. G. Rossmann, Structure and function of bacteriophage T4. *Future microbiology* **9** (12), 1319–1327 (2014).

51. P. Lu, C. Vogel, R. Wang, X. Yao, E. M. Marcotte, Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nature Biotechnology* **25** (1), 117–124 (2007), doi:10.1038/nbt1270.

52. R. Laurenceau, N. Raho, M. Forget, A. A. Arellano, S. W. Chisholm, Frequency of mispackaging of Prochlorococcus DNA by cyanophage. *The ISME Journal* **15** (1), 129–140 (2021), publisher: Nature Publishing Group, doi:10.1038/s41396-020-00766-0, `https://www.nature.com/articles/s41396-020-00766-0`.

53. J. Zhang, M. Inouye, MazG, a Nucleoside Triphosphate Pyrophosphohydrolase, Interacts with Era, an Essential GTPase in Escherichia coli. *Journal of Bacteriology* **184** (19), 5323–5329 (2002), publisher: American Society for Microbiology, doi:10.1128/jb.184.19.5323-5329.2002, `https://journals.asm.org/doi/10.1128/jb.184.19.5323-5329.2002`.

54. F. S. Foundation, GLPK - GNU Project - Free Software Foundation (FSF), `https://www.gnu.org/software/glpk/`.

Acknowledgments

We acknowledge the insightful discussions with Martin H. W. Gruebele, Angad Pankaj Mehta, Zaida Ann Luthey-Schulten, and Aaron Chan from the University of Illinois Urbana-Champaign, with John Casey from Lawrence Livermore National Laboratory, as well as with William Nelson, Jason McDermott, and Chris Oehmen from Pacific Northwest National Laboratory.

Funding: This work is supported by the NW-BRaVE for Biopreparedness project funded by the U. S. Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research, under FWP 81832. A portion of this research was performed on a project award (https://doi.org/10.46936/staf.proj.2023.61054/60012367) from the Environmental Molecular Sciences Laboratory, a DOE Office of Science User Facility sponsored by the Biological and Environmental Research program under Contract No. DE-AC05-76RL01830. Pacific Northwest National

Laboratory is a multi-program national laboratory operated by Battelle for the DOE under Contract DE-AC05-76RLO 1830.

Author contributions: **Conceptualization:** SF, RW; **Methodology:** JR, WS, SF; **Software:** SF, WS, JR; **Formal Analysis:** JR, WS, SF; **Investigation:** JR, WS, SF; **Resources:** SF, RW, DK, JE; **Data Curation:** WS, SF, JR; **Writing - Original Draft Preparation:** JR, WS, SF; **Writing - Review & Editing:** JR, WS, DK, PB, CJ, JE, WQ, RW, SF, DP, MC; **Visualization:** JR, WS, SF, CJ; **Supervision:** SF, RW, WQ; **Project Administration:** WQ, MC, DP; **Funding Acquisition:** MC, DP, WQ.

Competing interests: The authors declare no competing interests.

Supplementary materials

Supplementary Text

Figs. S1 to S5

Tables S1, S2

References *(7-54)*

Data S1, S2

Supplementary Materials for

Synergy and antagonism in a genome-scale model of metabolic hijacking by bacteriophage

Jordan C. Rozum[†], William Sineath[†], Doo Nam Kim, Pavlo Bohutskyi, Connah Johnson, James Evans, David Pollock, Wei-Jun Qian, Margaret S. Cheung, Ruonan Wu[*], Song Feng[*]

[*]Corresponding author. Email: song.feng@pnnl.gov (SF); ruonan.wu@pnnl.gov (RW)

[†]These authors contributed equally to this work.

This PDF file includes:

Supplementary Text

Figures S1 to S5

Tables S1, S2

Captions for Data S1, S2

Other Supplementary Materials for this manuscript:

Data S1, S2

Supplementary Text

Additional details about AMGs in *P. marinus*

In *Prochloroccocus marinus*, CP12 inhibits the Calvin-Benson cycle light-limited conditions to shift the metabolism in favor of the pentose phosphate pathway for nucleotide synthesis (*25*). Also favoring the pentose phosphate pathway, the *cobS* AMG affects the reaction responsible for cobalamin (B12) synthesis. Cobalamin is predicted to increase ribonucleotide diphosphatase function in the nucleotide metabolism, thus increasing DNA synthesis (*43*). Similarly working in favor of nucleotide synthesis, the AMG *talC* codes for transaldolase, and enzyme active in the pentose phosphate pathway (*25*).

The next set of AMGs has the largest correlation with the photosynthesis pathway in *Prochloroccocus marinus*. The *psbA* and *psbD* AMGs correlate with the D1 and D2 protein subunits, respectively, at the photosystem II reaction center. The D1 and D2 subunits bind the primary electron transfer chain along with the P680 reaction center chlorophyll special pair (*48*). These AMGs are predicted to have a large effect on energy binding in photosynthesis. The *ho1* AMG catalyzes the reaction responsible for the production of biliverdin, a precursor to light-harvesting pigments (*47*). Similarly, the *cpeT* and *cpcT* AMGs catalyze the production of the light-harvesting pigments phycocyanobilin (*44*). Further, the *pebS* AMG also codes for an enzyme that produces the light harvest pigment phycoerythrobilin (*44*). These AMGs work to increase pigments needed for photosynthesis. Along with AMGs predicted to increase photosynthesis, the *hlip* and *hsp20* AMGs code for proteins that dissipate heat at the photosynthesis reaction center (*45*). These last two genes do not directly participate in host metabolism; however, biologically these AMGs may play a thermodynamic role.

The function of *mazG* as an AMG is relatively unknown. In other organisms, it is predicted to regulate levels of (p)ppGpp, a putative alarmone, and trick the organism into a nutrient-replete state. This would cause the host to use less nutrients and allow the phage to increase replication (*28*). (*53*) argue that during infection of *S. elongatus*, *mazG* may degrade nucleotides to shift GC content in favor of the phage–*S. elongatus* has a GC content near 60%, while P-HM2 has a GC content of only 38%. In contrast, MED4 has an even lower GC content of just over 31%, so the role of *mazG* in this host is less clear. Additionally, the function of the *phoH* AMG remains relatively unclear. There is evidence that it functions as both a phosphate uptake gene and to code for ATPase enzymes (*12*).

However, *phoH* is significantly upregulated during phage infection and this could cause an increase of phosphate available for phage replication (*12*).

3.3.1 Flux balance analysis (FBA)

Flux analysis can be used to extract testable, mechanistic predictions from limited data (*30*). Reaction rate parameters are often unknown or poorly constrained, making a kinetic model nonviable. In contrast, the stoichiometric relationships encoded by chemical reaction networks are often tightly constrained.

For a set of metabolites M governed by reactions R, flux balance analysis (FBA) exploits stoichiometric constraints to determine the steady-state flux distributions that maximize a biological objective $F_{obi}(v_r)$. The *reaction flux* is the instantaneous rate of a reaction $r \in R$ per unit cell mass, here denoted $v_r \in \mathcal{R}$ (typically measured in units of mmol gDW^{-1} h^{-1}). The rate per unit cell mass at which a metabolite $m \in M$ is produced (or consumed) by a reaction $r \in R$ is given by the stoichiometric coefficient, $S_{r,m}$, times the flux v_r of r.

FBA is formulated as a linear optimization problem

$$\text{maximize:} \quad F_{obi}(v_r) \tag{S1}$$

subject to:

$$S_{r,m} v_r = 0 \tag{S2}$$

$$lb \leq v_r \leq ub \tag{S3}$$

where a biologically-inspired objective function, $F_{obi}(v_r)$ such as a biomass reaction encoding the metabolic components necessary to produce one unit (typically one gram) of biomass, is maximized to represent cell growth. The steady state assumption enforces $S_{r,m} v_r = 0$ constraints on flux values and reaction flux bounds, lower bounds $lb \in \mathcal{R}_{\geq 0}$ and upper bounds $ub \in \mathcal{R}_{\geq 0}$, can be used to constrain the solution. Solutions to $S_{r,m} v_r = 0$ that satisfy all specified constraints are called *feasible*; those that do not are called *infeasible*.

Secondary objectives can supplement FBA for more complex models. In *parsimonious* FBA (pFBA), the secondary objective is to minimize the total (absolute) flux of enzyme-catalyzed reactions. Thus, pFBA produces solutions that maximize biomass while simultaneously minimizing enzyme production. In *flux variability analysis* (FVA), each reaction flux is taken, in turn, to be an

objective to minimize or maximize providing the minimum and maximum fluxes that each reaction can support while maintaining optimal growth (with a chosen tolerance). In this study, we have peformed pFBA and FVA using COBRApy (*42*) with the GNU Linear Programming Kit (glpk) back-end (*54*).

Viral infection can be modeled as a perturbation to the host system. It has been demonstrated experimentally that gene knockout perturbation response often minimizes flux change rather than maximizes biomass production (*29, 32*). When considering the impact of AMG-hijacked reactions individually or pair-wise, we specify a target range for each hijacked reaction by taking 10% of the interval in which phage production can be supported above half its maximum value (see Figure 1). Following (*29*), to compare to a baseline flux distribution, we restrict the AMG-hijacked reaction(s) to the corresponding target range(s) and perform minimization of metabolic adjustment (MOMA) to find the flux distribution that satisfies the new constraints and is closest to the baseline (by L1 norm).

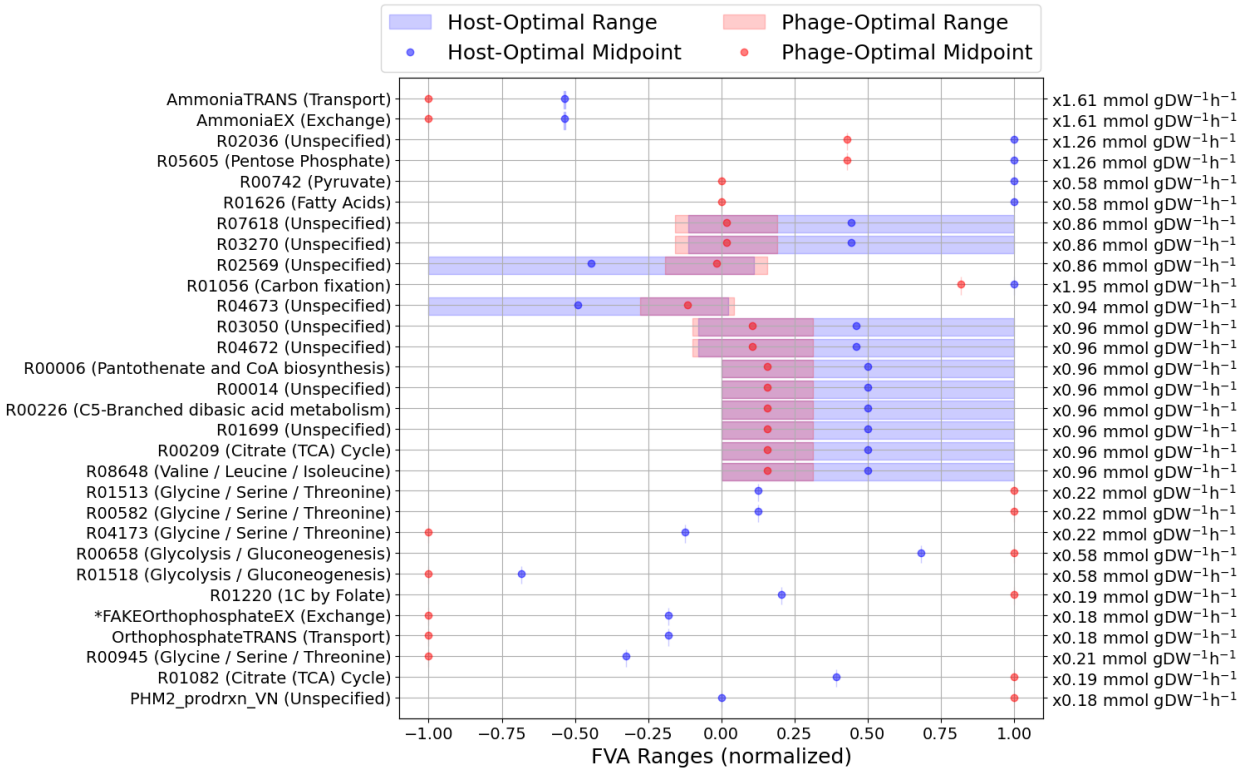

Figure S1: **FVA ranges for 30 reactions most affected by optimizing for phage vs host biomass.** Blue ranges indicate flux variability when maximizing for host biomass while red ranges are for the variability associated with maximizing the phage biomass. Any overlap is shown in purple. Ranges are normalized by the most extreme feasible flux across the two scenarios (normalization factors depicted along the right-hand side of the figure). Reactions with normalization below that of phage production (PHM2_prodrxn_VN) are not included in this analysis. Reaction IDs and subsystems are given along the left-hand side. Reactions are ordered by (unnormalized) midpoint distance. AMG-hijacked reactions are indicated with an asterisk.

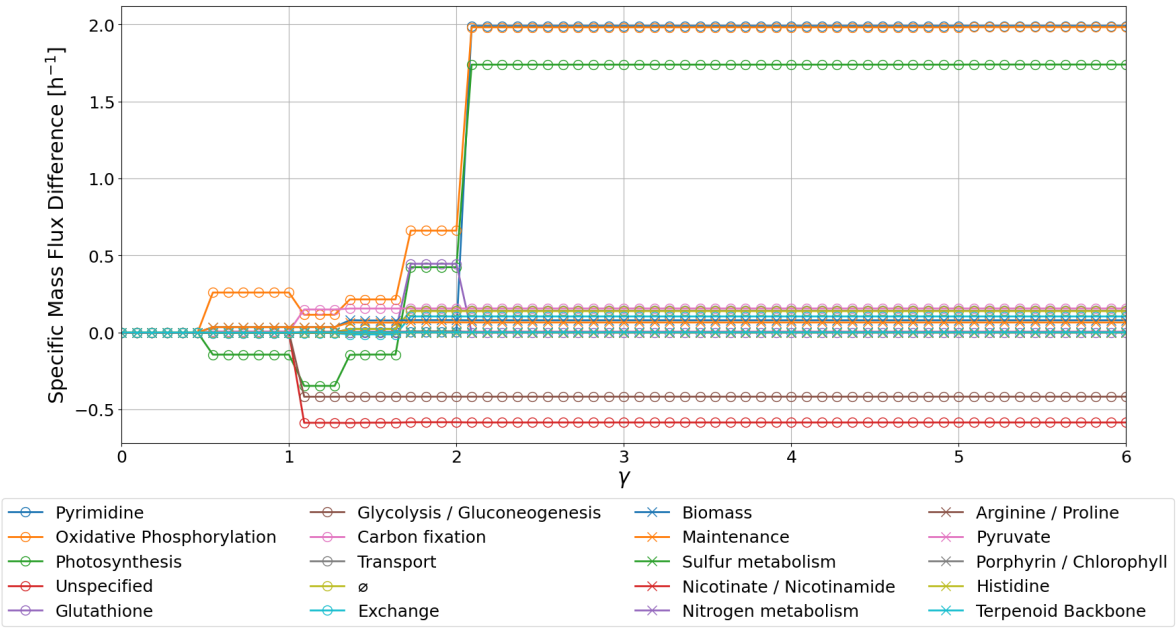

Figure S2: **Changes in specific mass flux through the 20 most impacted subsystems.** Results are shown as a function of AMG reaction emphasis factor, γ. The ten subsystems with the largest responses are marked with circles, the remaining ten use an X symbol. Increasing γ more heavily weights changing the flux of AMG-hijacked reactions relative to minimizing the flux through other gene-regulated reactions (see Methods). In contrast to Figure 4, the vertical axis uses a linear scale.

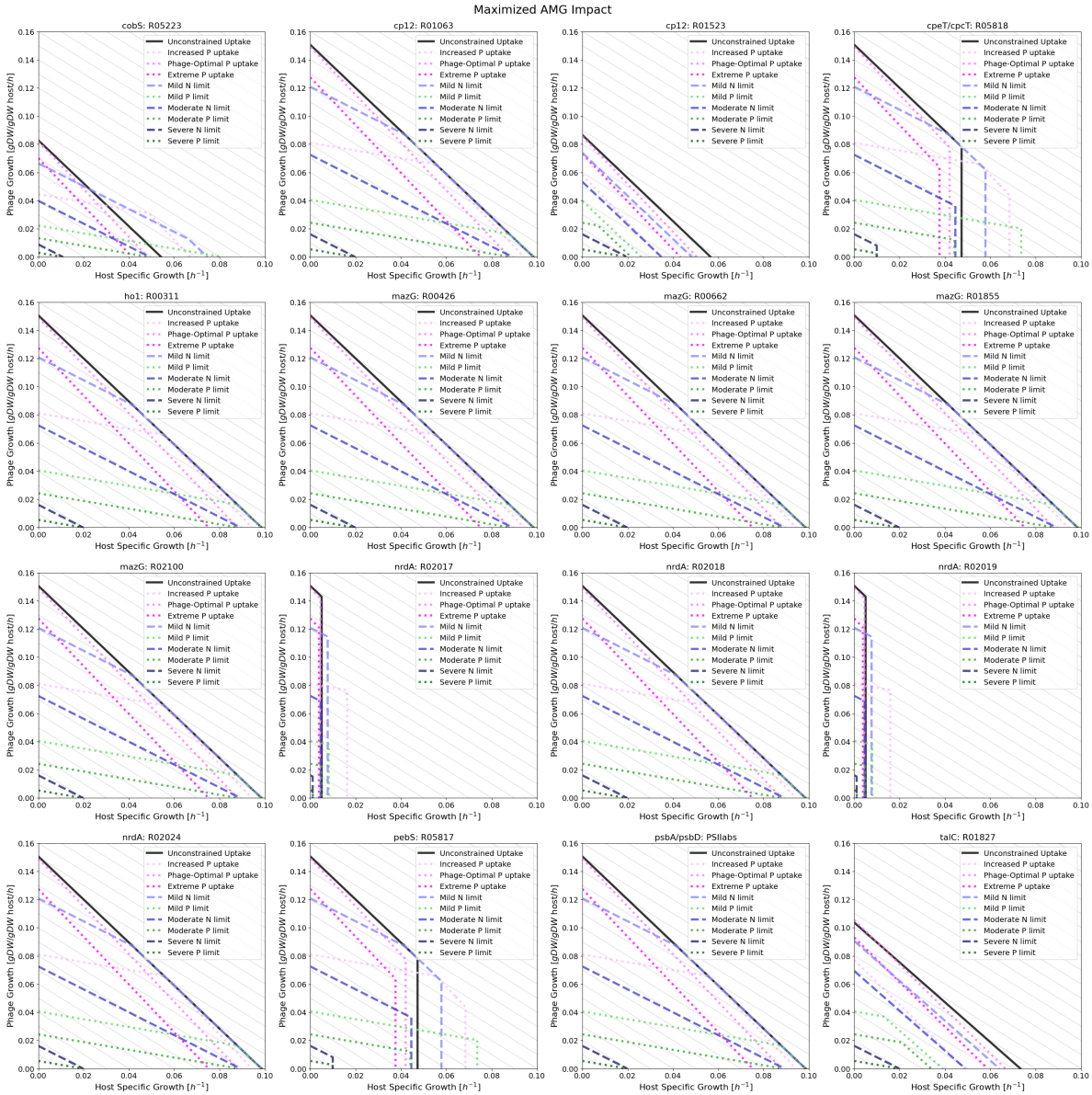

Figure S3: **Phage-host biomass Pareto fronts.** Full size image available online. Each panel depicts the instantaneous relationships between maximum (exponential) host specific growth rate and maximum (linear) phage growth rate for an AMG-hijacked reaction, as indicated in the figure titles. We show Pareto fronts for several nutrient-limiting scenarios: unconstrained uptake, and mild/moderate/severe limiting of N or P uptake. Mild, moderate, and severe limits correspond to 150%, 90%, and 20% of the values required for maximal host biomass production, respectively. When growth is not constrained by nutrient uptake, it is instead limited by RuBisCO efficiency. In addition, we also considered two scenarios in which the cell is forced to uptake more phosphate than is optimal (Increased P uptake: 3x host-optimal uptake; Phage-optimal P uptake: 6x host-optimal uptake; Extreme P uptake: 12x host-optimal uptake).

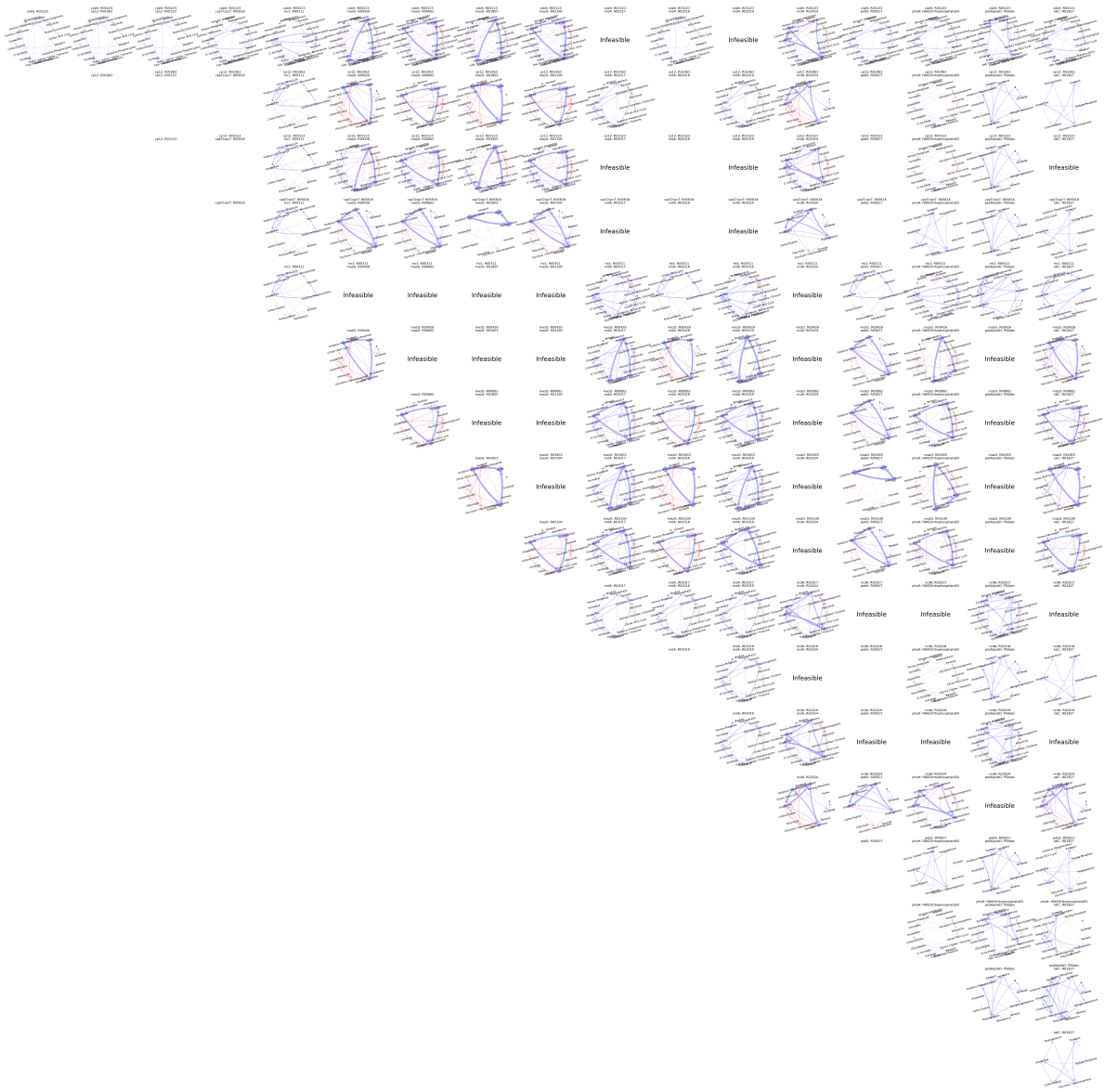

Figure S4: **Pairwise AMG-hijacked reaction impacts on subsystem fluxes.** Full size image available online. Each column and row corresponds to an AMG-hijacked reaction. Along the diagonal, a single reaction is hijacked. Along the upper off-diagonal, pairs of reactions are hijacked. The resulting MFD graphs (see Methods) are plotted, with blue (red) indicating increased (decreased) mass flux (changes below 0.01h^{-1} are not shown). The reactions hijacked are indicated with text labels on each graph plot. When simulatneous hijacking is not possible due to stoichiometric constraints, the text "Infeasible" appears in place of a graph plot.

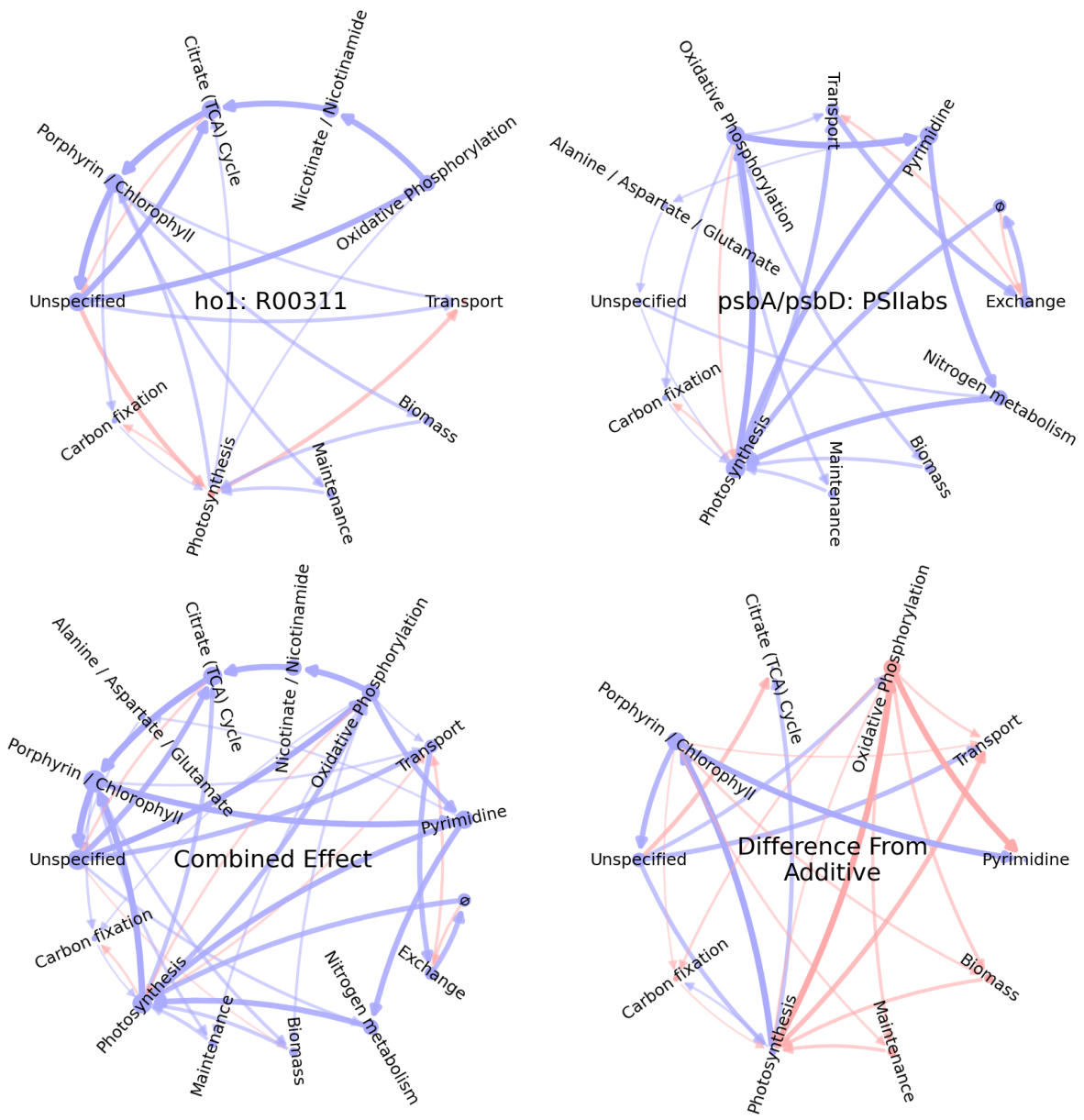

Figure S5: **Deviation from the additive impact of two AMG-hijacked reactions.** In the top row, the MFD graphs for PSIIabs and R00311 are shown. The bottom left graph is the MFD graph (see Methods) for their combined impact. The bottom right graph shows the deviation of the combined impact MFD from a purely additive impact. For example, the red edge from Oxidative Phosphorylation to Photosynthesis denotes that hijacking both reactions has a smaller impact on this flux than the sum of the individual impacts. In all graphs, edges with specific mass flux changes below 0.01h^{-1} are not shown, nor are nodes that lack incoming or outgoing edges above this threshold.

Gene: Reaction ID	Reaction Expression	Subsystem
cobS: R05223	[Adenosine-GDP-cobinamide] + [alpha-Ribazole] ⇀ [Cobamide coenzyme] + [GMP] + [H+]	Porphyrin / Chlorophyll
cp12: R01063	[D-Glyceraldehyde 3-phosphate] + [NADP+] + [Orthophosphate] ⇌ [3-Phospho-D-glyceroyl phosphate] + [H+] + [NADPH]	Unspecified
cp12: R01523	[ATP] + [D-Ribulose 5-phosphate] ⇌ [ADP] + [D-Ribulose 1,5-bisphosphate]	Carbon fixation
cpeT/cpcT: R05818	[Biliverdin] + [Reduced ferredoxin] ⇀ [15,16-Dihydrobiliverdin] + [Oxidized ferredoxin]	Porphyrin / Chlorophyll
ho1: R00311	[Heme] + 3.0 [Oxygen] + 3.0 [Plastoquinol] ⇀ [Biliverdin] + [CO] + [Fe2+] + 3.0 [H2O] + 3.0 [Plastoquinone]	Porphyrin / Chlorophyll
mazG: R00426	[GTP] + [H2O] ⇀ [Diphosphate] + [GMP]	Purine
mazG: R00662	[H2O] + [UTP] ⇀ [Diphosphate] + [UMP]	Pyrimidine
mazG: R01855	[H2O] + [dGTP] ⇀ [Diphosphate] + [dGMP]	Purine
mazG: R02100	[H2O] + [dUTP] ⇀ [Diphosphate] + [dUMP]	Pyrimidine
nrdA: R02017	[ADP] + [Thioredoxin] ⇀ [H2O] + [Thioredoxin disulfide] + d[ADP]	Purine
nrdA: R02018	[H2O] + [Thioredoxin disulfide] + [dUDP] ⇌ [Thioredoxin] + [UDP]	Unspecified
nrdA: R02019	[GDP] + [Thioredoxin] ⇀ [H2O] + [Thioredoxin disulfide] + d[GDP]	Purine
nrdA: R02024	[CDP] + [Thioredoxin] ⇀ [H2O] + [Thioredoxin disulfide] + d[CDP]	Pyrimidine
pebS: R05817	[Biliverdin] + [Reduced ferredoxin] ⇀ [(3Z)-Phycocyanobilin] + [Oxidized ferredoxin]	Porphyrin / Chlorophyll
phoH: FAKEOrthophosphateEX	[Orthophosphate] ⇌ ∅	Exchange
psbA/psbD: PSIIabs	[Photon] ⇀ [Photon absorbed by PSII]	Photosynthesis
talC: R01827	[D-Glyceraldehyde 3-phosphate] + [Sedoheptulose 7-phosphate] ⇌ [D-Erythrose 4-phosphate] + [D-Fructose 6-phosphate]	Pentose Phosphate

Table S1: **Chemical reaction expressions and subsystems for AMG-hijacked reactions.**

Amino Acid Consumed	mmol/g virus
Alanine	00.428
Arginine	00.195
Asparagine	00.312
Aspartate	00.300
Cysteine	00.023
Glutamate	00.247
Glutamine	00.180
Glycine	00.422
Histidine	00.050
Isoleucine	00.269
Leucine	00.310
Lysine	00.192
Methionine	00.090
Phenylalanine	00.185
Proline	00.164
Serine	00.389
Threonine	00.451
Tryptophan	00.039
Tyrosine	00.196
Valine	00.321

Nucleotide Consumed	mmol/g virus
ATP	20.455
CTP	00.005
GTP	00.006
UTP	00.007
dATP	00.378
dCTP	00.233
dGTP	00.233
dTTP	00.378

Nucleotide Produced	mmol/g virus
AMP	00.009
ADP	20.446
CMP	00.005
GMP	00.006
UMP	00.007

Small Molecule Consumed	mmol/g virus
H2O	20.446

Small Molecule Produced	mmol/g virus
H	20.446
Orthophosphate	20.446
Diphosphate	01.249

Table S2: **Biomass production reaction for P-HM2.** Phage mass is modeled as the mass deficit in this reaction due to amino acid and dNTP consumption. Stoichiometric coefficients and reaction velocity units are such that a reaction rate of 1 corresponds to 1 gram phage produced per hour per gram dry weight of host.

Component	MED4 (g/gDW host)	P-HM2 (g/g virus)
Protein	0.6492	0.6007
DNA	0.0131	0.3993
RNA	0.0525	0.0132
Lipids	0.1284	
Other	0.1824	

Table S3: ***Prochloroccocus marinus* and P-HM2 biomass components.** RNA mass for P-HM2 corresponds to mRNA synthesis needed for transcription of genes encoding structural proteins, and does not directly contribute to phage mass in this model.