Adaptation to a phenotypic trade-off requires an interplay between local and global regulators

Matthew Deyell, Andrew D. Griffiths, Sander J. Tans, Philippe Nghe

Abstract

The fitness of organisms depends critically on regulatory strategies that balance multiple phenotypes, including physiological, developmental, and pathogenic. An important example is the trade-off between motility and growth determine competition for habitat colonization by microorganisms and the invasiveness of metastatic cells in cancer. Phenotype regulation relies on hierarchical gene networks, but how gene hierarchy contributes to adaptation remains unclear. Here, we show that global upstream regulators allow cells to adjust the impact of downstream local regulators in the context of a phenotypic trade-off. We performed single and double knockdowns of transcription factors in *Escherichia Coli* and measured their impact on swimming and growth in different environments. We observed that perturbations of local regulators affect only a specific phenotype independently of the genetic background, whereas perturbations of global regulators affect either one or both phenotypes to varying degrees depending on the genetic background. Global and local regulators strongly interact giving rise to sign, reciprocal sign and masking epistasis. Epistasis between particular genes is modulated by the environment, but only a restricted number of epistatic patterns are observed across environments. These epistatic constraints suggests that that local regulators should be typically acquired before tuning of the phenotypic trade-off by global regulators. These findings highlight the importance of pleiotropic regulators for coordinating phenotypic responses to complex environments and providing flexibility during evolution.

Introduction

The fitness of biological organisms generally depends on multiple phenotypes, but it is not always possible to optimize all of them due to trade-offs [1]. An important example is the trade-off between growth and mobility, which has been demonstrated in bacteria and cancer cells [2-6]. Disseminated tumour cells are highly mobile, which promotes metastasis, but have a slow-cycling state that makes radiation therapy and chemotherapies poorly effective [2-5]. In bacteria, the trade-off between mobility and growth underlies the maintenance of population diversity through niche formation [6-10].

Trade-offs between phenotypes can be analysed using the concept of Pareto optimality, originally introduced in economics. A trade-off means that improving phenotype A may require the deterioration of phenotype B, and vice versa. When one organism excels at A and another at B, both these specialists may be maintained during evolution within a changing environment [11], or may coexist in distinct environmental niches [6, 12]. Such organisms that cannot be outcompeted on all functions simultaneously are said to be Pareto optimal. In graphs showing the phenotype B performance against the phenotype A performance for different organisms, Pareto optimal organisms fall on a line connecting the extremal or archetypical phenotypes that are specialized for a single task [1]. Hence, Pareto optimal organisms consist of a well-defined group of phenotype combinations that interpolate extremal archetypes, or specialists. Such set of optimal combinations of phenotypic values is called the Pareto front.

Regulation can partially mitigate trade-offs by modulating phenotypes in response to external or internal signals [11]. Genome-wide transcriptional regulation networks [13] have been shown to control growth

(e.g. by ribosome synthesis) and motility (e.g. by flagella expression) in bacteria [11], and hence determine the outcome of competition between strains [14]. Such networks are often hierarchical, with a few global regulators regulating many downstream local regulators [15, 16]. A well-known example is the lac operon, which is controlled by the local lacI and the global CRP regulator. Indeed, CRP participates in hundreds of operons in *E. coli* like several other global regulators do such as HNS and Fis, for instance. Local regulation is often combined with multiple global regulators, resulting in highly interconnected control networks. Given this complexity, it remains unclear how these hierarchical regulatory motifs allow for the coordination of multiple phenotypes.

While gene regulatory networks can mitigate phenotypic trade-offs, the network architecture may in turn also limit adaptation. For instance, to improve the overall phenotype of a hierarchical regulatory cascade, its regulatory genes may need to be mutated in a particular order, owing to sign epistasis between them [17]. More generally, any given regulatory network is bound by a limited range of phenotypes that it can produce, and overcoming such bounds may require evolving distinct network structures [18]. Global regulators are expected to be strongly constrained in their evolution due to their highly pleiotropic nature [19, 20].

To better understand the genetic basis of adaptations involving multiple phenotypes, we study here a model system where local and global regulators interact to determine the outcome of two quantitative phenotypes that are in a trade-off: growth and swimming in *E. coli*. We used CRISPR interference method to knock down combinations of transcription factors: CRP, Fis, HNS, mcaS, CsgD, FlhDC, and FliZ. These genes have already established hierarchical relationships in the gene regulatory network of *E. coli* [13] (Figure 1A). Additionally, these genes have been shown to contribute to the regulation of one or both measured phenotypes: mobility and growth. FliZ and FlhDC directly control expression of the flagella operon essential for swimming mobility [21]. CsgD directly controls expression of the Curli pathway and adhesion pili and represses swimming mobility [22]. CsgD is itself largely controlled by the small inhibitory RNA mcaS [23, 24]. All of these mobility regulating transcription factors are in part regulated by upstream global transcription factors CRP, Fis, and HNS [13]. CRP regulates the expression of secondary catabolites in response to cAMP, in addition to many other cellular processes [25]. Fis is one of the most highly expressed proteins during the exponential growth phase and is involved in the regulation of genes involved in translation and metabolism, among others [26]. HNS is traditionally thought to regulate horizontally acquired genes [27], and it also controls a wide range of genes involved in environmental adaptation and response to stress [28]. All three global transcription factors CRP, Fis, and HNS also contribute strongly to the nucleoid structure and maintenance [29].

We measure how single and pairwise knockdowns of local and global regulators alter these two phenotypes. This allows us to analyse the coupled variations of phenotypes in response to these genetic perturbations. We further examine the genetic interactions (epistasis) between pairs of local and global regulators for phenotypes taken separately or together. Epistasis is usually considered for a single fitness or phenotype and quantifies how much the effects of two mutations together deviate from the added effects of each mutation individually. However, epistasis is altered by the environment for specific gene pairs or specific phenotypes, showing the existence of high order interactions between genes and environment [30]. We find that epistasis and phenotypic couplings follow a restricted range of patterns for a hierarchy between regulatory genes, where global regulators adjust the balance between optimal phenotypes.

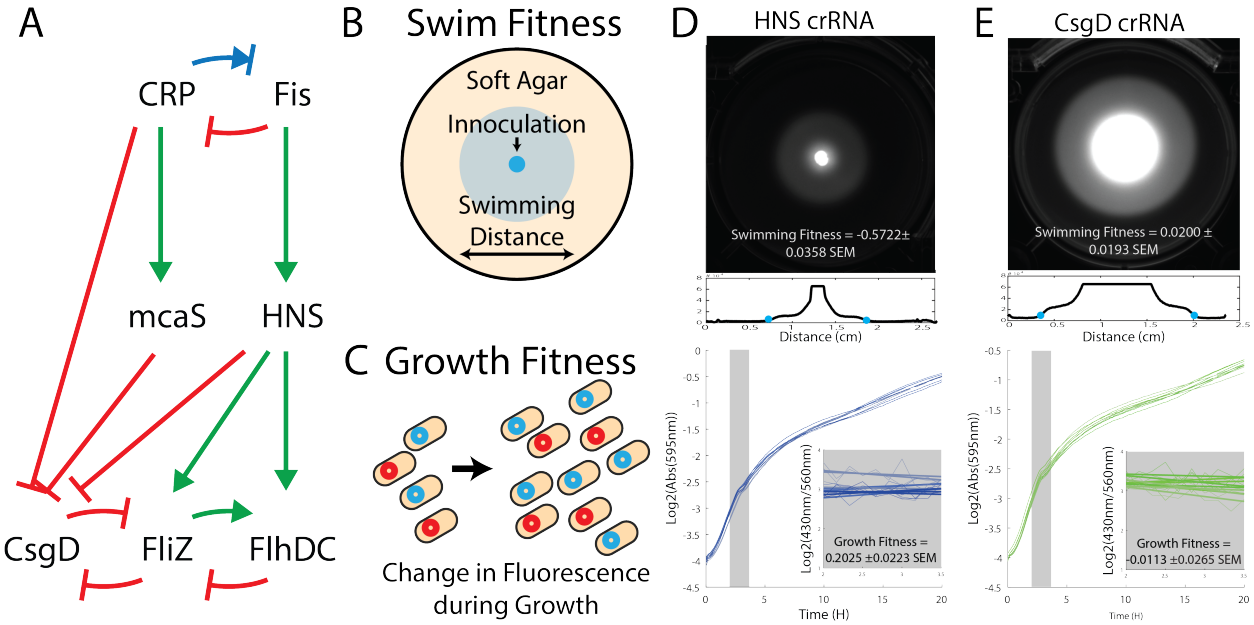

Figure 1: Quantification of swimming and growth phenotypes. (A) Transcription factors regulating mobility and growth are arranged in a hierarchical structure. The global transcriptional regulators CRP and Fis mutually regulate each other at the top of the pathway. At the bottom of the pathway, the transcription factors for twitching mobility (CsgD) and swimming mobility (FliZ) are arranged in a toggle switch motif. CRP can directly regulate the downstream CsgD transcription factor through a feed-forward repression motif with mcaS. In contrast, Fis can indirectly regulate each of the downstream mobility transcription factors through activation of the global transcriptional regulator HNS as an intermediate. Each of these transcription factors was targeted with CRISPRi perturbation, either alone or in combination. (B) Swimming fitness of each perturbation was determined through inoculation in the centre of soft agar plates and the diameter of the swimming front was measured. Swimming fitness was normalized to the swimming front of the reference control. (C) Growth fitness was determined as the change in fluorescence over time of the perturbation strain (cyan) grown in competition with a reference strain of another fluorescence (red). (D) Example of raw data from HNS perturbation, which results in a decrease in swimming fitness (top) but an increase in relative growth rate (bottom). (E) Example of raw data from CsgD perturbation, which results in a slight increase in swimming fitness (top) but a slight decrease in relative growth rate (bottom). HNS and CsgD are found on opposite ends of the Pareto-front formed between swimming and growth phenotypes.

Results

Characterizing relationships between genetic perturbations, phenotypes, and environments

We used CRISPR interference (CRISPRi) to generate single and double knockdown (KD) perturbations of 7 transcription factors. Cells were transformed with vectors expressing an inducable catalytically deactivated Cas9 (dCas9), mCerulean to fluorescently tag the cells, and a CRISPR RNA (crRNA) that targets the transcription start sequence of the relevant transcription factor (see Methods). CRISPRi provides a simple method for reliably repressing gene expression as the dCas9 acts as a programmable transcription factor [31]. We quantified two phenotypes which determine the fitness of *E. coli*: motility and growth (Figure 1). Motility was quantified using a migration assay in 0.3% soft agar plates. Mobile bacteria form concentric bands while swimming in soft agar plates (Figure 1B). These bands arise due to local gradients of nutrients formed by cell consumption [32, 33]. All mobile strains showed reproducible and uniform

swimming waves for each strain (supplementary figure 3). The diameter of disk resulting from these waves differed between strains, confirming the effectiveness of the KDs on this phenotype. We defined the motility performance as the diameter of the first swimming wave after 16 hours in LB media or 24 hours in M63 media.

Growth performance was quantified using a competition assay against an unperturbed reference strain in microwell plates under constant shaking agitation (figure 1C). Growth was determined based on three signals: optical density, mCerulean fluorescence of the query strain, and mCherry fluorescence of the reference stain. The relative growth rate is computed as the slope of the logarithm of the ratio between fluorescence signals and is restricted to the exponential phase as controlled using total OD. This growth rate is normalized using the OD and fluorescent measurements to remove contributions due to differences between constructs and fluorescent markers, including: maturation times, degradation rates, and expression costs (see Methods). Each fitness measurement was performed in 9 replicates, consisting of three biological replicates (different days and initial colonies), each comprising three technical replicates each (same day, different wells or plates). The growth values were taken as the median of all 9 measurements during the exponential growth phase. Most of the variability arose between biological replicates with much lower variation between technical replicate (SI Table 1).

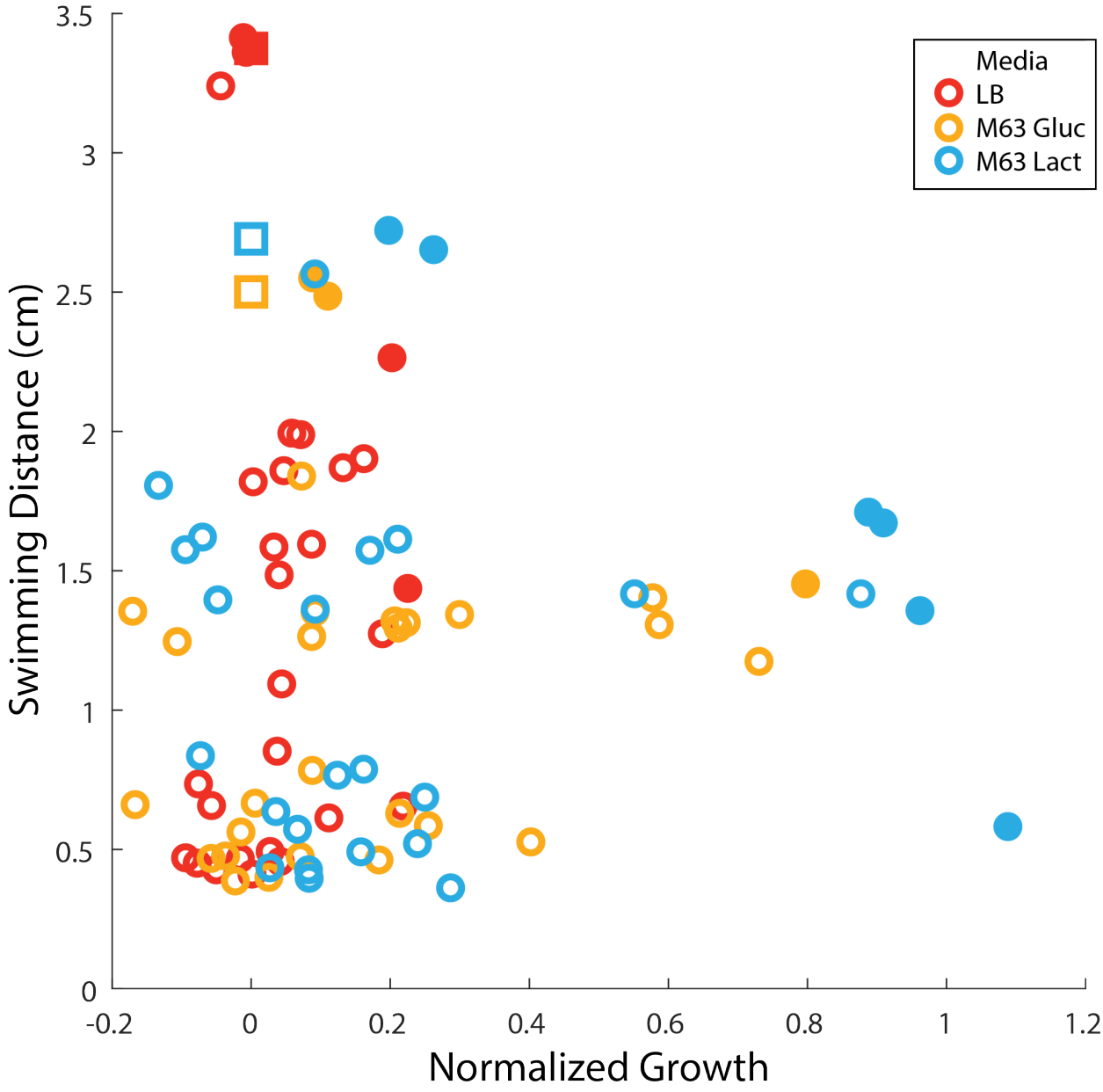

Figure 2: Coupled phenotype space displays a Pareto front. The distance that each perturbation is able to swim is plotted against the normalized growth fitness for each strain. Perturbations that are able to grow faster than the reference strain tend to also show a decrease in the distance that they are able to swim after 16 hours. Squares demonstrate the fitness of the reference strain. Solid dots indicate strains which are found along the Pareto front.

Motility network perturbations are bound by a growth-motility trade-off.

To understand how genetic perturbations impact phenotypes, we examined the distribution of phenotypes within a 2D space (Figure 2) consisting of the swimming distance (Y axis) and growth rate difference (x axis) for all 28 KD strains in all three environments. Single and double KDs cause mild to strong effects on both phenotypes, with growth rate differences ranging from -0.17 h^{-1} to 1.09 h^{-1}, and

swimming diameter ranging from 0.4 cm to 3.4 cm after 16 hours. These differences are measured in comparison with the reference strain, which has growth rates of: 0.56 doublings per hour in LB, 0.39 doublings per hour in M63 with Glucose, 0.40 doublings per hour in M63 with Lactose; and which swimming diameters are: 3.37cm in LB, 2.5 cm in M63 Glucose, 2.69 cm in M63 Lactose. Growth was found to both increase and decrease due to the genetic perturbations while swimming distance was found to only decrease. However, perturbation effects do not independently affect each phenotype. For example, mutations that increase growth rate tend to also decrease swimming distance (Figure 2), indicating a phenotypic trade-off. Pareto optimal points are such that growth does not increase without decreasing swimming and vice versa. For example, in figure 2 the reference strains (squares) are Pareto optimal because there are no other points in a given environment that are higher in both swimming and growth fitness.

The different environments have contrasting effects on the different phenotypes. On the one hand, they alter the range of accessible growth as seen in figure 2, with the growth varying less for higher quality nutrients. The largest relative changes in growth rates are observed in the poorest medium – lactose (1.24 fold change), followed by glucose (0.97) and finally LB (0.32). On the other hand, the range of accessible swimming phenotypes is quite similar across environments. Mainly, the best performing strain in LB swims 0.35-fold further than the best strains in M63 and glucose media. Importantly, a growth-migration trade-off is observed across all environments, with 4, 6, and 3 Pareto-optimal strains observed in LB, M63 Glucose and M63 Lactose respectively (out of 29 strains in total). Hence, these data indicate that genetic perturbations in the global motility-control network are bound by a growth-motility trade-off. The findings are consistent with growth-motility Pareto fronts observed during evolution [34, 35]. In contrast, our directed genetic perturbations allow us to cover a larger network of motility regulators (Figure 1A), and to also study non-optimal phenotypes. The latter constitute over 85% of our data set, cover a wide range of phenotypic values, and are key to understanding the role of network hierarchy and epistatic interactions.

Coupling between phenotypes depends on gene hierarchy.

To analyse the relationship between regulatory network structure and phenotypic effects, we first examined how changes in growth and swimming were associated with each given transcription factor. In figure 3, each panel reports the phenotypic effects of a given knockdown (KD) in all genetic backgrounds and environments. Here, a vector along the x-axis means the KD affect growth only, while a vector along the y-axis means the KD affects swimming only. Diagonal vectors indicate pleiotropy or a coupling between growth and swimming. If the coupling is positive, growth and swimming improve or deteriorate both for a single KD. If the coupling is negative, improved growth comes at the expense of deteriorated swimming, or vice versa, growth deteriorates while swimming improves in growth.

We found that the effects of the knockdowns depend both on genetic background and environment, with CRP, Fis, and HNS showing the most diverse responses (Figure 3). Indeed, KDs of these transcription factors affected growth exclusively, motility exclusively, or showed a coupling between both phenotypes. Perturbations of mcaS and csgD also had qualitatively varied effects, but these were lower in magnitude. Indeed, more than 50% of the phenotypic effects were not statistically significant ($p>0.05$ Welch t-test $n = 9$). In contrast, FliZ and FlhDC had strong effects on motility regardless of environment, while showing little impact on growth.

The three groups of transcription factors described above correspond to different positions within the regulatory network hierarchy (Figure 1). CRP, Fis, and HNS are upstream global regulators. Together, they accounted for 69% of perturbations that show negative growth-motility coupling (vectors that point top-left or bottom-right in Figure 3), mostly caused by HNS (n = 16). Negative couplings are of particular interest as they are causal manifestations of the growth-motility trade-off, at the level of single genetic perturbations. The regulators with the mildest effects, mcaS and csgD, occur in an intermediate position in the network hierarchy. Finally, FliZ and FlhDC both had the strongest impact on swimming but little impact on growth. Both are downstream transcription factors dedicated to direct flagella regulation. FliZ and FlhDC accounted for 60% (6/10) of the few cases (7% of all perturbations) where swimming and growth positively coupled, with both phenotypes decreasing.

This analysis indicates that downstream local transcription factors mostly modulate swimming, as may be expected, but upstream global transcription factors modulate both growth and swimming, inducing a large variety of ways the two phenotypes couple. However, it remains unclear at this stage how the diversity effects caused to perturbations in global regulators leads to the distribution phenotypes observed in figure 2.

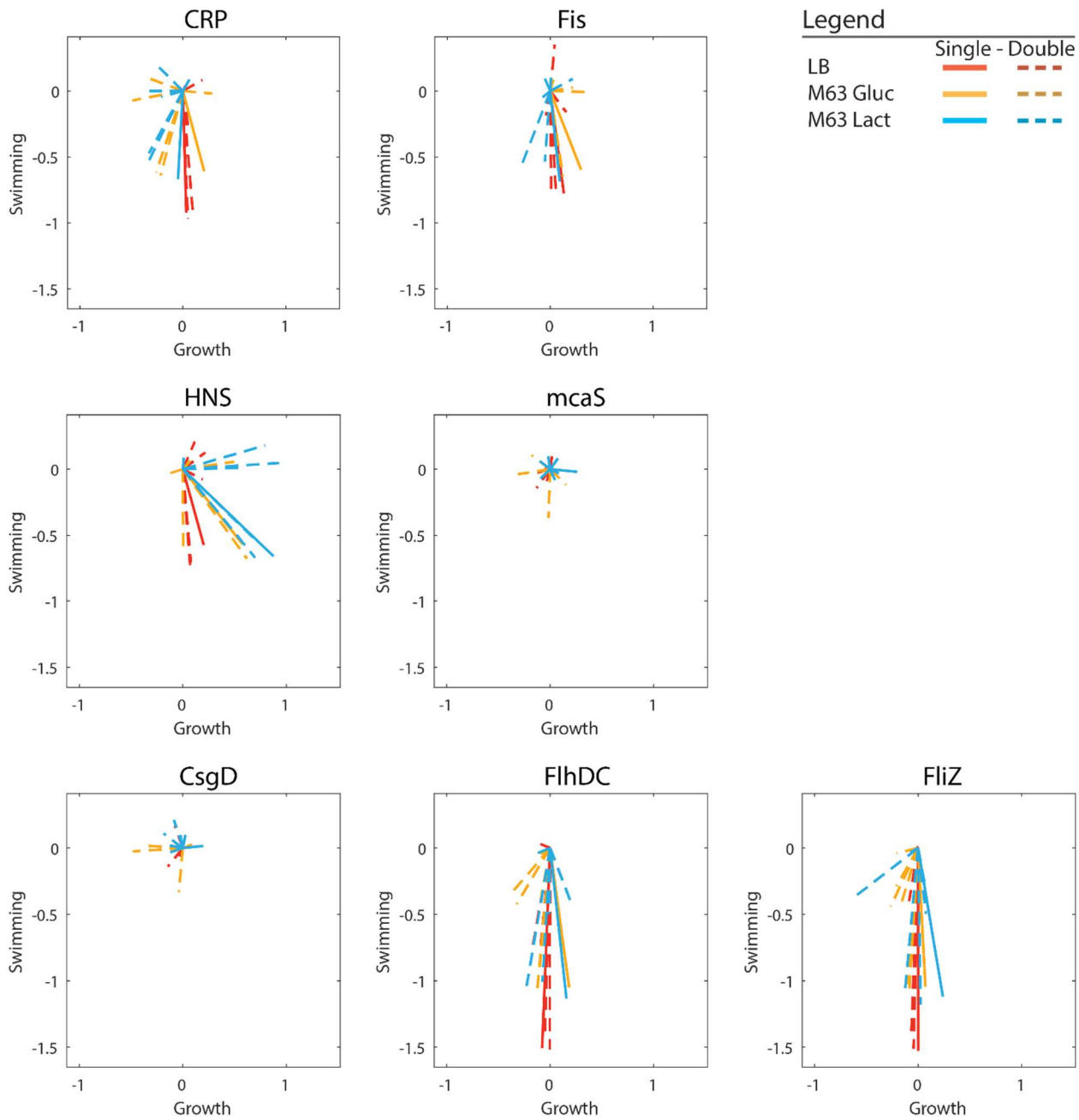

Figure 3: The regulatory structure influences the magnitude and direction of phenotypic effects. The vectors representing the effect of each perturbation in the combined growth/swimming phenotype space were all normalized to have the same origin of 0,0 to compare the direction and amplitude of each perturbation in different genetic backgrounds or growth media. The effects of perturbations to the global transcriptional regulators at the top of the hierarchy tend to be diverse in both amplitude and direction, while those of the flagella regulators FlhDC and FliZ tend to have very strong negative effects on swimming fitness, with more constrained effects on growth. In contrast to both groups, the intermediate regulators of mcaS and CsgD result in a highly diverse range of effects in terms of direction but were very weak.

Epistasis is pervasive and modulated by environment.

To understand the influence of regulatory interactions on phenotypic effects, we quantified the epistasis between each pair of transcription factors. Epistasis is indeed expected to depend on the architecture of gene networks [36-38]. Departure from the additive expectation indicates a genetic interaction which may take the various forms including magnitude, masking, or sign epistasis (Figure 4A). Magnitude epistasis conserves the sign of the fitness effect across backgrounds but not the amplitude of the effects. The extreme form of magnitude epistasis is masking epistasis in which a genetic perturbation has no further effect if another perturbation has already occurred. Sign epistasis refers to the case where a mutation causes a fitness increase in one background but a decrease in another background. When sign epistasis applies to both genes being perturbed, it is referred to as reciprocal sign epistasis. While one typically considers a single fitness or phenotype measure when studying epistasis, here we consider growth and swimming (Figure 4B and C). Epistasis can be represented for both growth and swimming, which can reveal coupling between the two (Figure 4B). Here, the black dot represents the reference strain, the grey dots the single KDs, and the white dot the double KD. The edges represent the perturbations induced by each KD of the corresponding colour (gene list on the left and bottom).

In figure 4, the particular shape taken by the polygon generated by the double KD trajectories characterizes the epistatic interactions [30]. For instance, the null additive case (no epistasis) for both phenotypes yield a parallelogram. The absence of significant epistasis for both phenotypes was only observed once among 63 double KD trajectories (1.6%). This shows that epistasis is pervasive and usually present in at least one of the two phenotypes. The classification of epistasis from each of the phenotypes independently confirms the presence of significant epistasis in 78% of cases (n = 63). Most of this epistasis is in the form of masking (60%) and sign epistasis (10%) (Figure 4). These forms of epistasis are strong as they nullify or invert the phenotypic effect respectively, in contrast to magnitude epistasis, which increases or decreases the size of the perturbation effect while maintaining the direction (either increasing or decreasing fitness).

Among all epistatic patterns, the most represented shape is the triangle. This shape results most often when one KD has negligible effects in one of the two genetic backgrounds. This pattern is observed in 9 out of 21 KD pairs in LB media: HNS-Fis, mcaS-Fis, mcaS-HNS, CsgD-CRP, CsgD-Fis, CsgD-HNS, FlhDC-CsgD, FliZ-mcaS, and FliZ-FlhDC. In other cases, the triangular shape is characterized by one genotype being positioned on the line connecting two others, as observed for another 4/21 double KDs: Fis-CRP, FlhDC-CRP, FlhDC-HNS, and FliZ-Fis. For example, in the latter case the double KD is mid-way between the single KDs and the overall pattern results from a combination of masking and sign epistasis. In other cases, a KD has a negligible effect on both fitness metrics, leading to a collapsed trajectory (mcaS-CRP, FlhDC-mcaS). In contrast, some polygons are quadrilaterals that occupy a large area such as HNS-CRP, FlhDC-Fis, and FliZ-HNS.

The above classification reveals strong genetic interactions. Notably, there is not a one-to-one relationship between a pair of perturbed genes on the one hand, and a specific epistatic interaction on the other hand. First, epistasis depends on the phenotype that is considered. Indeed, epistasis is in most cases dissimilar between growth and motility (17/21) as represented by the coloured squares on the top-right of each KD pair in figure 4C. Second, the epistatic interaction for a given KD pair is not conserved across environments (SI Figure 4), consistent with other studies [30].

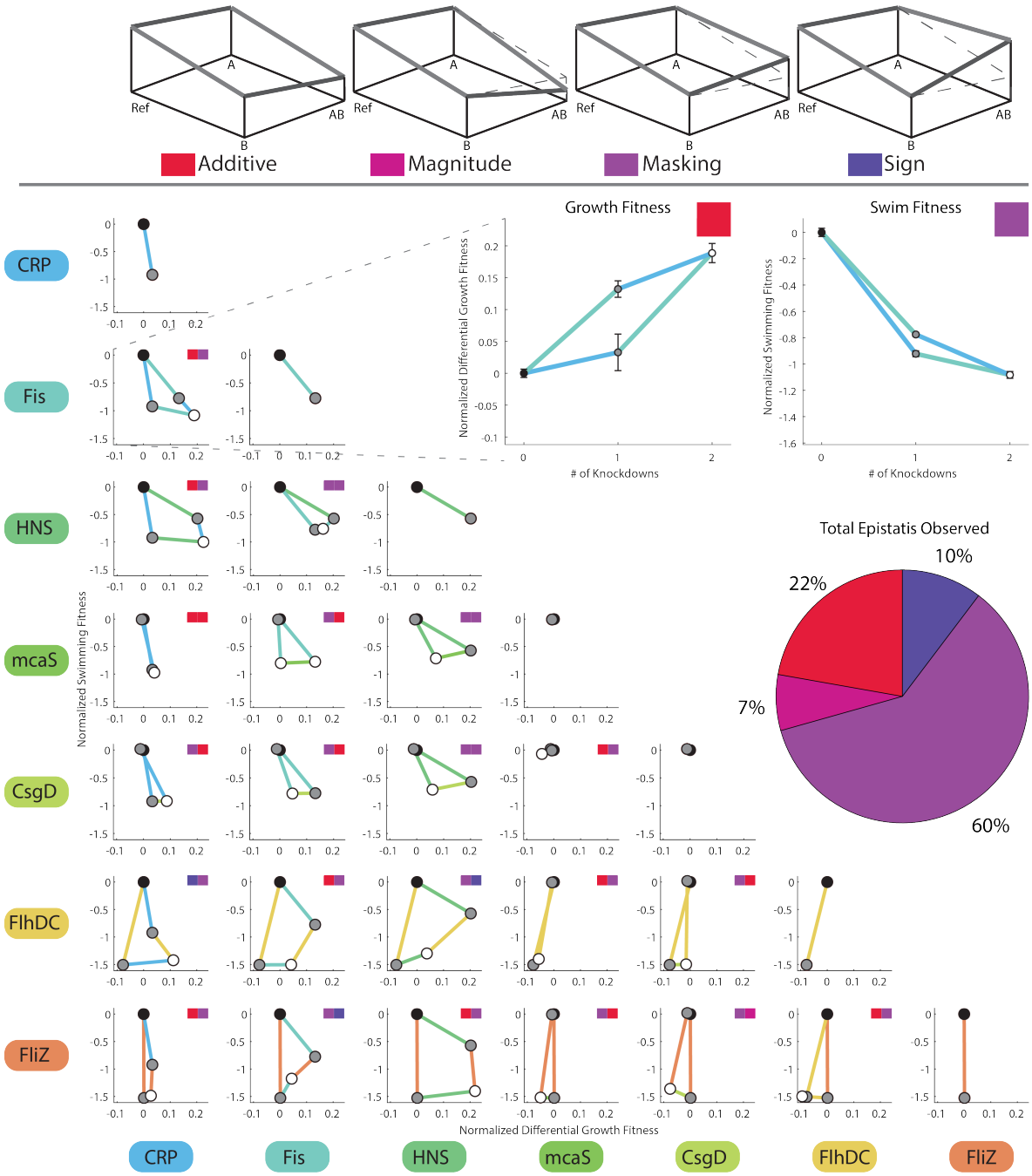

Figure 4: The epistasis of pairwise perturbations of transcription factors. (Top) Definitions of the observed epistatic interactions. If the effects of ref->A, B->AB and ref->B, A->AB are parallel, the effects are additive and no epistasis is observed. The effect of A does not depend on the presence of B and vice-versa. In contrast, if the effects of A are increased or decreased by the presence of B, but the direction of the effect is the same, we refer to this as magnitude epistasis. If the effect of A is lost in the presence of B, such that AB = A ≠ ref, we refer to this as masking epistasis. If the direction of the effect of A is changed by the presence of B, we refer to this as sign epistasis. (Bottom) The plotted growth and swimming fitness of each combination of mutants in LB media with the corresponding observed epistasis pattern for either growth (left insert) or swimming (right insert). The epistasis pattern tends to not be conserved between each measured fitness. (Top insert) Example of epistasis calculation of CRP-Fis combination for both growth (left) and swimming (right). (Bottom insert) The total epistasis patterns observed across all perturbation combinations and media.

Epistasis is restricted to only a few patterns determined by regulatory hierarchy.

Despite the apparent complexity of epistasis and its modulation by the environment, recurring patterns are apparent when considering regulatory hierarchy: (i) the upstream global regulators CRP, Fis, and HNS; (ii) the intermediates mcaS and CsgD; (iii) the downstream local regulators FlhDC and FliZ. For instance, interactions between global regulators and other global or intermediate regulators (first 3 columns, top 5 rows) overall follow a triangular pattern. These triangular patterns arise from the moderate reduction of swimming in strains that contain at least one perturbation of a global regulator which also have the largest growth effects. In contrast, the strains that are not perturbed by a global regulator show higher swimming performance but negligible differences in growth. Although the details of the epistasis depend on the precise orientation of these triangles, the overall patterns seem to arise from limits to growth or the growth-swimming trade-off.

Another ensemble of consistent epistatic interactions is between local (bottom 2 rows) and global (first 3 columns) regulators. Interactions within this ensemble yield large perturbations in both swimming and growth. These combinations therefore seem to be the least constrained by phenotypic limitations. The last ensemble is between local regulators (bottom 2 rows) and other local/intermediate regulators (last 4 columns). Here, genetic perturbations are dominated by strong decreases in swimming with little impact on growth. Unlike for epistasis between specific pairs of genes, these same overall patterns are observed across all 3 environments (SI).

To understand how the stereotypical patterns of genetic interactions interplay with phenotypic limitations, we re-examined the overall set of perturbations in the phenotypic space with edges representing individual perturbations (Figure 5). We coloured each edge by which of the three groups each perturbation belonged to, corresponding to the hierarchy of the regulatory network. These were global (yellow), intermediate (purple), or local (blue) regulators. In this representation, the overall structure of the epistatic effects is notably conserved across all environments (Figure 5A-C) even as particular epistasis for specific pair of KDs may not be (Figure 4 and SI). The vertical edges are mainly attributed to local (blue) or global (yellow) perturbations, which produce strong or moderate decreases in swimming fitness respectively. Horizontal and negatively slanted edges are dominated by global regulators, as would be expected from regulators linked to a phenotypic trade-off. Horizontal edges occur in moderate to low swimming phenotypes and result in significant changes to growth. In contrast, oblique edges include all the perturbations which move between the extrema of the Pareto front. Finally, intermediate regulator perturbations generally lead to fine-tuned changes in both phenotypes, independently of the background. This indicates a distinct role for each group of regulators.

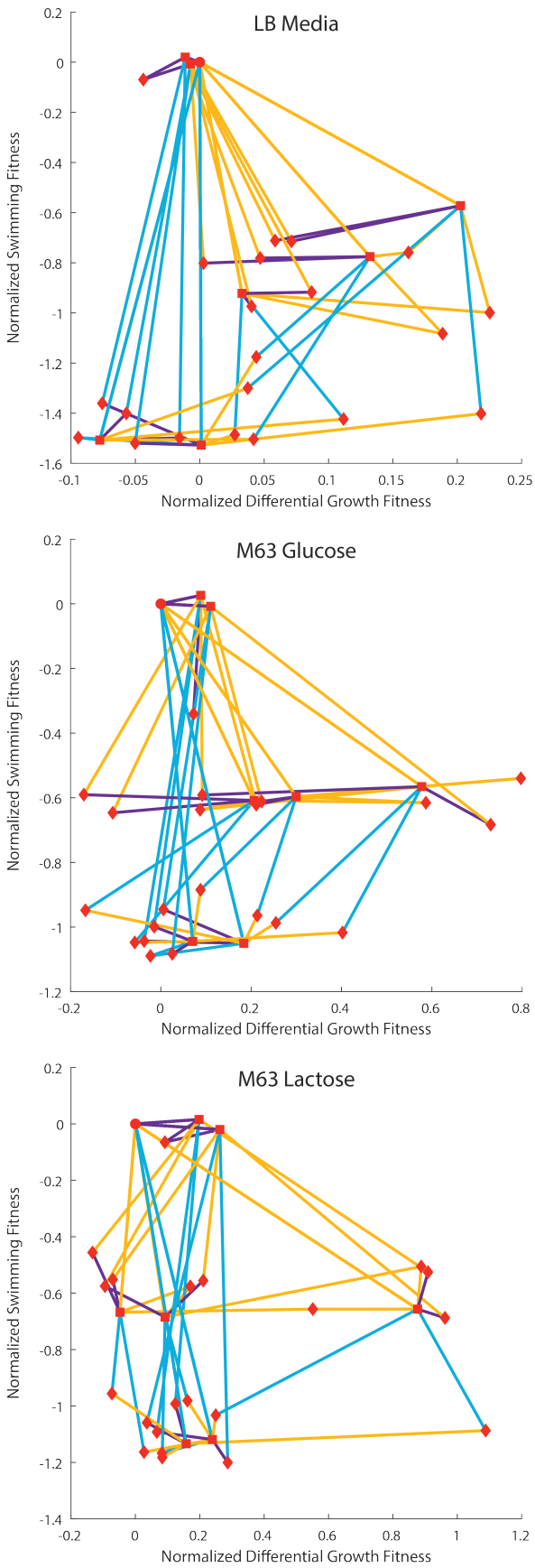

Figure 5: Relationship between regulatory network structure and response in phenotype space. The trajectories for each perturbation are coloured by the group of the transcriptional regulator: Global (yellow), Intermediate (Purple), or Local (blue). Reference strain is represented as a circle, single perturbations are represented as a box, and double perturbations are represented as a diamond. The constraints on phenotype space are conserved across all conditions with the same over-all pattern observed. Movements along the Pareto front are typically perturbations in the global regulators, while perturbations that move on or off of the Pareto front are typically local regulators.

Discussion

Interconnected networks and phenotypic trade-offs are ubiquitous in biological systems, from plants to cancer cells [39, 40]. The complexity of both of these are a major hurdle in our ability to explain and predict the evolution of biological systems [41]. In this work, we characterized how the interplay between global and local regulators determines two phenotypes, motility and growth, which are bound by a trade-off [6-9, 11]. We found that local regulators impact mostly the particular phenotype they are expected to control, swimming in this case, whereas global regulators are flexible in the way they couple phenotypic variations. Epistasis is pervasive and environment-dependent, confirming former large scale screens of combined genetic perturbations [42]. However, only a few types of epistatic patterns (primarily masking) were found when considering interactions between groups of genes. Genetic interactions for multiple phenotypes are in fact most easily explained when examining phenotypes altogether (rather than separately) and as a function of the position of genes in the regulatory hierarchy. Epistatic constraints and phenotypic limitations lead to evolutionary scenarios that require both local and global regulators for adaptation, the latest steps of which rely on adjusting the trade-off by global regulators.

These findings invite us to reconsider possible explanations for the origin of hierarchical structures in regulatory networks and the existence of global regulators. One interpretation for the existence of global regulation disregards functional details and suggests that the observed regulatory structures merely result from a probabilistic evolutionary process of network wiring [43]. For instance, rules of preferential attachment, where a highly connected gene would have more chance to become the regulator of a newly appeared gene (e.g. following gene duplication-divergence events) should lead to the existence of regulatory 'hubs' [44]. Such a framework could indeed explain the existence of global regulators, but it is unclear whether the probabilistic rules assumed in such scenario are mechanistically justified. Additionally, this explanation ignores the relationship between the network wiring process and organismal fitness.

A contrasting interpretation of regulatory hierarchy is purely functional. It suggests that hierarchy reflects cellular decision-making. For instance, alternative sugar operons, such as the lactose operon, have been shown to be triggered only in the absence of glucose when CRP is activated. This indicates a priority related to the quality of the carbon source and invokes optimality principles in metabolism [45]. But this interpretation also has its limitations. First, diauxie and the hierarchy in carbon source usage are not observed for all sugars [46, 47]. A possible explanation is that evolving a network able to discern every carbon source combination may be too costly as compared to fitness gains. Second, this reasoning may not necessarily apply to global regulators other than CRP, or to phenotypes other than growth. For instance, HNS is believed to control newly acquired genes from horizontal gene transfer, and shut them off by default [48]. It coordinates with another global regulator, FIS, to control the overall supercoiling and nucleoid density in opposing growth phases [49]. At the moment, it unknown whether their regulation corresponds to any form of priority in physiological requirements.

Conciliating the evolutionary and functional interpretations of global regulation is not a priori obvious. On the one hand, global regulators being connected to many genes is needed to re-balance expression at a genome-wide scale, thus seems necessary for adaptation. On the other hand, global regulators are highly pleiotropic, thus a source of evolutionary constraints (genetic perturbations may cause a large number of deleterious effects) [19].

The present study considers an additional ingredient, which contributes to resolve the above contraction: the existence of phenotypic trade-offs. Adapting to trade-offs requires the simultaneous and coordinated optimization of multiple phenotypes. Evolution then becomes dependent on genetic changes that couple variations in multiple phenotypes. In this context, accessible evolutionary trajectories require some degree of pleiotropy, as we observed when only global regulators can adjust two optimal phenotypes. Consistently, global transcription factors are frequently modified when cells undergo directed evolution for specific, often unrelated functions [50]. Pleiotropy may be a necessary evil to adapt to phenotypic trade-offs, thus illustrating how network complexity may find its roots in functional complexity [51].

Methods

Creation of vectors

pCRRNA mCherry and pCRRNA EGFP were made by Gibson Assembly. PCR was performed on pMD019 (pGFP) and pMD024 (pmCherry) with primers oMD546 and oMD547, and on pCRRNA with primers oMD545 and oMD548. PCR products were purified with PCR clean up kit from Macherey-Nagel. Equal molar concentrations of PCR product were mixed (one from either pMD019 or pMD024 and one from pCRRNA) and 5 µL of mix was added to 15 µL of Gibson Master mix to create pCRRNA Green and pCRRNA Red respectively. Golden Gate Assembly was used to replace the TrrnB terminator from pMD019 and pMD024 with B0014 as the TrrnB terminator contained 2 BsaI sites. PCR was done on pCRRNA Green and pCRRNA Red with primers oMD609 and oMD610, and on pCKDL with primers oMD607 and oMD608. PCR products where joined with Golden Gate Assembly using BsmBI enzyme to create pCRRNA EGFP and pCRRNA mCherry. pCRRNA mCerulean was created by Gibson Assembly. PCR was performed on pMD027

(pCerulean) with primers oMD704 and oMD705, and on pCRRNA mCherry with primers oMD706 and oMD707). PCR products were purified with PCR clean up kit from Macherey-Nagel. Equal molar concentrations of PCR product were mixed and 5 µL of mix was added to 15 µL of Gibson Master mix. All vectors were sequenced by GATC Biotech prior to use.

Growth Conditions of Cultures

The host strain for all pCRRNA vectors is MG1655 with the pdCas9 vector from Stanley Qi (provided by Lun Cui and David Bikard). Glycerol stocks of each culture were streaked onto individual Lysogeny Broth (LB) agar plates containing 34 µg/mL of Chloramphenicol and 50 µg/mL of Kanamycin. Single colonies were inoculated into 2 mL of selected media (either LB or M63 supplemented with 0.4% Glucose or Lactose) containing 34 µg/mL Chloramphenicol and 50 µg/mL Kanamycin. Cultures were placed in a 37°C incubator for either overnight for 16 hours for LB cultures or for 24 hours for M63 cultures.

Growth Competition Assays

Assays were performed by diluting 220 µL pCRRNA mCherry NT pre-culture into 22 mL of selected media containing 34 µg/mL Chloramphenicol, 50 µg/mL Kanamycin, and 25 ng/mL anhydrotetracycline. A Greiner, 96 Well, PS, F-Bottom, µCLEAR, Black microplate was filled with 198 µL of diluted culture per well. For each pCRRNA EGFP knockdown vector, 2 µL of pre-culture was inoculated into 3 individual wells. For the Non-targeting Control strain, 2 µL of pre-culture was inoculated into 15 individual wells. A volume of 60 µL of mineral oil was added to each well of the microplate. The absorbance at 595 nm, as well as fluorescence at 480/510 nm and 580/610 nm (excitation/emission) were recorded every 10 minutes for 20 hours with a SpectraMax i3x. Microplates were incubated at 37°C and shook for 90 seconds before and

after each measurement. The absolute growth rate of the reference strain was determined by the change in the log2(OD595nm) / time of the exponential growth phase (50mins to 170mins in LB, 50mins to 230mins in M63). To determine the fitness measurement of each knockdown, the ratio of the green and red fluorescence of the strain was divided by the median ratio of the green and red fluorescence of the 12 additional wells for the non-targeting control. Fitness measurements for each knockdown were made for LB media, and M63 media containing 0.4% of either glucose or lactose.

Swimming Fitness Assays

For each knockdown, 10 µL of pre-culture was spotted into the middle of 15 mL of selected media soft agar (0.3%) plates. Plates where then incubated at 37°C for either 16 hours for LB plates or 24 hours for M63 plates. Plates where then imaged using a USB Camera and python viewer. An image of a non-inoculated plate was subtracted from each image. The fitness was determined by the ratio of the swimming area diameter of each knockdown strain to that of the non-targeting control strain.

References

1. Shoval, O., et al., *Evolutionary Trade-Offs, Pareto Optimality, and the Geometry of Phenotype Space.* Science, 2012. **336**(6085): p. 1157-1160.
2. Gallaher, J.A., J.S. Brown, and A.R.A. Anderson, *The impact of proliferation-migration tradeoffs on phenotypic evolution in cancer.* Sci Rep, 2019. **9**(1): p. 2425.
3. Vittadello, S.T., et al., *Examining Go-or-Grow Using Fluorescent Cell-Cycle Indicators and Cell-Cycle-Inhibiting Drugs.* Biophys J, 2020. **118**(6): p. 1243-1247.
4. Garay, T., et al., *Cell migration or cytokinesis and proliferation?--revisiting the "go or grow" hypothesis in cancer cells in vitro.* Exp Cell Res, 2013. **319**(20): p. 3094-103.
5. Hatzikirou, H., et al., *'Go or grow': the key to the emergence of invasion in tumour progression?* Math Med Biol, 2012. **29**(1): p. 49-65.
6. Gude, S., et al., *Bacterial coexistence driven by motility and spatial competition.* Nature, 2020. **578**(7796): p. 588-592.
7. Narla, A.V., J. Cremer, and T. Hwa, *A traveling-wave solution for bacterial chemotaxis with growth.* Proceedings of the National Academy of Sciences, 2021. **118**(48): p. e2105138118.
8. Kuhn, T., et al., *Spatial scales of competition and a growth–motility trade-off interact to determine bacterial coexistence.* Royal Society Open Science, 2022. **9**(12): p. 211592.
9. Liu, W., et al., *The spatial organization of microbial communities during range expansion.* Current Opinion in Microbiology, 2021. **63**: p. 109-116.
10. Keegstra, J.M., F. Carrara, and R. Stocker, *The ecological roles of bacterial chemotaxis.* Nature Reviews Microbiology, 2022. **20**(8): p. 491-504.
11. Kim, J., et al., *Trade-offs between gene expression, growth and phenotypic diversity in microbial populations.* Current Opinion in Biotechnology, 2020. **62**: p. 29-37.
12. Baquero, F., et al., *The Origin of Niches and Species in the Bacterial World.* Frontiers in Microbiology, 2021. **12**.
13. Santos-Zavaleta, A., et al., *RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in E. coli K-12.* Nucleic Acids Research, 2019. **47**(D1): p. D212-D220.
14. Cremer, J., et al., *Chemotaxis as a navigation strategy to boost range expansion.* Nature, 2019. **575**(7784): p. 658-663.
15. Shen-Orr, S.S., et al., *Network motifs in the transcriptional regulation network of Escherichia coli.* Nature Genetics, 2002. **31**(1): p. 64-68.

16. Roy, S., et al., *Motifs enable communication efficiency and fault-tolerance in transcriptional networks.* Scientific Reports, 2020. **10**(1): p. 9628.
17. Nghe, P., M. Kogenaru, and S.J. Tans, *Sign epistasis caused by hierarchy within signalling cascades.* Nature Communications, 2018. **9**(1): p. 1451.
18. Kogenaru, M., et al., *Predicting Evolutionary Constraints by Identifying Conflicting Demands in Regulatory Networks.* Cell Systems, 2020. **10**(6): p. 526-534.e3.
19. Galis, F. and J.A. Metz, *Evolutionary novelties: the making and breaking of pleiotropic constraints.* Integrative and Comparative Biology, 2007. **47**(3): p. 409-419.
20. Hu, H., et al., *Constrained vertebrate evolution by pleiotropic genes.* Nature ecology & evolution, 2017. **1**(11): p. 1722-1730.
21. Fitzgerald, D.M., R.P. Bonocora, and J.T. Wade, *Comprehensive mapping of the Escherichia coli flagellar regulatory network.* PLoS genetics, 2014. **10**(10): p. e1004649-e1004649.
22. Pesavento, C., et al., *Inverse regulatory coordination of motility and curli-mediated adhesion in Escherichia coli.* Genes Dev, 2008. **22**(17): p. 2434-46.
23. Mika, F. and R. Hengge, *Small RNAs in the control of RpoS, CsgD, and biofilm architecture of Escherichia coli.* RNA biology, 2014. **11**(5): p. 494-507.
24. Andreassen, P.R., et al., *sRNA-dependent control of curli biosynthesis in Escherichia coli: McaS directs endonucleolytic cleavage of csgD mRNA.* Nucleic Acids Research, 2018. **46**(13): p. 6746-6760.
25. Botsford, J.L. and J.G. Harman, *Cyclic AMP in prokaryotes.* Microbiological reviews, 1992. **56**(1): p. 100-122.
26. Bradley, M.D., et al., *Effects of Fis on Escherichia coli gene expression during different growth stages.* Microbiology, 2007. **153**(9): p. 2922-2940.
27. Lucchini, S., et al., *H-NS mediates the silencing of laterally acquired genes in bacteria.* PLoS pathogens, 2006. **2**(8): p. e81.
28. Laurent-Winter, C., et al., *Role of Escherichia Coli Histone-Like Nucleoid-Structuring Protein in Bacterial Metabolism and Stress Response.* European Journal of Biochemistry, 1997. **244**(3): p. 767-773.
29. Dillon, S.C. and C.J. Dorman, *Bacterial nucleoid-associated proteins, nucleoid structure and gene expression.* Nature Reviews Microbiology, 2010. **8**(3): p. 185-195.
30. de Vos, M.G.J., et al., *Environmental Dependence of Genetic Constraint.* PLOS Genetics, 2013. **9**(6): p. e1003580.
31. Bikard, D., et al., *Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system.* Nucleic acids research, 2013. **41**(15): p. 7429-7437.
32. Adler, J., *Chemotaxis in Bacteria.* Science, 1966. **153**(3737): p. 708-716.
33. Lauffenburger, D., C.R. Kennedy, and R. Aris, *Traveling bands of chemotactic bacteria in the context of population growth.* Bulletin of Mathematical Biology, 1984. **46**(1): p. 19-40.
34. Fraebel, D.T., et al., *Environment determines evolutionary trajectory in a constrained phenotypic space.* eLife, 2017. **6**: p. e24669.
35. Yi, X. and A.M. Dean, *Phenotypic plasticity as an adaptation to a functional trade-off.* eLife, 2016. **5**: p. e19307.
36. Gjuvsland, A.B., et al., *Statistical Epistasis Is a Generic Feature of Gene Regulatory Networks.* Genetics, 2007. **175**(1): p. 411-420.
37. Macía, J., R.V. Solé, and S.F. Elena, *THE CAUSES OF EPISTASIS IN GENETIC NETWORKS.* Evolution, 2012. **66**(2): p. 586-596.
38. Domingo, J., P. Baeza-Centurion, and B. Lehner, *The Causes and Consequences of Genetic Interactions (Epistasis).* Annual Review of Genomics and Human Genetics, 2019. **20**(1): p. 433-460.

39. Hausser, J. and U. Alon, *Tumour heterogeneity and the evolutionary trade-offs of cancer.* Nature Reviews Cancer, 2020. **20**(4): p. 247-257.
40. Messier, J., et al., *Interspecific integration of trait dimensions at local scales: the plant phenotype as an integrated network.* Journal of Ecology, 2017. **105**(6): p. 1775-1790.
41. Hall, A.E., et al., *Environment changes epistasis to alter trade-offs along alternative evolutionary paths.* Evolution, 2019. **73**(10): p. 2094-2105.
42. Costanzo, M., et al., *Environmental robustness of the global yeast genetic interaction network.* Science, 2021. **372**(6542): p. eabf8424.
43. Boguñá, M., et al., *Network geometry.* Nature Reviews Physics, 2021. **3**(2): p. 114-135.
44. Torres-Sosa, C., S. Huang, and M. Aldana, *Criticality Is an Emergent Property of Genetic Networks that Exhibit Evolvability.* PLOS Computational Biology, 2012. **8**(9): p. e1002669.
45. Peebo, K., et al., *Proteome reallocation in Escherichia coli with increasing specific growth rate.* Molecular BioSystems, 2015. **11**(4): p. 1184-1193.
46. Aidelberg, G., et al., *Hierarchy of non-glucose sugars in Escherichia coli.* BMC systems biology, 2014. **8**(1): p. 1-12.
47. Wang, X., et al., *Growth strategy of microbes on mixed carbon sources.* Nature communications, 2019. **10**(1): p. 1279.
48. Dorman, C.J., *H-NS, the genome sentinel.* Nature Reviews Microbiology, 2007. **5**(2): p. 157-161.
49. Kahramanoglou, C., et al., *Direct and indirect effects of H-NS and Fis on global gene expression control in Escherichia coli.* Nucleic acids research, 2011. **39**(6): p. 2073-2091.
50. Liu, R., et al., *Engineering regulatory networks for complex phenotypes in E. coli.* Nature Communications, 2020. **11**(1): p. 4050.
51. Gros, P.-A., H. Le Nagard, and O. Tenaillon, *The evolution of epistasis and its links with genetic robustness, complexity and drift in a phenotypic model of adaptation.* Genetics, 2009. **182**(1): p. 277-293.