bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

What influences selection of native phosphorelay architectures?

Rui Alves [1,*], Baldiri Salvado [1,*], Ron Milo [2], Ester Vilaprinyo [1], Albert Sorribas [1]

*Equal contributors and co-corresponding authors

1 – Departament de Ciencies Mèdiques Bàsiques, Universitat de Lleida, Edifici de Recerca Biomèdica II, Av. Rovira Roure 80, 25198, Lleida, Spain

2 – Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehoovot, Israel.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract: Phosphorelays are signal transduction circuits that combine four different phosphorylatable protein domains for sensing environmental changes and use that information to adjust cellular metabolism to the new conditions in the milieu. Five alternative circuit architectures account for more than 99% of all phosphorelay operons annotated in over 9000 fully sequenced genomes, with one of those architectures accounting for more than 72% of all cases.

Here we asked if there are biological design principles that explain the selection of preferred phosphorelay architectures in nature and what might those principles be. We created several types of data-driven mathematical models for the alternative phosphorelay architectures, exploring the dynamic behavior of the circuits in concentration and parameter space, both analytically and through over 10^8 numerical simulations. We compared the behavior of architectures with respect to signal amplification, speed and robustness of the response, noise in the response, and transmission of environmental information to the cell.

Clustering analysis of massive Monte Carlo simulations suggests that either information transmission or metabolic cost could be important in selecting the architecture of the phosphorelay. A more detailed study using models of kinetically well characterized phosphorelays (Spo0 of *Bacillus subtilis* and Sln1-Ypd1-Ssk1-Skn7 of *Saccharomyces cerevisiae*) shows that information transmission is maximized by the natural architecture of the phosphorelay. In view of this we analyze seventeen additional phosphorelays, for which protein abundance is available but kinetic parameters are not. The architectures of 16 of these are also consistent with maximization of information transmission.

Our results highlight the complexity of the genotype (architecture, parameter values, and protein abundance) to phenotype (physiological output of the circuit) mapping in phosphorelays. The results also suggest that maximizing information transmission through the circuit is important in the selection of natural circuit genotypes.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Background

Organisms and cells use signal transduction circuits to detect environmental changes and make decisions on how to adjust their internal milieu to better survive those changes. Those circuits modulate the cellular response and its metabolic adjustments.

Phosphorelays (PR) are a large and important class of signal transduction circuits in microorganisms and some plants[1–4]. A self-phosphorylating Sensor Kinase (SK) modulates its own phosphorylation state in response to the environmental signal. The phosphorylated SK transfers its phosphate to an aspartate residue in a Response Regulator (RR) protein. This response regulator transfers its phosphate to a Histidine phosphotransfer (Hpt) protein domain, which then transfers the phosphate to a final RR (Figure 1). The phosphorylation state of the circuit modulates cellular activity and adaptation. PR are important for making life or death decisions about sporulation [5], in adapting to various stressors [6], such as changing levels of oxygen [7], or in developmental decisions made by many plants [8].

The four-step nature of the PR phosphotransfer cascade creates steeper signal response curves than those of two component systems (TCS), a simpler signal transduction alternative to PR with only two phosphotransfer steps [9–15]. This difference could explain the evolution of PR as an alternative to TCS signaling.

Still, PR circuits can assemble the four-step phosphotransfer cascade using alternative protein domain assembly architectures [12–14,16–18]. This raises the question of whether alternative PR architectures may have irreducible physiological differences for the behavior of the circuit or if those differences may be quenched easily by evolving appropriate parameter values [9,10,19–21]. In this context, can we identify general physiological requirements that dominate the selection of PR circuit architectures and/or parameters?

To address this question we used a comprehensive census of PR circuit architectures, and analyzed over 9000 fully sequenced microbial and plant genomes. We found that more than 72% of all PR coded in a single operon have the same circuit architecture (M3 of Figure 1) [16]. Four additional architectures account for another 27% of all other PR (architectures M1, M2, M2', an M4 of Figure 1) [16].

To understand the physiological differences between the five PR architectures we consider a set of physiological variables that are important for the function of PR. We analyze how adjusting those variables for improved circuit performance depends on the architecture, the protein abundance, and the parameter values of the PR. Our results suggest that optimizing different aspects of environmental information transmission to the cell could be crucial for selection of a specific PR architecture. They also suggest that fine tuning of parameter values and protein abundances could, in many cases, overcome intrinsic differences between the physiological responses of alternative circuit architectures. Together, the results reported here have consequences for understanding the evolution of PR signal transduction circuits and for the development of synthetic biology PR circuit applications.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Results

Physiological variables as a proxy for circuit performance in signal transduction

The overall fitness of organisms is affected by how their molecular components organize into biological circuits [22,23]. The organization of molecular circuits with common biological functions in different organisms may have alternative architectures. Often, these alternatives lead to improved circuit performance in the context of the organisms where they are observed. Understanding how the alternative architectures affect circuit performance requires that the physiological variables that are important for the biological function of the circuit are understood. Over the last few decades, several of these physiological variables were identified as important in determining the performance of molecular signal transduction circuits. We compiled those variable and summarize them in Table 1.

Comparing physiological variables between architectures

To investigate if any physiological variable or combination thereof might explain the genomic frequency differences between the five alternative PR architectures shown in Figure 1, we started by creating mathematical models that permit estimating the variables from Table 1 for each of the architectures. Then, we analyzed those models and compared the variables between architectures in a controlled manner.

We note that the comparisons consider variables that are physiologically relevant in two domains for the dynamic response of the circuits (Table 1): (i) the deterministic domain (≥~1000 copies of the circuit per cell); (ii) the stochastic domain (<1000 copies of the circuit per cell). The methods used to calculate the different physiological variables depend on the regime being analyzed.

In the deterministic domain, signal amplification, robustness, speed, and metabolic cost can be calculated analytically for each architecture (see Datafile_S1.zip for the analytical expressions). We compare which architecture has the highest value for each of these variables by taking the ratio of that variable between all possible pairs of architectures. A detailed analysis of the analytical ratios shows that the best performing architecture with respect to any variable depends on the parameter values.

To investigate if, statistically, there were physiological differences between architecture with respect to the different physiological variables we used a Monte Carlo approach. We randomly generated sets of parameter values, tested the system they created and selected those that create realistic behavior for each architecture and permit comparing that behavior between architectures in a controlled manner (see methods for details). We stopped generating new sets when five thousand realistic systems were obtained for each architecture. Overall, we scanned for more than one hundred and sixty million parameter sets.

Then, for each variable and set of parameter values, we calculated the numerical value of the physiological variables. Finally, for each variable, we selected the 10% of systems that performed better and calculated the percentage of each architecture contained in this high-performance set.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The percentages of systems for each architecture in the best performing systems are very different from the relative genomic frequency of the architectures. Because of this we ranked the five architectures from best (1) to worst (5). Results are summarized in Table 2.

We then clustered the vectors of architecture ranks (Figure 2) using a Canberra distance metric. We find that Information transmission and cost are the physiological variable that generate a rank that is the most similar to that generated by genomic frequency.

Controlling for the possibility of bias in parameter space: data-driven comparative analysis of the five architectures

The previous results suggest that either cost of operating the circuit or information transmission or both could justify that architecture M3 is the most frequently found in genomic PRs. Still, given the percentage differences in architectural abundance between high performance and genomic PRs, it could be that we are analyzing parameter regions and protein abundances that are unexplored by nature, biasing our results.

To control for this possibility, we searched the literature for experimentally determined parameter values for the individual reactions of PR. A similar search was done for the total amount of the individual proteins in the circuit. Then, using a Latin hypercube approach (see methods) we created parameter values sets combining those values. Overall, five hundred combinations of parameters and concentrations were analyzed, percentages of abundance for each architecture in the set of systems that have the top 10% best performance were calculated as described above and clusters were built. The results are similar to those described before (Figure 2).

In order to clarify if cost or information transmission through the circuit are important factors in selecting PR architecture we performed two additional data driven sets of experiments.

What physiological variables are optimized by native architectures in kinetically well characterized PRs?

First, we focus on two systems for which abundant quantitative and kinetic information is available: The Spo0 PR in *Bacillus subtilis* and the Sln1 system in *Saccharomyces cerevisiae*.

The Spo0 phosphorelay of *Bacillus subtilis*

We first focused on the Spo0 phosphorelay in *B. subtilis* because this is the PR system with the most abundant and reliable experimental determinations of kinetic information [25–29]. The architecture of this system is of type M1, and supplementary Appendix S1 gives the protein abundances and parameter values for the reactions. We then created four alternative PR architectures (M2 - M4) with the same parameter values and protein abundances adjusted to minimize differences in dynamic behavior with respect to the native circuit architecture. Then, we compared the architectures with respect to the variables in Table 1. We found that the cost of protein synthesis is an order of magnitude lower for architecture M3 than for M1 (Figure 3). Protein costs for architectures M1, M2', and M4 are an order of magnitude higher than in M3 and one order of magnitude lower than in M2. The robustness of signal transmission to environmental fluctuations in parameters is smaller in architectures M1 and M2' and larger in M2, M3, and M4 (Figure 3). Architecture M4 is the fastest in responding to phosphorylating signals, and M4 and M1 are faster in a similar percentage of cases when the signal dephosphorylates the system (Figure 3). We clearly see that M1 transmits the highest amounts

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of information about the environment over the regulatable signal range for the Spo0 phosphorelay, whether the signal comes at the level of the SK (modulation of k1) or at the level of regulating dephosphorylation of the final response regulator (modulation of k26) (Figure 4). We note that the amount of information lost through the channel is smaller for architecture M2.

The Sln1/Ypd1/Ssk1 and Sln1/Ypd1/Skn7 phosphorelays of *Saccharomyces cerevisiae*

S. cerevisiae uses a PR with architecture M2 to sense changes in the osmolarity of the medium and regulate its internal metabolism and membrane composition, adapting to those changes. The hybrid sensor kinase Sln1 senses changes in membrane curvature and ultimately modulates the phosphorylation state of the terminal response regulators Ssk1 and Skn7. Phosphorylation of Ssk1 acts as a molecular switch in controlling the yeast's osmosensing mitogen-activated protein (MAP) kinase cascade[30], while phosphorylation of Skn7 directly affects the transcription of SLN1-SKN7-responsive genes [31]. Skn7 is also involved in regulating heat shock response in a Sln1-Ypd1 independent way [31]. Reliable kinetic parameter values and protein abundances are available for this PR.

We use those values to create a PR model where we consider the simultaneous presence of the terminal response regulators Ssk1 and Skn7. Then, create four independent alternative models, with architectures M1, M2', M3, and M4 and parameter values equal to those of the native circuit architecture. Finally, we optimize protein abundance to minimize differences in the dynamic signal-response curve of each architecture with respect to that of the native architecture of the system.

We find that the cost of architecture M2' is one order of magnitude higher than that of each of the other architectures, which have a similar metabolic cost (Figure 3). The sensitivity of the steady state concentrations for the Ssk1 RR are low and similar in all alternative architectures except M4, where sensitivity is high. In contrast, the sensitivity of Skn7 concentration is highest in architectures M1 and M2 (Figure 3). Architecture M4 most frequently responds faster to signals that increase the phosphorylation levels of the proteins, followed by M2 (Figure 3). In contrast, architecture M1 is always the fastest if the change in environmental conditions decreases phosphorylation levels of the circuit. When it comes to transmitting information about the environment to Skk1 over the regulatable range of the circuit, architecture M2 is the best for a wider range of parameters. It surprised us that all non-native architectures are better information transmitters to Skn7 than M2 (Figure 4). It is interesting to note that only very high rates of dephosphorylation for one RR affect the information transmitted by the circuit to the other, increasing it significantly.

Overall, these two systems suggest that transmission of information is an important determinant of architectural selection in PRs.

Analyzing real PR examples: How do different PR architectures influence information transmission?

A performance atlas for information transmission in the parameter space of PR architectures

With the hypothesis that information transmission is important in the selection of PR architecture, we created a data-driven atlas that describes how architecture, protein

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

abundance, and parameter values influence information transmission in PRs. To create that atlas we compiled experimentally determined parameter values and proteins abundances for as many PR systems as we could find in the primary literature and in public databases (see methods). Then we used those parameters and protein abundances to calculate the capacity to transmit information of each alternative architecture with all possible combinations of protein abundances and kinetic values. Finally we ranked the five architectures with respect to increasing capacity of information transmission. The atlas is summarized in Supplementary Table S2.

Experimentally determined protein abundances are consistent with amount of transmitted information being an important driver of PR architecture selection

We identified and obtained protein abundances for seventeen PR systems that are present in the PAXDB database and belong to seven different organisms[24] (Table S3). Then we matched these abundances with the performance in the atlas of amount of transmitted information as a function of parameter values, protein abundance and circuit architecture given in Table S2. In 16 out of the 17 cases we find that the native architecture of the PR systems is consistent with maximizing the through-circuit transmitted information under experimentally known parameter conditions (Table 3).

Discussion

Implications

Mutations in the genomes of organisms led to the emergence of variant architectures for biological circuits with similar biological function. Natural selection acting upon this variability led to fixing those architectures that are more efficient for the function of the circuit in the genome, contingent on life history. Elucidating the reasons why these architectures improve performance of the circuit reveals biological design principles that constrain evolution and explain the observed patterns.

In some cases, biological design principles are general and apply to a whole class of biological circuits. For example, overall feedback inhibition of the final product to the initial reaction of a linear biosynthetic pathway leads to molecular systems that are faster, more responsive, and less sensitive to fluctuations than any other regulatory alternatives [32–34]. As such that regulatory solution is widespread in the tree of life. However, in many cases, design principles appear casuistic and are specific to a single circuit. For example, the circuit that regulates competence in *B. subtilis* appears to have been selected over alternative designs because of its noisy response [35], which enables bet edging in the competence response of the bacterium and improves its survival chances.

In the case of PR, a specific architecture is observed in over 72% of all genomic PR [16] in a sample of more than 9000 organisms [4]. This suggested that indeed there could be design principles for selecting the architecture of these circuits. To test this hypothesis we needed to understand what physiological variables influence the performance of the circuit. Then, we needed to study how architecture affects those variables, allowing natural selection between alternative PR architectures.

Initially we compiled a list of physiological variables that are important for the function of these signal transduction circuits [9,10,42–47,12,27,36–41]. We compared the alternative architectures to understand how they differed with respect to each of those variables. We then ranked the

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

alternative designs with respect to their performance for each variable. Our results show that no unique variable could be used to rank all the alternative PR architectures in the same order as their genome abundance. Still, an analysis of large scale "simulomics" results shows that maximizing information transmission through the circuit and minimizing metabolic cost leads to an abundance of PR architectures that is the closest to the observed genomic frequency of those architectures.

This result led us to hypothesize that optimization of information transmission is an important driver of selection for PR architectures. To investigate this hypothesis we took a three step approach. First, we focused on two PR that are experimentally well characterized in terms of kinetic parameters and protein abundances. Simulation experiments revealed that the native architecture was the one that maximizes information transmission by the circuit. Second, we created a data-driven atlas of how architecture, protein abundance, and parameter values influence information transmission capacity. To create that atlas we compiled experimentally determined parameter values and proteins abundances for as many PR systems as we could find in the primary literature and in public databases. Then we used those parameters and protein abundances to calculate the capacity to transmit information of each alternative architecture with all possible combinations of protein abundances and kinetic values. Third, we took all PR for which experimental protein abundances are known and verified that the native architecture of the system is the one that maximizes information transmission in sixteen out of seventeen cases.

Overall, the results of our analysis for well characterized PR circuits are consistent with the hypothesis that information transmission is a crucial driver of PR circuit selection in nature. This was previously suggested to be the case for MAP kinase signaling and other metabolic circuits in eukaryotes [48–55]. Still, architecture selection is a complex function of the interplay between the architecture, protein amounts and kinetic parameters, and evolution can play with these "genotypic variables" to find instances of each architectures that are almost equivalent information transmission circuits. Although we did look for additional specific examples of PR, we could not find any for which a complete set of parameter values had been measured either by the same or by independent labs. For some of them there were partial sets of parameter values (The ArcA – ArcB, the EvgS – EvgA, and the BvgS – BVgA PR). Nevertheless, creating models for these systems would have required using more than half of the parameters from other systems. This would have created a situation where the model would not be driven from the data, but rather from other estimates.

Our results provide evidence about design principles that could explain the evolutionary emergence of specific PR architectures. Those results also have consequences for synthetic biology. If we want to build synthetic PR that are stable in the genome, and have a significant level of quantitative understanding about the operating ranges for the circuit, then it should be designed using the architecture that has the highest capacity to transmit information over that range.

Speculations

Our analysis invites some speculation regarding the ranges of operability for many PR. If 72% of all PR have a M3 architecture, as a first hypothesis we could expect that these systems operate in the ranges where the M3 architecture outperforms the others. By looking at our performance in information transmission atlas, this implies that M3 architecture systems are likely to be operating at abundance ranges of the PR that are below 100 molecules per cell for the protein containing the SK domain. This is amenable to testing in future proteomic experiments. The M3

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

circuits are also expected to be operating over modulatable ranges of the self-phosphorylation and self-dephosphorylation rate constants of SK above 10^{-1} min^{-1}. Similarly, 18% of the identified PR might be operating mostly over the range where architecture M2' transmits a high amount of information about the system, and 6% over the range where M4 is a better information transmitter.

In addition, we speculate that the type of environmental stimulus is also important in selecting the architecture. If the stimulus changes in a graded way and cells can be adjusted in a similar graded way, it makes sense that the architecture should allow for a high capacity transmission. For example, an architecture that allows the cell to distinguish between n+1 states (that is, with a capacity to transmit information of n+1 bits) provides for a better design that another architecture that only distinguishes between n states. On the other hand, if a very sharp response is required, architectures that can distinguish between a small number of states over a short operability range might be more effective.

Another tempting speculation arises from our modelling of the Sln1 phosphorelay. Given its information transmission profile, it could be that Sln1-mediated activation of Skn7 and its dependent genes may not be an important function of the circuit. For the experimentally determined parameter values and protein concentrations, other architectures would transmit more information to that RR. This suggests to us that Skn7's role in heat shock response might be much more important for the cell than its role in osmoregulation.

Limitations

Our study has several limitations. Here we discuss those we think are the most important. First, we may have underestimated the percentage of PR with architecture M1, as these PR may be coded under different promoters in separate parts of the genome. This could also be true for designs M2, M2', and M3. Nevertheless, the amount of proteins of type SKRR, SKHpt, or SKRRHpt that are found without the remaining operon in genomes is much smaller than that of the same type of proteins that are assembled in operons with the remaining putative cognate proteins of the PR [4], which indicates that any underestimation of these architectures is likely to be small.

Second, the performance landscape of PR can be strongly affected by how its expression is regulated [45,56–58]. Nevertheless, transcriptional regulation occurs on a timescale of tens to hundreds of minutes, while the regulation of phosphorylation levels of the PR occurs on a timescale of minutes and this timescale can also strongly influence the performance of the gene expression regulatory circuit [57,59].

Third, there might be a wider range of parameter values for each reaction of the PR in the wild and this could change our phenotypical mapping of PR behavior onto the phase and architectural spaces of the circuits. Nevertheless, if this is so, our results would still be valid for the regions that were analyzed and considering that expanded set of values would only increase the "genotypic space" without changing the mapping we present here.

Fourth, our analysis of the capacity to transmit information focuses on steady state shifts. One could argue that the transient capacity could be a more important determinant of system performance. However, we tested how the transient capacity to transmit information compares to the corresponding steady state capacity and found that the latter is always a majorant of the former.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Methods

Mathematical Modelling

PR shown in Figure 1 represent more than 99% of all PR circuits present in the fully sequenced genomes of more than 9000 organisms. This information was obtained from [4]. Two types of models were created for each architecture.

S-system and GMA models

We used the power law formalism for the large-scale mapping of dynamic behavior into parameter and concentration space. This formalism can be used to create quantitative models of circuits when little kinetic and/or mechanistic information is available [60].

First, we created models for the five architectures using the S-system representation. This representation allows us to calculate analytical expressions for the steady state values of each of the deterministic physiological variables that we needed to evaluate. Then we used the generalized mass action representation to create models that could be evaluated numerically for each of the five architectures. These models were used to map the physiological behavior of each architecture in parameter space. All models are given in supplementary Appendix S1.

Mechanistic Models

We used a mass action description of reactions to create a mechanistically more detailed model for each architecture [60]. These models are given in supplementary Appendix S1. They were parameterized with experimentally determined parameter values and protein concentrations compiled from the primary literature. We used these models to create an atlas of performance with information transmission through the PR as a function of protein amount, parameter values, and circuit architecture.

Estimating protein concentrations

Proteins occupy approximately 20% of cell volume across the tree of life [61,62]. Taking into account average protein sizes, average protein masses and factoring in cell volumes one can estimate that total protein concentration in cells is of the order of ~ 1 μM (ranging between 0.4-1.4) [63], or 6.23×10^6 proteins per μm^3. We then use the average cell volume for the different cell types to estimate the total amount of proteins per cell.

In order to estimate biologically relevant protein amounts for each circuit topology we used whole-proteome relative abundance determinations (Table S3) reported by PaxDB [24]. The database information is given in parts per million. By multiplying these number by total number of cell proteins, we calculate how many proteins exist for each experimental PR system, within the experimental error.

With this information we can further estimate typical orders of magnitude for the ratios between protein abundances within a PR circuit. In general, we find that protein abundances and ratios between protein abundances are bound and have a limited range in PR (Table S3).

Parameter values for the mass action models

To find experimentally determined parameter values for the individual reactions of the PR, we searched Medline and the Biomodels database [64]. We searched for previous mathematical models for TCS and signal transduction PR. In addition we also search the primary literature for

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

known circuits that have been characterized biochemically and collected all different parameter values for the various reactions in the PR circuit. This information is compiled in Supplementary Table S4, revealing that kinetic parameters for corresponding reactions in experimentally well-characterized systems are quite similar and, in most cases, within the same order of magnitude.

Calculating the dynamic behavior of each model

Steady state concentrations and stability

We obtained steady state concentrations by numerically solving the GMA and mass action models for each set of parameter values. We then calculated the jacobian of the ODE system and the eigenvalues of that jacobian to determine the stability of the steady state. We compared the steady state stability of the PR circuits in two ways. As we selected for sets of parameter values that generate systems with stable steady states (see methods), all real parts of the eigenvalues of the steady state are negative. Taking this into account we compared the minimum of the real parts of the eigenvalues in each pair of models. This comparison allows us to compare the fastest time scale in which the two circuits respond to a transient perturbation to their steady states. We also compared the maximum of the real parts of the eigenvalues between the two circuits in each pair. This allows us to compare the slowest time scale in which the two circuits respond to a transient perturbation to their steady states.

Logarithmic gains and parameter sensitivities

To estimate the response of a physiological variable to changes in environmental signals or model parameters, we compute [60]

$$L(X,\theta) = \frac{dLog(X)}{dLog(\theta)} = \left(\frac{\theta}{X} \times \frac{dX}{d\theta}\right)_{SS}$$

Here, ss indicates evaluation at a reference steady state.

Signal amplification

To estimate signal amplification we calculated the logarithmic gains of the concentration of the phosphorylated form of the final response regulator (RR_2) with respect to the cognate signal of the system (Signal), as given by:

$$L\left(RR2_p, Signal\right) = \frac{dLog(RR2_P)}{dLog(Signal)} = \frac{Signal}{RR2_P}\frac{d\ RR2_P}{d\ Signal}$$

Sensitivity to the total amount of protein

To estimate sensitivity of the different phosphorylated forms of the PR proteins (Prot-P) with respect to the total amount of circuit proteins (Pr_{tot}), we calculated the logarithmic sensitivities, as given by:

$$L(Prot-P, Signal, Pr_{tot}) = \frac{dLog(Prot-P\ \ )}{dLog(Pr_{tot})} = \frac{Pr_{tot}}{Prot-P}\frac{d\ Prot-P}{d\ Pr_{tot}}$$

Global sensitivity to fluctuations in parameters

To estimate sensitivity of the different phosphorylated forms of the PR proteins (Prot-P) with respect to fluctuations in each of the parameters (Par_i), we calculated the logarithmic sensitivities, as given by:

$$L(Prot-P, Signal, Par_i) = \frac{dLog(Prot-P\ \ )}{dLog(Par_i)} = \frac{Par_i}{Prot-P}\frac{d\ Prot-P}{d\ Par_i}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Then, for each protein, we calculate the norm of the vector whose coordinates are the individual sensitivities of that protein, as described in [60]. This is an aggregated measure of the robustness of circuits to fluctuations in parameter values [47,65].

Metabolic cost of the circuit

To calculate the metabolic cost of running a circuit we focused on two aspects. First we estimated the average size of each type of protein in each architecture from the data in [4]. Then, we multiplied the number of amino acids of the individual proteins in the phosphorelay by the number of proteins in the cell and used this number as a proxy for the cost of each circuit. Then, we also focused on the reactions of the circuit that consume ATP (self-phosphorylation of the SK domain).

Response times

We performed two independent experiments per architecture and per set of parameter values to calculate the response times. First, we start with the system fully dephosphorylated and run a time course simulation using the same scan procedure for the self-phosphorylation and self-dephosphorylation rate constant of SK. We then calculate the time it takes each of the architectures to reach 90% of the new steady state values. We repeated this experiment starting with fully phosphorylated proteins.

Steady state noise: the coefficient of variation

To estimate the intrinsic noise for each architecture and set of parameter values we numerically calculated the deterministic concentration of the systems, converted this concentration into number of copies of the PR, and ran twenty five stochastic simulations of twenty minutes. Then, for each variable we calculated the coefficient of variation (CV) as given by:

$$CV_{Variable\ i} = \frac{Standard\ Deviation\ (Variable\ i)}{Mean\ value\ (Variable\ i)}$$

Amount of information transmitted through the circuit

To calculate the information transmitted through the PR we simulated that system to steady state using Gillespie's algorithm for stochastic simulation one hundred times per set of parameter values [66]. The proxy for changes in the environmental information was taken to be the number of phosphorylated molecules of the SK domain, as this is the sensing domain of the PR. The output of the PR was considered to be the number of phosphorylated final response regulators. The information transmission through the circuit was estimated using the standard formula for mutual information between the two variables:

$$MI(SKP \rightarrow RR2P) = \sum_{SKP_i} \sum_{RR2P_j} p(SKPi, RR2Pj) \log\left(\frac{p(SKPi, RR2Pj)}{p(SKPi)p(RR2Pj)}\right)$$

The information transmission from the SK to the first RR was estimated from:

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

$$MI(SKP \rightarrow RR1P) = \sum_{SKP_i} \sum_{RR1P_j} p(SKP_i, RR1P_j) \log\left(\frac{p(SKP_i, RR1P_j)}{p(SKP_i)p(RR1P_j)}\right)$$

The information transmission from the Hpt domain to the final RR was estimated from:

$$MI(HptP \rightarrow RR2P) = \sum_{HptP_i} \sum_{RR2P_j} p(HptP_i, RR2P_j) \log\left(\frac{p(HptP_i, RR2P_j)}{p(HptP_i)p(RR2P_j)}\right)$$

In these formulas, p(x) is the relative frequency of x, and p(x,y) is the joint relative frequency of x and y.

We compare the information transmitted through the various steps of the circuit as a measure of information loss. The closer $MI(SKP \rightarrow RR1P), MI(RR1P \rightarrow HptP),$ and $MI(HptP \rightarrow RR2P)$ are, the smaller the amount of lost information.

Comparison of dynamic properties between architectures

The comparison between alternative architectures was done in a pairwise manner, to facilitate interpretation.

Analytical comparison

We calculated the analytical solutions for the S-system model of each architecture. Then we took the ratio of corresponding properties between all possible pairs of architectures, controlling the comparison in such a way that the phosphorylated levels of the final RR were the same in both elements of the pair. If the ratios could be demonstrated to be always larger or smaller than one, this would mean that the property being analyzed would always be large in one of the architectures.

Numerical comparison

We took a reference architecture and generate a set of parameter values drawing from a random distribution of parameters with realistic boundaries. We did the same sampling for protein concentrations. Then we calculated the numerical properties of the pair of models being compared, discarding all sets of parameter values that led to systems with negative signal amplifications, unstable steady states, aggregated sensitivities to fluctuations in parameters and sensitivities of phosphorylated forms to total amounts of circuit protein that were too high, and response times that were too slow. We stopped generating sets of parameter values when 5000 of those set met all criteria of realist performance described above. Overall, we tested several hundred million sets of parameter values in all comparisons.

Mathematical modelling of the Spo0 and Sln1 phosphorelays

Detailed mechanistic models were created for the Spo0 and Sln1 phosphorelays, as described in supplementary Appendix S1. These models use experimental data from [29] for Spo0 and from the SGD [67] for Sln1.

Then, for each of the PR, four additional equivalent models were created, each assuming that the PR would have a different architecture. The parameter values for these alternative architectures were considered to be the same as those for the native circuit architecture. The concentrations for the proteins were optimized by first simulating the original model at different

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

signal intensities and calculating the steady state phosphorylation state of the final RR. Using these results we then performed the same experiments for each alternative architecture and allowed for the proteins that got fused or separated with respect to the cognate architecture to change in concentration by up to one order of magnitude below or above that in the natural architecture. We then selected the concentrations that led to the most similar signal-response curves using a least square minimum criteria for the differences between phosphorylated final RR.

Calculating the cost of alternative architectures for the Spo0 and Sln1 phosphorelays

To calculate the cost of alternative architectures for the Spo0 and Sln1 we first counted the number of amino acids in the sequence for each protein in the original architecture. Then, we assumed that any alternative architecture would result from fusing or splitting the original proteins while conserving the same number of amino acids. Finally, we multiplied the number of amino acids of the individual proteins in the phosphorelay by the number of proteins in the cell and used this number as a proxy for the cost of each architecture.

Software

All models and analysis were done using Mathematica [68]. The notebooks containing all code to generate each figure are given as supplementary data pack S1.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Bibliography

1. Ulrich, L. E. & Zhulin, I. B. The MiST2 database: a comprehensive genomics resource on microbial signal transduction. *Nucleic Acids Res.* **38**, D401-7 (2010).

2. Porter, S. L., Wadhams, G. H. & Armitage, J. P. Signal processing in complex chemotaxis pathways. *Nat. Rev. Microbiol.* **9**, 153–65 (2011).

3. Kobir, A. *et al.* Protein phosphorylation in bacterial signal transduction. *Biochim. Biophys. Acta* **1810**, 989–94 (2011).

4. Salvado, B., Vilaprinyo, E., Sorribas, A. & Alves, R. A survey of HK, HPt, and RR domains and their organization in two-component systems and phosphorelay proteins of organisms with fully sequenced genomes. *PeerJ* **2015**, (2015).

5. Burbulys, D., Trach, K. A. & Hoch, J. A. Initiation of sporulation in B. subtilis is controlled by a multicomponent phosphorelay. *Cell* **64**, 545–52 (1991).

6. Bekker, M., Teixeira de Mattos, M. J. & Hellingwerf, K. J. The role of two-component regulation systems in the physiology of the bacterial cell. *Sci. Prog.* **89**, 213–42 (2006).

7. Georgellis, D., Kwon, O. & Lin, E. C. Quinones as the Redox Signal for the Arc Two-Component System of Bacteria. *Science (80-.).* **292**, 2314–2316 (2001).

8. Kieber, J. J. & Schaller, G. E. Cytokinin signaling in plant development. *Development* **145**, dev149344 (2018).

9. Igoshin, O. A., Alves, R. & Savageau, M. A. Hysteretic and graded responses in bacterial two-component signal transduction. *Mol. Microbiol.* **68**, 1196–1215 (2008).

10. Alves, R. & Savageau, M. A. Comparative analysis of prototype two-component systems with either bifunctional or monofunctional sensors: differences in molecular structure and physiological function. *Mol. Microbiol.* **48**, 25–51 (2003).

11. Salvado, B., Vilaprinyo, E., Karathia, H., Sorribas, A. & Alves, R. Two component systems: physiological effect of a third component. *PLoS One* **7**, e31095 (2012).

12. Knudsen, M., Feliu, E. & Wiuf, C. Exact analysis of intrinsic qualitative features of phosphorelays using mathematical models. *J. Theor. Biol.* **300**, 7–18 (2012).

13. Jovanovic, G. *et al.* Phosphorelay of non-orthodox two component systems functions through a bi-molecular mechanism in vivo: the case of ArcB. *Mol. Biosyst.* **11**, 1348–59 (2015).

14. Kothamachu, V. B., Feliu, E., Wiuf, C., Cardelli, L. & Soyer, O. S. Phosphorelays Provide Tunable Signal Processing Capabilities for the Cell. *PLoS Comput. Biol.* **9**, e1003322 (2013).

15. Sáez, M., Wiuf, C. & Feliu, E. Graphical reduction of reaction networks by linear elimination of species. *J. Math. Biol.* **74**, 195–237 (2017).

16. Salvado, B., Vilaprinyo, E., Sorribas, A. & Alves, R. A survey of HK, HPt, and RR domains and their organization in two-component systems and phosphorelay proteins of organisms with fully sequenced genomes. *PeerJ* **3**, e1183 (2015).

17. Kim, J.-R. & Cho, K.-H. The multi-step phosphorelay mechanism of unorthodox two-component systems in E. coli realizes ultrasensitivity to stimuli while maintaining

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

robustness to noises. *Comput. Biol. Chem.* **30**, 438–444 (2006).

18. Feliu, E., Knudsen, M. & Wiuf, C. Signaling Cascades: Consequences of Varying Substrate and Phosphatase Levels. in 81–94 (Springer, New York, NY, 2012). doi:10.1007/978-1-4419-7210-1_4

19. Samadani, A., Mettetal, J. & van Oudenaarden, A. Cellular asymmetry and individuality in directional sensing. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 11549–54 (2006).

20. Hazelbauer, G. L. & Lai, W.-C. Bacterial chemoreceptors: providing enhanced features to two-component signaling. *Curr. Opin. Microbiol.* **13**, 124–32 (2010).

21. Forst, S. A. & Roberts, D. L. Signal transduction by the EnvZ-OmpR phosphotransfer system in bacteria. *Res. Microbiol.* **145**, 363–73

22. Savageau, M. A. *Biochemical systems analysis : a study of function and design in molecular biology*. (Adisson-Wesley, 1976).

23. Alon, U. *An introduction to systems biology : design principles of biological circuits*. (Chapman & Hall/CRC, 2007).

24. Wang, M., Herrmann, C. J., Simonovic, M., Szklarczyk, D. & von Mering, C. Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. *Proteomics* **15**, 3163–3168 (2015).

25. Grimshaw, C. E. *et al.* Synergistic kinetic interactions between components of the phosphorelay controlling sporulation in Bacillus subtilis. *Biochemistry* **37**, 1365–1375 (1998).

26. Narula, J., Devi, S. N., Fujita, M. & Igoshin, O. A. Ultrasensitivity of the Bacillus subtilis sporulation decision. *Proc. Natl. Acad. Sci.* **109**, E3513–E3522 (2012).

27. Sen, S., Garcia-Ojalvo, J. & Elowitz, M. B. Dynamical consequences of bandpass feedback loops in a bacterial phosphorelay. *PLoS One* **6**, e25102 (2011).

28. Eswaramoorthy, P., Dinh, J., Duan, D., Igoshin, O. A. & Fujita, M. Single-cell measurement of the levels and distributions of the phosphorelay components in a population of sporulating Bacillus subtilis cells. *Microbiology* **156**, 2294–304 (2010).

29. Narula, J. *et al.* Chromosomal Arrangement of Phosphorelay Genes Couples Sporulation and DNA Replication. *Cell* **162**, 328–337 (2015).

30. Janiak-Spens, F., Cook, P. F. & West, A. H. Kinetic Analysis of YPD1-Dependent Phosphotransfer Reactions in the Yeast Osmoregulatory Phosphorelay System †. *Biochemistry* **44**, 377–386 (2005).

31. Fassler, J. S. & West, A. H. Histidine phosphotransfer proteins in fungal two-component signal transduction pathways. *Eukaryot. Cell* **12**, 1052–60 (2013).

32. Alves, R. & Savageau, M. A. Comparing systemic properties of ensembles of biological networks by graphical and statistical methods. *Bioinformatics* **16**, 527–33 (2000).

33. Alves, R. & Savageau, M. A. Effect of overall feedback inhibition in unbranched biosynthetic pathways. *Biophys. J.* **79**, 2290–304 (2000).

34. Alves, R. & Savageau, M. A. Irreversibility in unbranched pathways: preferred positions based on regulatory considerations. *Biophys. J.* **80**, 1174–85 (2001).

35. Çağatay, T., Turcotte, M., Elowitz, M. B., Garcia-Ojalvo, J. & Süel, G. M. Architecture-

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Dependent Noise Discriminates Functionally Analogous Differentiation Circuits. *Cell* **139**, 512–522 (2009).

36. Salvado, B., Vilaprinyo, E., Karathia, H., Sorribas, A. & Alves, R. Two component systems: physiological effect of a third component. *PLoS One* **7**, e31095 (2012).

37. Kothamachu, V. B., Feliu, E., Wiuf, C., Cardelli, L. & Soyer, O. S. Phosphorelays provide tunable signal processing capabilities for the cell. *PLoS Comput. Biol.* **9**, e1003322 (2013).

38. Amin, M. *et al.* Phosphate sink containing two-component signaling systems as tunable threshold devices. *PLoS Comput. Biol.* **10**, e1003890 (2014).

39. Feng, S., Sáez, M., Wiuf, C., Feliu, E. & Soyer, O. S. Core signalling motif displaying multistability through multi-state enzymes. *J. R. Soc. Interface* **13**, 20160524 (2016).

40. Shinar, G. & Feinberg, M. Concordant chemical reaction networks and the Species-Reaction Graph. *Math. Biosci.* **241**, 1–23 (2013).

41. Shinar, G. & Feinberg, M. Design principles for robust biochemical reaction networks: what works, what cannot work, and what might almost work. *Math. Biosci.* **231**, 39–48 (2011).

42. Shinar, G. & Feinberg, M. Structural sources of robustness in biochemical reaction networks. *Science* **327**, 1389–91 (2010).

43. Csikász-Nagy, A., Cardelli, L. & Soyer, O. S. Response dynamics of phosphorelays suggest their potential utility in cell signalling. *J. R. Soc. Interface* **8**, 480–488 (2011).

44. Uschner, F. & Klipp, E. Information processing in the adaptation of Saccharomyces cerevisiae to osmotic stress: an analysis of the phosphorelay system. *Syst. Synth. Biol.* **8**, 297–306 (2014).

45. Ray, J. C. J., Tabor, J. J. & Igoshin, O. A. Non-transcriptional regulatory processes shape transcriptional network dynamics. *Nat. Rev. Microbiol.* **9**, 817–28 (2011).

46. Savageau, M. A. Concepts relating the behavior of biochemical systems to their underlying molecular properties. *Arch. Biochem. Biophys.* **145**, 612–21 (1971).

47. Savageau, M. A. Parameter sensitivity as a criterion for evaluating and comparing the performance of biochemical systems. *Nature* **229**, 542–4 (1971).

48. Cheong, R., Rhee, A., Wang, C. J., Nemenman, I. & Levchenko, A. Information Transduction Capacity of Noisy Biochemical Signaling Networks. *Science (80-.).* **334**, 354–358 (2011).

49. Nemenman, I. Information Theory and Adaptation. in *Quantitative Biology* 73–92 (CRC Press, 2013).

50. Paulsson, J. & Ehrenberg, M. Random signal fluctuations can reduce random fluctuations in regulated components of chemical regulatory networks. *Phys. Rev. Lett.* **84**, 5447–50 (2000).

51. Paulsson, J., Berg, O. G. & Ehrenberg, M. Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 7148–53 (2000).

52. Pedraza, J. M. & Paulsson, J. Effects of molecular memory and bursting on fluctuations in gene expression. *Science* **319**, 339–43 (2008).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

53. Hilfinger, A. & Paulsson, J. Separating intrinsic from extrinsic fluctuations in dynamic biological systems. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 12167–72 (2011).

54. Lestas, I., Vinnicombe, G. & Paulsson, J. Fundamental limits on the suppression of molecular fluctuations. *Nature* **467**, 174–8 (2010).

55. Pedraza, J. M. & Paulsson, J. Random timing in signaling cascades. *Mol. Syst. Biol.* **3**, 81 (2007).

56. Narula, J., Smith, A. M., Gottgens, B. & Igoshin, O. A. Modeling reveals bistability and low-pass filtering in the network module determining blood stem cell fate. *PLoS Comput. Biol.* **6**, e1000771 (2010).

57. Tiwari, A. & Igoshin, O. A. Coupling between feedback loops in autoregulatory networks affects bistability range, open-loop gain and switching times. *Phys. Biol.* **9**, 055003 (2012).

58. Cağatay, T., Turcotte, M., Elowitz, M. B., Garcia-Ojalvo, J. & Süel, G. M. Architecture-dependent noise discriminates functionally analogous differentiation circuits. *Cell* **139**, 512–22 (2009).

59. Tiwari, A., Balázsi, G., Gennaro, M. L. & Igoshin, O. A. The interplay of multiple feedback loops with post-translational kinetics results in bistability of mycobacterial stress response. *Phys. Biol.* **7**, 036005 (2010).

60. Alves, R., Vilaprinyo, E., Hernández-Bermejo, B. & Sorribas, A. Mathematical formalisms based on approximated kinetic representations for modeling genetic and metabolic pathways. *Biotechnol. Genet. Eng. Rev.* **25**, (2008).

61. Dill, K. A., Ghosh, K. & Schmit, J. D. Physical limits of cells and proteomes. *Proc. Natl. Acad. Sci.* **108**, 17876–17882 (2011).

62. Milo, R. & Phillips, R. *Cell biology by the numbers*. (Garland Science, 2015).

63. Uri, M. & Milo, R. Average concentration of a protein in cell - Budding yeast Saccharomyces ce - BNID 104520. (2010). Available at: http://bionumbers.hms.harvard.edu/bionumber.aspx?&id=104520&ver=10&trm=cell protein concentration. (Accessed: 12th April 2018)

64. Glont, M. *et al.* BioModels: expanding horizons to include more modelling approaches and formats. *Nucleic Acids Res.* **46**, D1248–D1253 (2018).

65. Kitano, H. Computational systems biology. *Nature* **420**, 206–210 (2002).

66. Gillespie, D. T., Hellander, A. & Petzold, L. R. Perspective: Stochastic algorithms for chemical kinetics. *J. Chem. Phys.* **138**, 170901 (2013).

67. Hirschman, J. E. Genome Snapshot: a new resource at the Saccharomyces Genome Database (SGD) presenting an overview of the Saccharomyces cerevisiae genome. *Nucleic Acids Res.* **34**, D442–D445 (2006).

68. Wolfram Research, I. *Mathematica*. (Wolfram Research, Inc, 2014).

69. Berens, M. L., Berry, H. M., Mine, A., Argueso, C. T. & Tsuda, K. Evolution of Hormone Signaling Networks in Plant Defense. *Annu. Rev. Phytopathol.* **55**, 401–425 (2017).

70. Espinar, L., Dies, M., Cagatay, T., Suel, G. M. & Garcia-Ojalvo, J. Circuit-level input integration in bacterial gene regulation. *Proc. Natl. Acad. Sci.* **110**, 7091–7096 (2013).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

71. Kuchina, A. *et al.* Temporal competition between differentiation programs determines cell fate choice. *Mol. Syst. Biol.* **7**, 557 (2011).

72. Kærn, M., Elston, T. C., Blake, W. J. & Collins, J. J. Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.* **6**, 451–464 (2005).

73. Raj, A. & van Oudenaarden, A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216–26 (2008).

74. Kaufmann, B. B. & van Oudenaarden, A. Stochastic gene expression: from single molecules to the proteome. *Curr. Opin. Genet. Dev.* **17**, 107–12 (2007).

75. Shimell, M. *et al.* Prothoracicotropic hormone modulates environmental adaptive plasticity through the control of developmental timing. *Development* **145**, dev159699 (2018).

76. Noor, E. *et al.* The Protein Cost of Metabolic Fluxes: Prediction from Enzymatic Rate Laws and Cost Minimization. *PLOS Comput. Biol.* **12**, e1005167 (2016).

77. Pereira, T. *et al.* Quantitative Operating Principles of Yeast Metabolism during Adaptation to Heat Stress. *Cell Rep.* **22**, 2455–2468 (2018).

78. Weiss, R., Kukora, J. & Adams J. The relationship between enzyme activity, cell geometry, and fitness in Saccharomyces cerevisiae. - PubMed - NCBI. *Proc. Natl. Acad. Sci. USA* **72**, 794–798 (1975).

79. Dénervaud, N. *et al.* A chemostat array enables the spatio-temporal analysis of the yeast proteome. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15842–7 (2013).

80. Wu, J.-Q. & Pollard, T. D. Counting Cytokinesis Proteins Globally and Locally in Fission Yeast. *Science (80-.).* **310**, 310–314 (2005).

81. Mitchison, J. M. The growth of single cells: I. Schizosaccharomyces pombe. *Exp. Cell Res.* **13**, 244–262 (1957).

82. Gunaratne, J. *et al.* Extensive Mass Spectrometry-based Analysis of the Fission Yeast Proteome. *Mol. Cell. Proteomics* **12**, 1741–1751 (2013).

83. García-Rivera, J., Chang, Y. C., Kwon-Chung, K. J. & Casadevall, A. Cryptococcus neoformans CAP59 (or Cap59p) is involved in the extracellular trafficking of capsular glucuronoxylomannan. *Eukaryot. Cell* **3**, 385–92 (2004).

84. Santi, L. *et al.* Proteomic Profile of *Cryptococcus neoformans* Biofilm Reveals Changes in Metabolic Processes. *J. Proteome Res.* **13**, 1545–1559 (2014).

85. Melaragno, J. E., Mehrotra, B. & Coleman, A. W. Relationship between Endopolyploidy and Cell Size in Epidermal Tissue of Arabidopsis. *Plant Cell* **5**, 1661–1668 (1993).

86. Castellana, N. E. *et al.* Discovery and revision of Arabidopsis genes by proteogenomics. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 21034–8 (2008).

87. Kubitschek, H. E. & Friske, J. A. Determination of bacterial cell volume with the Coulter Counter. *J. Bacteriol.* **168**, 1466–7 (1986).

88. Loferer-Krössbacher, M., Klima, J. & Psenner, R. Determination of bacterial cell dry mass by transmission electron microscopy and densitometric image analysis. *Appl. Environ. Microbiol.* **64**, 688–94 (1998).

89. Wiśniewski, J. R. & Rakus, D. Multi-enzyme digestion FASP and the 'Total Protein

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Approach'-based absolute quantification of the Escherichia coli proteome. *J. Proteomics* **109**, 322–331 (2014).

90. Gonçalves, A. *et al.* Comparative proteomics of an extended spectrum β-lactamase producing Escherichia coli strain from the Iberian wolf. *J. Proteomics* **104**, 80–93 (2014).

91. Pieper, R. *et al.* Analysis of the proteome of intracellular Shigella flexneri reveals pathways important for intracellular growth. *Infect. Immun.* **81**, 4635–48 (2013).

92. Braisted, J. PRIDE Archive. Available at: https://www.ebi.ac.uk/pride/archive/projects/PRD000545;jsessionid=019280191DA68B869DC1ED3D42CC1C06. (Accessed: 13th April 2018)

93. Lynch, M. & Marinov, G. K. The bioenergetic costs of a gene. *Proc. Natl. Acad. Sci.* **112**, 201514974 (2015).

94. Toyofuku, M., Roschitzki, B., Riedel, K. & Eberl, L. Identification of proteins associated with the Pseudomonas aeruginosa biofilm extracellular matrix. *J. Proteome Res.* **11**, 4906–15 (2012).

95. Robador, A., LaRowe, D. E., Finkel, S. E., Amend, J. P. & Nealson, K. H. Changes in Microbial Energy Metabolism Measured by Nanocalorimetry during Growth Phase Transitions. *Front. Microbiol.* **9**, 109 (2018).

96. Fievet, A. *et al.* Single-Cell Analysis of Growth and Cell Division of the Anaerobe Desulfovibrio vulgaris Hildenborough. *Front. Microbiol.* **6**, 1378 (2015).

97. Zhang, W. *et al.* LC-MS/MS based proteomic analysis and functional inference of hypothetical proteins in Desulfovibrio vulgaris. *Biochem. Biophys. Res. Commun.* **349**, 1412–9 (2006).

98. Zhang, W. *et al.* Global transcriptomic analysis of Desulfovibrio vulgaris on different electron donors. *Antonie Van Leeuwenhoek* **89**, 221–37 (2006).

99. Narula, J., Devi, S. N., Fujita, M. & Igoshin, O. A. Ultrasensitivity of the Bacillus subtilis sporulation decision. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E3513-22 (2012).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure captions

Figure 1 – The five most abundant phosphorelay circuit architectures, as inferred from operon structure, account for over 99% of all detected phosphorelays. Architecture M1 is for a circuit where the four phosphorylatable domains exist in independent proteins. Architecture M2 is for a circuit with a hybrid Sensor Kinase (SK), which contains the SK and the first Response Regulator (RR_1) domain in the same protein, while the remaining phosphorylatable domains exist in independent proteins. Architecture M2' is for a circuit where the SK and the Hpt domains are in the same protein, while both RR domains exist in independent proteins. Architecture M3 is for a circuit where the SK, RR_1 and the Hpt domains are in the same protein, while the final RR, RR_2, is in an independent protein. Architecture M4 is for a circuit where all phosphorylatable domains exist in the same protein. A total of 5219 PR operons were surveyed, out of which 5182 fall in one of the five architectures shown here.

Figure 2 – Clustering each architecture according to their abundance ranks in the ten percent best systems with respect to each of the physiological variables in Table 2. Variables that are absent in the clustering have an equal percentage of all architectures and therefore have equivalent ranking.

Figure 3 – Effects of alternative architectures in the deterministic regime for the Spo0 phosphorelay of *Bacillus subtilis* and the Sln1 phosphorelay of *Saccharomyces cerevisiae*. In all cases, the protein amounts of the alternative architectures were optimized to make the steady state signal-response curves be as similar as possible (see methods). The black arrow indicates the native architecture. Panels A, B, and C pertain to the Spo0 phosphorelay. Panels D, E, and F pertain to the Sln1 phosphorelay. A, D – Cost of synthesizing the circuit under different architectures. X – axis: PR architecture. Y – axis: total metabolic cost of the circuit proteins (arbitrary units). B, E – normalized sensitivity of the steady state concentration of the final response regulator to changes in parameters. X – axis: PR architecture. Y – axis: euclidean norm of the sensitivities vector of the response regulator. C, F – Percentage of transient responses where each architecture is the fastest (outer donut) or slowest (inner donut) in reaching 90% of the new steady state under equivalent conditions. Left-side donuts from responses to phosphorylating signals. Right-side donuts from responses to dephosphorylating signals.

Figure 4 – Effects of alternative architectures in the transmission of information through the Spo0 phosphorelay of *Bacillus subtilis* and the Sln1 phosphorelay of *Saccharomyces cerevisiae*. In all cases, the protein amounts of the alternative architectures were optimized to make the steady state signal-response curves be as similar as possible. **A** – Spo0 phosphorelay. Y-axis represents the accumulated mutual information over a range of 6 orders of magnitude for the self-dephosphorylation rate constant of kinA between variations in the number of phosphorylated kinA molecules and that of phosphorylated Spo0E molecules. k1 represents modulation of the kinA phosphorylation rate, while k18 represents modulation of SpoE dephosphorylation. Architecture M1 transmits the most information for comparable parameter values. **B** – Sln1 phosphorelay. Y-axis represents the accumulated mutual information over a range of 6 orders of magnitude for the self-dephosphorylation rate constant of Sln1 between variations in the number of phosphorylated Sln1 molecules and that of phosphorylated Ssk1 or Skn7 molecules. k1 represents modulation of the Sln1 phosphorylation rate, k18 represents modulation of Ssk1 dephosphorylation, and k26 represents modulation of Skn7 dephosphorylation. Architecture M2 transmits the most information to Ssk1 for comparable parameter values.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table 1 – Physiological variables used as proxy for performance in signal transduction circuit.

Variables	Circuit performance improves with*:	Experimental support
Signal amplification	Higher amplification	[9,10,36,69–72]
Noise attenuation^	Attenuated noise	[57,72–74]
Information transmission^	Higher transmission	[57,72–74]
Robustness to changes in parameter values	High robustness (low sensitivity)	[9,10,36,58,69]
Speed of response to changes	Rate of adaptation	[9,10,29,36,48,71,75]
Metabolic cost of circuit	Low cost	[76,77]

*as a general trend; ^ Relevant in the stochastic domains of dynamic behavior.

Table 2 – Comparative results for the various physiological variables. Dark green indicate best performance. Dark yellow indicates worst performance.

Physiological Variable	M1	M2	M2'	M3	M4
Signal amplification	4.3 %	10.3 %	36.2 %	33.8 %	15.4 %
Robustness	21.6 %	26.1 %	18.5 %	17.5 %	16.3 %
Intrinsic Noise SK	1.6 %	5.4 %	8.6 %	30.3 %	54. %
Intrisic Noise RR2	19.8 %	17.8 %	18.1 %	23.5 %	20.8 %
INRR2/INSK	6.1 %	15.3 %	15.1 %	31.2 %	32.3 %
SK->RR1 MI	28.9 %	24.9 %	24 %	13.3 %	9 %
SK->RR2 MI	29.4 %	24.3 %	24.5 %	11.1 %	10.7 %
SK->RR2/SK->RR1 MI	9.6 %	8.8 %	18.5 %	14.9 %	48.2 %
Response time	29. %	21.9 %	25.4 %	14. %	9.6 %
Cost	1.9 %	5.1 %	8.3 %	31.2 %	53.4 %
Genomic frequency	***0.6*** %	***1.6*** %	***18.3*** %	***72.3*** %	***6.5*** %

Physiological Variable	Rank M1	Rank M2	Rank M2'	Rank M3	Rank M4
Signal amplification	5	4	1	2	3
Robustness	2	1	3	4	5
Intrinsic Noise SK	5	4	3	2	1
Intrisic Noise RR2	3	5	4	1	2
INRR2/INSK	5	3	4	2	1
SK->RR1 MI	1	2	3	4	5
SK->RR2 MI	1	2	2	4	5
SK->RR2/SK->RR1 MI	4	5	2	3	1
Response time	1	3	2	4	5
Cost	5	4	3	2	1
Genomic frequency rank	***5***	***4***	***2***	***1***	***3***

Table 3 – Observed native architectures and predictions of where in parameter space their observation is expected.

Organism	Phosphorelay	Ratios of abundance (order of magnitude)	Estimated proteins per cell[1]	Predicted operational regions
	TorS:TorR (M3)	1:1:1:10	100:1000	***k1, k2>10^{-1} s^{-1} ****
	EvgS:EvgA (M3)	1:1:1:100	100:10000	***k1, k2>10^{-1} s^{-1} ****
	BarA:UvrY (M3)	1:1:1:10	1000:10000	***k1, k2>10^{-1} s^{-1} ****
Escherichia coli	ArcB:ArcA (M3)	1:1:1:10	10000:1000000	***k1, k2>10^{-1} s^{-1} ****
	RcsC:RcsD:RcsB (M2)	1:1:1:100	1000:1000:100000	***k1>10^{-1} s^{-1}****
	BarA:UvrY (M3)	1:1:1:1	1000:1000	M1 or M2’
Shigella flexneri	ArcB:ArcA (M3)	1:1:1:10	10000: 100000	***k1>10^{-1} s^{-1}*+**
Shewanella oneidensis	SO0859:SO0860 (M3)	1:1:1:1	100000:100000	***k1>10^{-1} s^{-1}*++**
Desulfovibrio vulgaris	DVU_3062:DVU_3061 (M3)	1:1:1:1	100000:100000	***k1>10^{-1} s^{-1}*++**
	Sln1:Ypd1:Ssk1 (M2)	1:1:1:1	1000:1000:1000	*****
Saccharomyces cerevisiae	Sln1:Ypd1:Skn7 (M2)	1:1:1:1	1000:1000:1000	*****
	Mak1:Mpr1:Mcs4 (M2)	1:1:1:10	1000:1000:10000	******
	Mak2:Mpr1:Mcs4 (M2)	1:1:1:10	1000:1000:10000	******
Schizosaccharomyces pombe	Mak3:Mpr1:Mcs4 (M2)	1:1:1:10	1000:1000:10000	******
Bacillus subtilis	KinA:Spo0F:Spo0B:Spo0A (M1)	1:100:1:100	12:4200:110:1700	***k1, k2<10^{-1} s^{-1}. *+**
	KinB:Spo0F:Spo0B:Spo0A (M1)	1:100:1:100	93:4200:110:1700	***k1, k2<10^{-1} s^{-1}. *+**
	KinC:Spo0F:Spo0B:Spo0A (M1)	1:100:1:100	82:4200:110:1700	***k1, k2<10^{-1} s^{-1}. *+**

[1] These numbers are calculated by multiplying cell volume (μm^3), average number of proteins in cells per μm^3, and the protein abundance in parts per million: $Cell\ Volume\ \times 6.23 \times 10^8 \times Protein\ abundances \times 10^{-6}$

* As compared to architectures M1, M2, M2'. M4 does not allow for the observed abundance ratio between signal transduction domains.

** As compared to M1, the only other architecture that allows for the observed ratio between abundances of signal transduction domains. Regulation by the environment expected at the SK phosphorylation step.

+ As compared to architectures M1, M2, M2'. M4 does not allow for the observed abundance ratio between signal transduction domains. Regulation by the environment expected at the SK phosphorylation step.

++ As compared to architectures M1, M2, M2' and M4. Regulation by the environment expected at the SK phosphorylation step.

***** See analysis of the system as a Sln1-Ypd1-Skn7-Ssk1 PR in the main text.

******Outside the range of protein abundances tested in this work. Nevertheless, comparing the trends of similar abundance ratios for one order of magnitude less suggests that M2 would be the preferred architecture if we pool the abundances of Mac1, Mac2 and Mac 3 proteins together.

$^{*+}$ Comparison between architectures M1 and M2', which are the only ones that allow for this ratio of abundance between domains. Consistent with experimental determinations of the rate constants (supplementary materials).

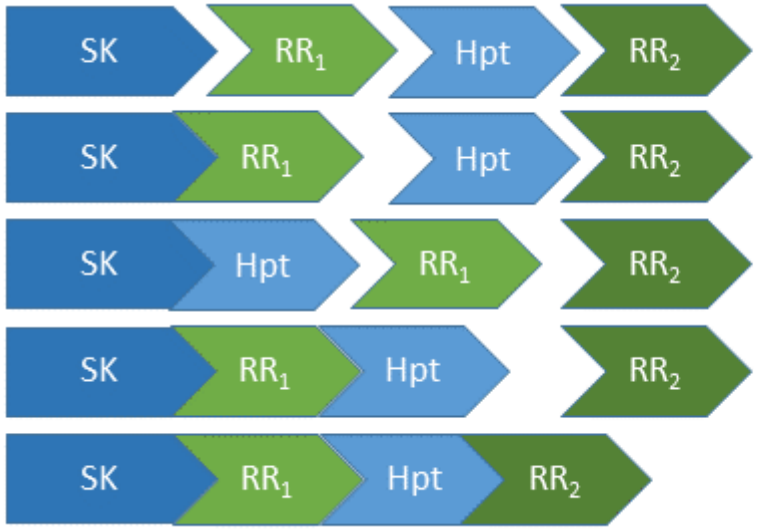


Figure 1

Architecture M1 (0.6%)

Architecture M2 (1.6%)

Architecture M2’ (18.3%)

Architecture M3 (72.3%)

Architecture M4 (6.5%)

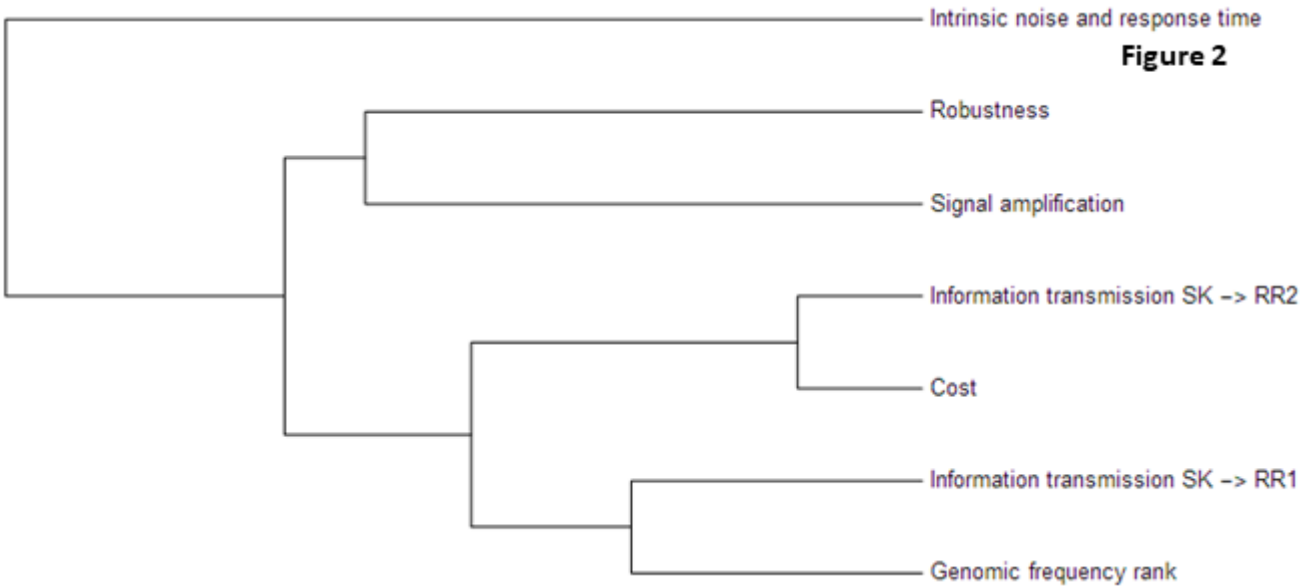
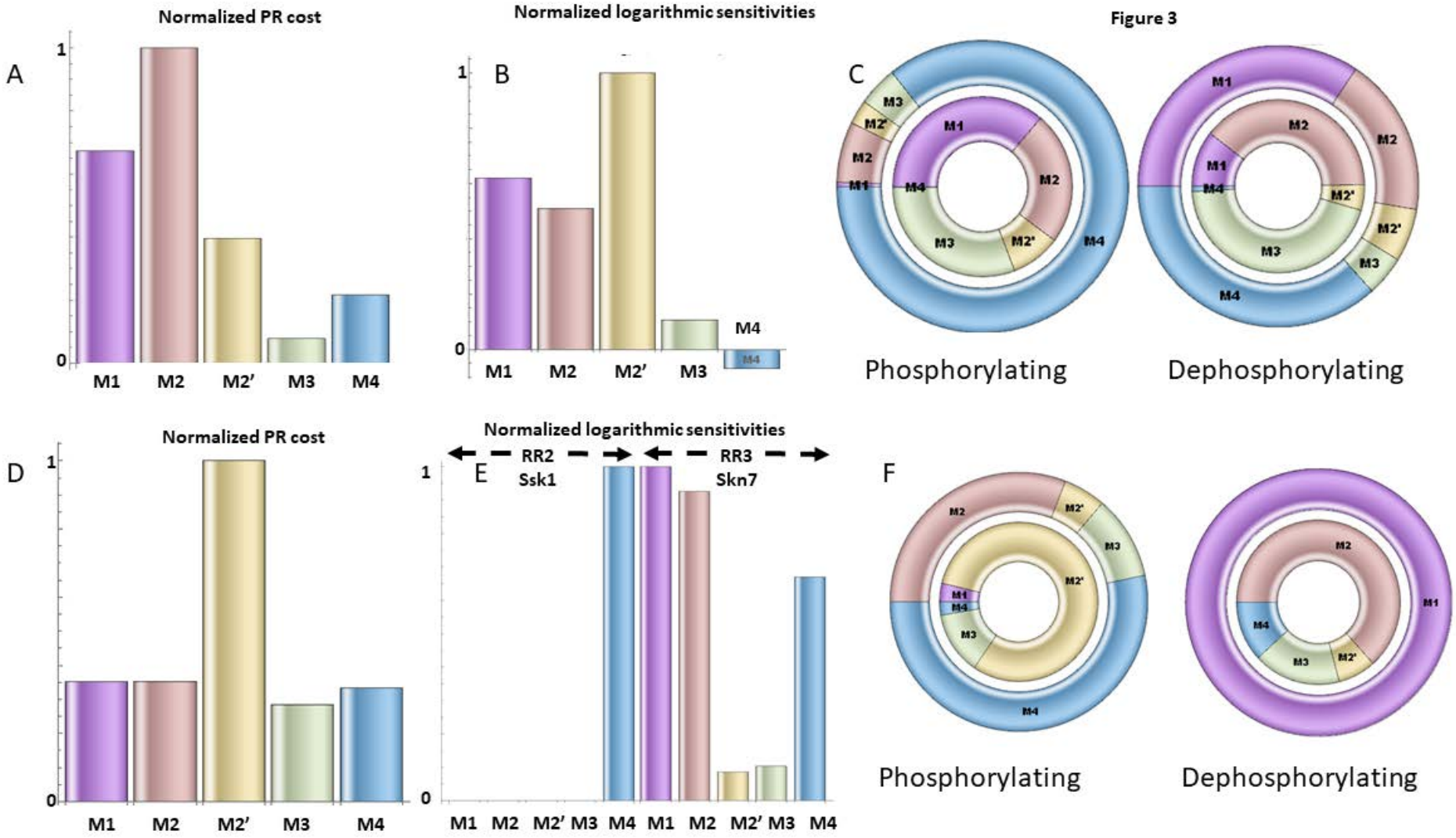

Intrinsic noise and response time
Robustness
Signal amplification
Information transmission SK –> RR2
Cost
Information transmission SK –> RR1
Genomic frequency rank

Figure 2

Figure 3


A
Normalized PR cost
1
0
M1
M2
M2'
M3
M4
B
Normalized logarithmic sensitivities
1
0
M1
M2
M2'
M3
M4
C
Phosphorylating
Dephosphorylating
D
Normalized PR cost
1
0
M1
M2
M2'
M3
M4
E
Normalized logarithmic sensitivities
RR2
Ssk1
RR3
Skn7
1
0
M1 M2 M2' M3 M4 M1 M2 M2' M3 M4
F
Phosphorylating
Dephosphorylating

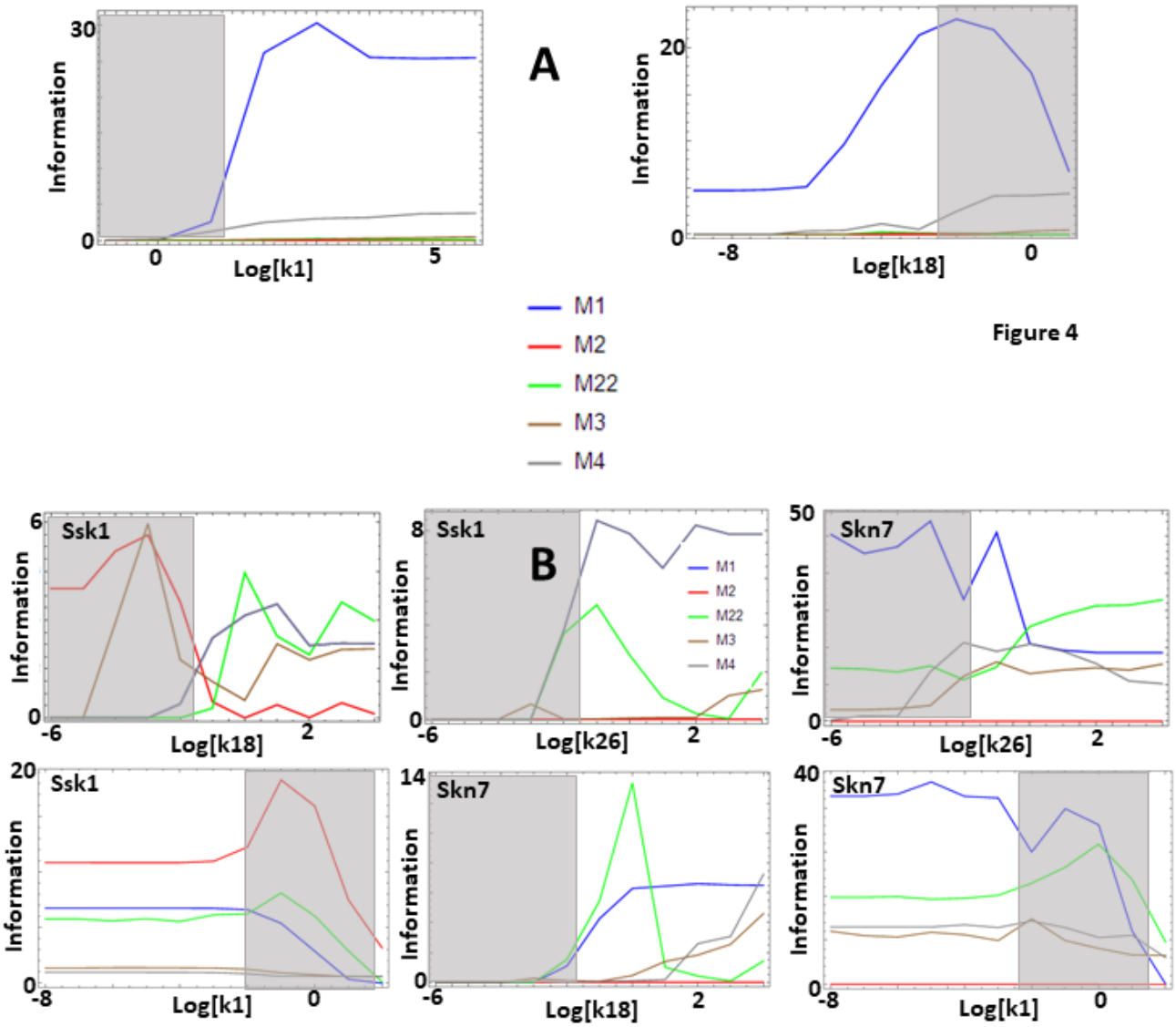


Figure 4

Table S1 – Correlation analysis of variable in different architectures: Spearman correlation

Architecture M1

	L[X4,X5]	L[X4,X6]	L[X4,Xt]	S(X4,.)	Noise X4	Noise X4/Noise X1	Slowest eigenvalue	tau on	tau off	MIX1X2	MIX1X4
L[X4,X5]	1.			0.356138							
L[X4,X6]		1.		-0.207535							
L[X4,Xt]			1.	0.253767						-0.225825	-0.229017
S(X4,.)				1.			-0.472012	-0.30121			
Noise X4					1.	0.590393					
Noise X4/Noise X1						1.				0.485753	0.496824
Slowest eigenvalue							1.	0.465043	0.287862	-0.298759	-0.295689
tau on								1.	0.57938		
tau off									1.		
MIX1X2										1.	0.979351
MIX1X4											1.

Architecture M2

	L[X4,X5]	L[X4,X6]	L[X4,Xt]	S(X4,.)	Noise X4	Noise X4/Noise X1	Slowest eigenvalue	tau on	tau off	MIX1X2	MIX1X4
L[X4,X5]	1.	-0.516553	0.286692	0.529419							
L[X4,X6]		1.	-0.400306	-0.498105							
L[X4,Xt]			1.	0.349106		-0.220468				-0.210107	-0.3032
S(X4,.)				1.			-0.267596	-0.251508			
Noise X4					1.	0.807993				0.6077	0.69371
Noise X4/Noise X1						1.				0.59542	0.67683
Slowest eigenvalue							1.	0.499252	0.302342	-0.231062	-0.2987
tau on								1.	0.575868		
tau off									1.		
MIX1X2										1.	0.83992
MIX1X4											1.

Architecture M2’

	L[X4,X5]	L[X4,X6]	L[X4,Xt]	S[X4,.]	Noise X4	Noise X4/Noise X1	Slowest eigenvalue	tau on	tau off	MIX1X2	MIX1X4
L[X4,X5]	1.	-0.480821	0.595986	0.526117			0.32277				
L[X4,X6]		1.	-0.359513	-0.597799			-0.320074			0.216402	0.214074
L[X4,Xt]			1.	0.36662	-0.217077		0.298417			-0.333982	-0.33202
S[X4,.]				1.							
Noise X4					1.	0.810644	-0.265015			0.680968	0.696602
Noise X4/Noise X1						1.				0.662303	0.652146
Slowest eigenvalue							1.	0.324938	0.302736	-0.395	-0.40638
tau on								1.	0.513862		
tau off									1.	-0.240163	-0.25476
MIX1X2										1.	0.947977
MIX1X4											1.

Architecture M3

	L[X4,X5]	L[X4,X6]	L[X4,Xt]	S[X4,.]	Noise X4	Noise X4/Noise X1	Slowest eigenvalue	tau on	tau off	MIX1X2	MIX1X4
L[X4,X5]	1.	-0.453484	0.527709	0.625872							
L[X4,X6]		1.	-0.449944	-0.622401							
L[X4,Xt]			1.	0.613068			0.200176				-0.247475
S[X4,.]				1.							
Noise X4					1.	0.708935				0.560928	0.727969
Noise X4/Noise X1						1.				0.438144	0.456246
Slowest eigenvalue							1.	0.32365	0.22171	-0.319489	-0.40039
tau on								1.	0.67316		
tau off									1.		
MIX1X2										1.	0.687972
MIX1X4											1.

Architecture M4

	L[X4,X5]	L[X4,X6]	L[X4,Xt]	S(X4,.)	Noise X4	Noise X4/Noise X1	Slowest eigenvalue	tau on	tau off	MIX1X2	MIX1X4
L[X4,X5]	1.		0.441695	0.472136							
L[X4,X6]		1.	-0.231123	-0.320598							
L[X4,Xt]			1.	0.500891							
S(X4,.)				1.							
Noise X4					1.	0.724457				0.322111	0.329864
Noise X4/Noise X1						1.	0.252565				
Slowest eigenvalue							1.	0.31013	0.220807	-0.455544	-0.513793
tau on								1.	0.717326		
tau off									1.		
MIX1X2										1.	0.792357
MIX1X4											1.

Table S2 – Performance atlas with respect to information transmission through the phosphorelay as a function of architecture, protein amounts, and parameter values

Table S3 – Phosphorelay circuits for which there is abundance information.

Organism (average cell volume in μm^3)	Phosphorelay	Protein abundances (ppm)	Estimated number of proteins per cell[2]	Ratios of abundance (order of magnitude)
Saccharomyces cerevisiae (~ 70 [78])	Sln1:Ypd1:Ssk1 HKRR:Hpt:RR	151:188:177 [79]	66000: 82000: 77000	1:1:1:1
	Sln1:Ypd1:Skn7 HKRR:Hpt:RR	151:188:175 [79]	66000: 82000: 77000	1:1:1:1
Schizosaccharomyces pombe (~ 100 [80,81])	Mak1:Mpr1:Mcs4 HKRR:Hpt:RR	44:62:470 [82]	11000: 39000: 293000	1:1:1:10
	Mak2:Mpr1:Mcs4 HKRR:Hpt:RR	53:62:470 [82]	33000: 39000: 293000	1:1:1:10
	Mak3:Mpr1:Mcs4 HKRR:Hpt:RR	53:62:470 [82]	33000: 39000: 293000	1:1:1:10
Cryptococcus neoformans (~ 50-70 [83])	CNH02350 HKRRHptRR	2.63 [84]	980	1:1:1:1
	CNC03340:CNM01530:CNB03090 HKRR:Hpt:RR	42:367:14 [84]	16000: 137000: 5000	10:10:100:1
	CNA03400:CNM01530:CNB03090 HKRR:Hpt:RR	3:367:14 [84]	1000: 137000: 5000	1:1:100:1

[2] This number is calculated by multiplying cell volume (μm^3) by the average number of proteins in cells per μm^3 times the protein abundance in parts per million: $\frac{Cell\ Volume \times 6.23\times10^6\times Protein\ abundances}{10^6}$

Organism (average cell volume in μm^3)	Phosphorelay	Protein abundances (ppm)	Estimated number of proteins per cell[3]	Ratios of abundance (order of magnitude)
Arabidopsis thaliana (2000 [85])	AKH2:AHP1-AHP4 HKRRRR:Hpt	0.15:8-50 [86]	2000: 99000-623000	1:1:10:1 1:1:100:1
	AKH3:AHP1-AHP4 HKRRRR:Hpt	1: 8-50 [86]	12000: 99000-623000	1:1:10:1 1:100:1:1
	AKH4:AHP1-AHP4 HKRRRR:Hpt	0.62: 8-50 [86]	8000:99000-623000	1:1:10:1 1:1:100:1
	AHK1:AHP2-AHP3:ARR1-3,5,7,9,14,16 HKRR:Hpt:RR	0.3:15-50:0.01-2 [86]	4000: 187000-623000:100-25000	1:1:100:1 1:1:100:10 10:10:1000:1
	AHK5:AHp1-3, AHP5-6:ARR1-2 HKRR:Hpt:RR	0.1:9-50:2 [86]	1200: 112000-623000: 25000	1:1:100:10
	ETR1:AHP5:RR1-2 HKRR:Hpt:RR	2.5:0.02:2 [86]	31000: 250-623000: 25000	1:1:1:1 10:10:1:10

[3] This number is calculated by multiplying cell volume (μm^3) by the average number of proteins in cells per μm^3 times the protein abundance in parts per million: $\frac{Cell\ Volume \times 6.23\times10^6\times Protein\ abundances}{10^6}$

Organism (average cell volume in μm^3)	Phosphorelay	Protein abundances (ppm)	Estimated number of proteins per cell[4]	Ratios of abundance (order of magnitude)
Escherichia coli (~ 0.1-3 [87,88])	TorS:TorR HKRRHpt:RR	0.34:6 [89,90]	212:3728	1:1:1:10
	EvgS:EvgA HKRRHpt:RR	0.31:88 [89,90]	193:54800	1:1:1:100
	BarA:UvrY HKRRHpt:RR	3.5:55 [89,90]	2180:34265	1:1:1:10
	ArcB:ArcA HKRRHpt:RR	51:1065 [89,90]	30:6600	1:1:1:10
	RcsC:RcsD:RcsB HKRR:Hpt:RR	3.:7.4:375 [89,90]	19:46: 2340	1:1:1:100
Shigella flexneri (~ 0.1-3 [87,88])	BarA:UvrY HKRRHpt:RR	3.94:8.74 [91,92]	24:54	1:1:1:1
	ArcB:ArcA HKRRHpt:RR	34.9:1094	220:.6800	1:1:1:100
Pseudomonas aeruginosa (~ 0.5 [93])	PA0413:PA0416 HKRRHpt:RR	158:BDL [90,94]	580:-	1:<1
	PA4112 HKRRHptRR	7.2[90,94]	26	1:1:1:1
	SO3207:SO3206:SO3209 HKHpt:RR:RR	452:17.1:768 [90]	2800: 110: 4800	10:1:10:10 10:1:10:100
Shewanella oneidensis (1-4 [95])	SO0859:SO0860 HKRRHpt:RR	88:79 [90]	550:490	1:1:1:1
Desulfovibrio vulgaris (2.4 [96])	DVU_3062:DVU_3061 HKRRHpt:RR	80:79 [97,98]	1200:1200	1:1:1:1

[4] These numbers are calculated by multiplying cell volume (μm^3), average number of proteins in cells per μm^3, and the protein abundance in parts per million: $\frac{Cell\ Volume \times 6.23\times10^6\times Protein\ abundances}{10^6}$

Organism (average cell volume in μm^3)	Phosphorelay	Protein abundances (ppm)	Estimated number of proteins per cell[5]	Ratios of abundance (order of magnitude)
Salmonella enterica (0.8 [93])	RcsC:RcsD:RcsB HKRR:Hpt:RR	3.4:3.9:327 [90]	17:19:1600	1:1:1:100
Bacillus subtilis (1.5 [93])	KinA:Spo0F:Spo0B:Spo0A HK:RR:Hpt:RR	1.3:446:12.4:185 [90]	12:4200:110:1700	1:100:1:100
	KinB:Spo0F:Spo0B:Spo0A HK:RR:Hpt:RR	10:446:12.4:185 [90]	93:4200:110:1700	1:100:1:100
	KinC:Spo0F:Spo0B:Spo0A HK:RR:Hpt:RR	8.8:446:12.4:185 [90]	82:4200:110:1700	1:100:1:100

BDL: Below detection limit

[5] These numbers are calculated by multiplying cell volume (μm^3), average number of proteins in cells per μm^3, and the protein abundance in parts per million: $\frac{Cell\ Volume \times 6.23\times10^6\times Protein\ abundances}{10^6}$

Table S4 – Numerical values of the rate constants of the 19 reactions in the PR models used in the stochastic simulations, represented in Figure 4. These values are based on [29]. We found a range of variation in four of the rate constants (k1, k2, k10 and k19) [25–27,43]. For the generality of the results, we used both the minimum and the maximum value of these four rate constants, obtaining 16 different combinations of numerical values. Each one of these 16 combinations of rate constant values was used in each one of the 5 PR models to obtain 100 stochastic trajectories.

Rate Constant	Experimental values	Scanning	Type of scanning
k1	Min= 0.002 s^{-1} Middle = 0.028 s^{-1} Max = 20 s^{-1}	Yes	Continuous, from 10^{-4} s^{-1} to 10^{3} s^{-1}
k2	Min= 0 s^{-1} Max = 0.00055 s^{-1}	Yes	Continuous, from 10^{-4} s^{-1} to 10^{3} s^{-1}
k3, k8, k11, k14, k17	0.000139 nM^{-1} s^{-1}	No	–
k4	0.083 s^{-1}	No	–
k5	0.083 s^{-1}	No	–
k9, k18	0.028 s^{-1}	No	–
k10, k19	Min=0.0056 s^{-1} Max=0.028 s^{-1}	Yes	Discrete, with all possible combinations of k10 and k19 and $k10 \in \{0.0056, 0.028\}$ and $k19 \in \{0.0056, 0.028\}$
k12	0.139 s^{-1}	No	–
k13, k16	0.22 s^{-1}	No	–
k15	0.056 s^{-1}	No	–

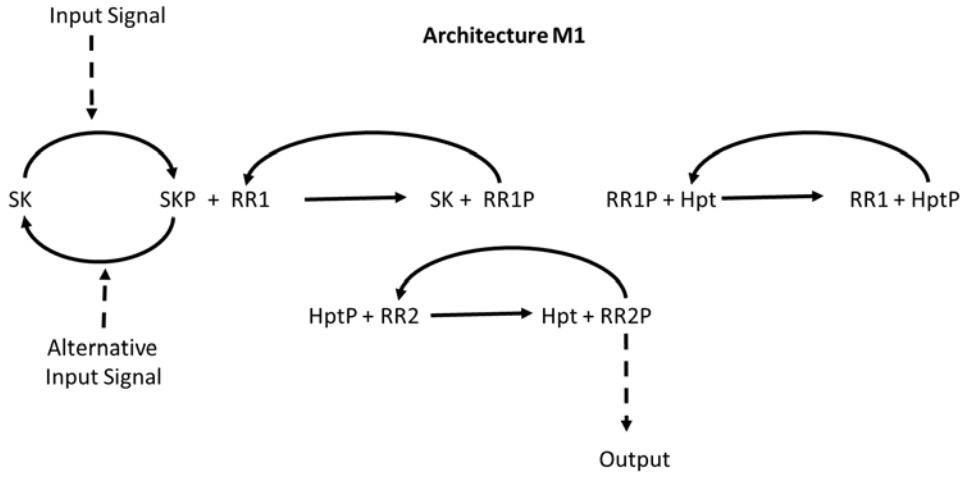
bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary text S1 – Mathematical modelling

Power law models

Architecture M1

The conceptual scheme we used to derive the power law mathematical models for architecture M1 is as follows



S-Systems mathematical model notation:
X1 – SKP; X2 – RR1P; X3 – HptP; X4 – RR2P; X5 – Input Signal; X6 – Alternative Input Signal; X7 – SK; X8 – RR1; X9 – Hpt; X10 – RR2
X1t – SK + SKP; X2t – RR1 + RR1P; X3t – Hpt +HptP; X4t – RR2 + RR2P

GMA Model

$$\frac{dX1}{dt} = -\frac{dX7}{dt} = \alpha_1 {X_7}^{g17} X_5^{g15} - \alpha_2 {X_1}^{g21} X_6^{g26} - \alpha_3 {X_1}^{g31} X_8^{g38}$$

$$\frac{dX2}{dt} = -\frac{dX8}{dt} = \alpha_3 {X_1}^{g31} X_8^{g38} - \alpha_4 {X_2}^{g42} X_9^{g49} - \alpha_5 {X_2}^{g52}$$

$$\frac{dX3}{dt} = -\frac{dX9}{dt} = \alpha_4 {X_2}^{g42} X_9^{g49} - \alpha_6 {X_3}^{g63} - \alpha_7 {X_3}^{g73} X_{10}^{g710}$$

$$\frac{dX4}{dt} = -\frac{dX10}{dt} = \alpha_7 {X_3}^{g73} X_{10}^{g710} - \alpha_8 {X_4}^{g84}$$

S-system Model

$$\frac{dX1}{dt} = -\frac{dX7}{dt} = \alpha_1 {X_7}^{g17} X_5^{g15} - \alpha_2{}' {X_1}^{g21'} X_6^{g26'} X_8^{g28'}$$

$$\frac{dX2}{dt} = -\frac{dX8}{dt} = \alpha_3 {X_1}^{g31} X_8^{g38} - \alpha_4{}' {X_2}^{g42'} X_9^{g49'}$$

$$\frac{dX3}{dt} = -\frac{dX9}{dt} = \alpha_5 {X_2}^{g52} X_9^{g59} - \alpha_6{}' {X_3}^{g63'} X_{10}^{g610'}$$

$$\frac{dX4}{dt} = -\frac{dX10}{dt} = \alpha_7 {X_3}^{g73} X_{10}^{g710} - \alpha_8 {X_4}^{g84}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

$$\alpha_2' = \left.\frac{v_2 + v_3}{{X_1}^{g21'} X_6^{g26'} X_8^{g28'}}\right|_{Steady\ state};$$

$$v_2 = \alpha_2 {X_1}^{g21} X_6^{g26}\big|_{steady\ state}; v_3 = \alpha_3 {X_1}^{g31} X_8^{g38}\big|_{steady\ state};$$

$$g_{21}' = \frac{v_2 g_{21}}{v_2 + v_3};\ g_{26}' = \frac{v_2 g_{26}}{v_2 + v_3}; g_{28}' = \frac{v_3 g_{28}}{v_2 + v_3}$$

$$\alpha_4' == \left.\frac{v_4 + v_5}{{X_2}^{g42'} X_9^{g49'}}\right|_{Steady\ state};$$

$$v_4 = \alpha_4 {X_2}^{g42} X_9^{g49}\big|_{steady\ state}; v_5 = \alpha_5 {X_2}^{g52}\big|_{steady\ state};$$

$$g_{42}' = \frac{v_4 g_{42} + v_5 g_{52}}{v_4 + v_5};\ g_{49}' = \frac{v_4 g_{49}}{v_4 + v_5}$$

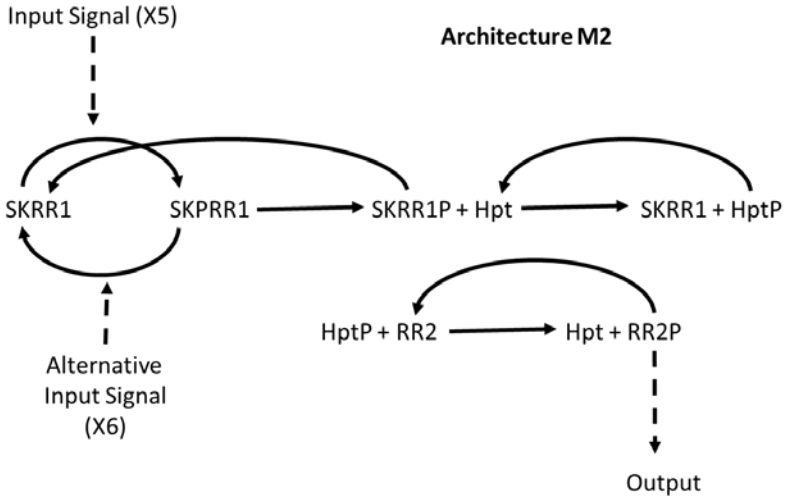
$$\alpha_6' = \left.\frac{v_6 + v_7}{{X_3}^{g63'} X_{10}^{g610'}}\right|_{Steady\ state};$$

$$v_6 = \alpha_6 {X_3}^{g63}\big|_{steady\ state}; v_7 = \alpha_7 {X_3}^{g73} X_{10}^{g710}\big|_{steady\ state};$$

$$g_{63}' = \frac{v_6 g_{63} + v_7 g_{73}}{v_6 + v_7};\ g_{610}' = \frac{v_7 g_{710}}{v_6 + v_7}$$

Architecture M2

The conceptual scheme we used to derive the power law mathematical models for architecture M2 is as follows



S-Systems mathematical model notation:
X1 – SKPRR1; X2 – SKRR1P; X3 – HptP; X4 – RR2P; X5 – Input Signal; X6 – Alternative Input Signal; X7 – SKRR1; X8 –Hpt; X9 – RR2
X1t – SKRR1 + SKPRR1 + SKRR1P; X2t – Hpt +HptP; X3t – RR2 + RR2P

GMA Model

$$\frac{dX1}{dt} = \alpha_1 {X_7}^{g17} X_5^{g15} - \alpha_2 {X_1}^{g21} X_6^{g26} - \alpha_3 {X_1}^{g31}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

$$\frac{dX2}{dt} = \alpha_3 {X_1}^{g31} - \alpha_4 {X_2}^{g42} X_8^{g48} - \alpha_5 {X_2}^{g52}$$

$$\frac{dX7}{dt} = -\frac{dX1}{dt} - \frac{dX2}{dt}$$

$$\frac{dX3}{dt} = -\frac{dX8}{dt} = \alpha_4 {X_2}^{g42} X_8^{g48} - \alpha_6 {X_3}^{g63} - \alpha_7 {X_3}^{g73} X_9^{g79}$$

$$\frac{dX4}{dt} = -\frac{dX9}{dt} = \alpha_7 {X_3}^{g73} X_9^{g79} - \alpha_8 {X_4}^{g84}$$

S-system Model

$$\frac{dX1}{dt} = -\frac{dX7}{dt} = \alpha_1 {X_7}^{g17} X_5^{g15} - \alpha_2' {X_1}^{g21'} X_6^{g26'}$$

$$\frac{dX2}{dt} = -\frac{dX8}{dt} = \alpha_3 {X_1}^{g31} - \alpha_4' {X_2}^{g42'} X_9^{g49'}$$

$$\frac{dX7}{dt} = -\frac{dX1}{dt} - \frac{dX2}{dt}$$

$$\frac{dX3}{dt} = -\frac{dX9}{dt} = \alpha_5 {X_2}^{g52} X_9^{g59} - \alpha_6' {X_3}^{g63'} X_{10}^{g610'}$$

$$\frac{dX4}{dt} = -\frac{dX10}{dt} = \alpha_7 {X_3}^{g73} X_{10}^{g710} - \alpha_8 {X_4}^{g84}$$

$$\alpha_2' = \left.\frac{v_2 + v_3}{{X_1}^{g21'} X_6^{g26'}}\right|_{Steady\ state};$$

$$v_2 = \alpha_2 {X_1}^{g21} X_6^{g26}\big|_{steady\ state}; v_3 = \alpha_3 {X_1}^{g31}\big|_{steady\ state};$$

$$g_{21}' = \frac{v_2 g_{21} + v_3 g_{31}}{v_2 + v_3};\ g_{26}' = \frac{v_2 g_{26}}{v_2 + v_3}$$

$$\alpha_4' = \left.\frac{v_4 + v_5}{{X_2}^{g42'} X_9^{g49'}}\right|_{Steady\ state};$$

$$v_4 = \alpha_4 {X_2}^{g42} X_9^{g49}\big|_{steady\ state}; v_5 = \alpha_5 {X_2}^{g52}\big|_{steady\ state};$$

$$g_{42}' = \frac{v_4 g_{42} + v_5 g_{52}}{v_4 + v_5};\ g_{49}' = \frac{v_4 g_{49}}{v_4 + v_5}$$

$$\alpha_6' = \left.\frac{v_6 + v_7}{{X_3}^{g63'} X_9^{g69'}}\right|_{Steady\ state};$$

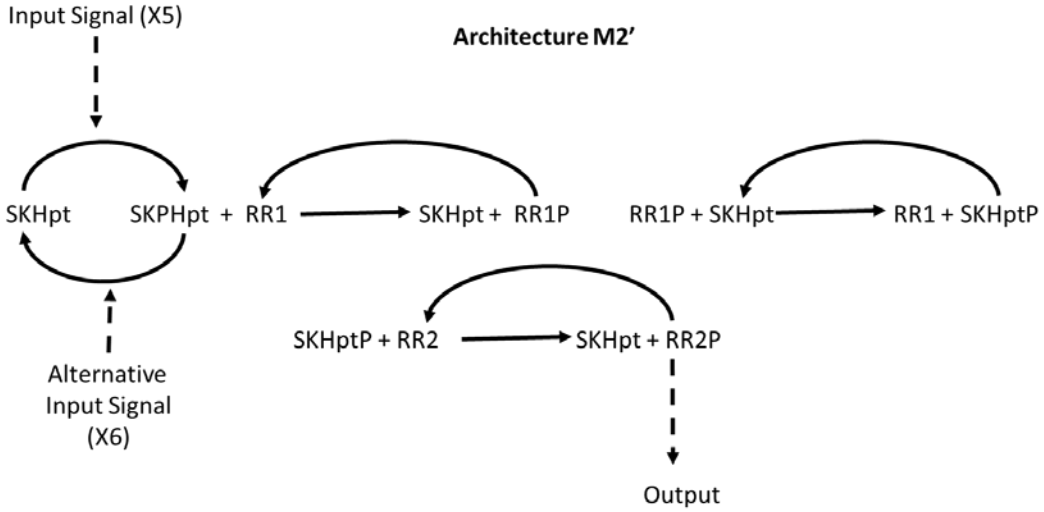
$$v_6 = \alpha_6 {X_3}^{g63}\big|_{steady\ state}; v_7 = \alpha_7 {X_3}^{g73} X_9^{g79}\big|_{steady\ state};$$

$$g_{63}' = \frac{v_6 g_{63} + v_7 g_{73}}{v_6 + v_7};\ g_{69}' = \frac{v_7 g_{79}}{v_6 + v_7}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Architecture M2'

The conceptual scheme we used to derive the power law mathematical models for architecture M2' is as follows



S-Systems mathematical model notation:
X1 – SKPHpt; X2 – RR1P; X3 – SKHptP; X4 – RR2P; X5 – Input Signal; X6 – Alternative Input Signal; X7 – SKHpt; X8 – RR1; X9 – RR2
X1t – SKHpt + SKPHpt + SKHptP; X2t – RR1 + RR1P; X3t – Hpt +HptP; X4t – RR2 + RR2P

GMA Model

$$\frac{dX1}{dt} = \alpha_1 {X_7}^{g17} X_5^{g15} - \alpha_2 {X_1}^{g21} X_6^{g26} - \alpha_3 {X_1}^{g31} X_8^{g38}$$

$$\frac{dX2}{dt} = -\frac{dX8}{dt} = \alpha_3 {X_1}^{g31} X_8^{g38} - \alpha_4 {X_2}^{g42} X_7^{g47} - \alpha_5 {X_2}^{g52}$$

$$\frac{dX3}{dt} = \alpha_4 {X_2}^{g42} X_7^{g47} - \alpha_6 {X_3}^{g63} - \alpha_7 {X_3}^{g73} X_9^{g79}$$

$$\frac{dX7}{dt} = -\frac{dX1}{dt} - \frac{dX3}{dt}$$

$$\frac{dX4}{dt} = -\frac{dX9}{dt} = \alpha_7 {X_3}^{g73} X_9^{g79} - \alpha_8 {X_4}^{g84}$$

S-system Model

$$\frac{dX1}{dt} = \alpha_1 {X_7}^{g17} X_5^{g15} - \alpha_2{}' {X_1}^{g21'} X_6^{g26'} X_8^{g28'}$$

$$\frac{dX2}{dt} = -\frac{dX8}{dt} = \alpha_3 {X_1}^{g31} X_8^{g38} - \alpha_4{}' {X_2}^{g42'} X_7^{g47'}$$

$$\frac{dX3}{dt} = \alpha_5 {X_2}^{g52} X_7^{g57} - \alpha_6{}' {X_3}^{g63'} X_9^{g69'}$$

$$\frac{dX7}{dt} = -\frac{dX1}{dt} - \frac{dX3}{dt}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

$$\frac{dX4}{dt} = -\frac{dX9}{dt} = \alpha_7 {X_3}^{g73} X_9^{g79} - \alpha_8 {X_4}^{g84}$$

$$\alpha_2' = \left.\frac{v_2 + v_3}{{X_1}^{g21'} X_6^{g26'} X_8^{g28'}}\right|_{Steady\ state};$$

$$v_2 = \alpha_2 {X_1}^{g21} X_6^{g26}\big|_{steady\ state}; v_3 = \alpha_3 {X_1}^{g31} X_8^{g38}\big|_{steady\ state};$$

$$g_{21}' = \frac{v_2 g_{21}}{v_2 + v_3};\ g_{26}' = \frac{v_2 g_{26}}{v_2 + v_3}; g_{28}' = \frac{v_3 g_{28}}{v_2 + v_3}$$

$$\alpha_4' = \left.\frac{v_4 + v_5}{{X_2}^{g42'} X_7^{g47'}}\right|_{Steady\ state};$$

$$v_4 = \alpha_4 {X_2}^{g42} X_7^{g47}\big|_{steady\ state}; v_5 = \alpha_5 {X_2}^{g52}\big|_{steady\ state};$$

$$g_{42}' = \frac{v_4 g_{42} + v_5 g_{52}}{v_4 + v_5};\ g_{47}' = \frac{v_4 g_{47}}{v_4 + v_5}$$

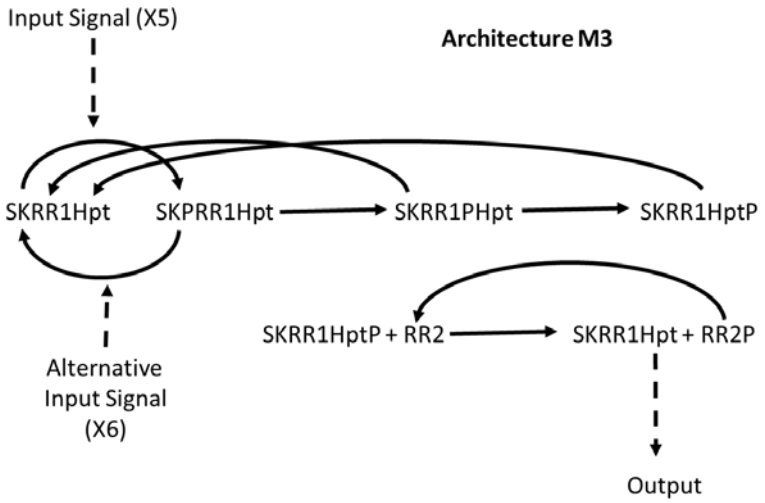
$$\alpha_6' = \left.\frac{v_6 + v_7}{{X_3}^{g63'} X_9^{g69'}}\right|_{Steady\ state};$$

$$v_6 = \alpha_6 {X_3}^{g63}\big|_{steady\ state}; v_7 = \alpha_7 {X_3}^{g73} X_9^{g79}\big|_{steady\ state};$$

$$g_{63}' = \frac{v_6 g_{63} + v_7 g_{73}}{v_6 + v_7};\ g_{69}' = \frac{v_7 g_{79}}{v_6 + v_7}$$

Architecture M3

The conceptual scheme we used to derive the power law mathematical models for architecture M3 is as follows



S-Systems mathematical model notation:
X1 – SKPRR1Hpt; X2 – SKRR1PHpt; X3 – SKRR1HptP; X4 – RR2P; X5 – Input Signal; X6 – Alternative Input Signal; X7 – SKRR1Hpt; X8 – RR2
X1t – SKRR1Hpt + SKPRR1Hpt + SKRR1PHpt + SKRR1HptP; X2t – RR2 + RR2P

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

GMA Model

$$\frac{dX1}{dt} = \alpha_1 {X_7}^{g17} X_5^{g15} - \alpha_2 {X_1}^{g21} X_6^{g26} - \alpha_3 {X_1}^{g31}$$

$$\frac{dX2}{dt} = \alpha_3 {X_1}^{g31} - \alpha_4 {X_2}^{g42} - \alpha_5 {X_2}^{g52}$$

$$\frac{dX3}{dt} = \alpha_4 {X_2}^{g42} - \alpha_6 {X_3}^{g63} - \alpha_7 {X_3}^{g73} X_8^{g78}$$

$$\frac{dX7}{dt} = -\frac{dX1}{dt} - \frac{dX2}{dt} - \frac{dX3}{dt}$$

$$\frac{dX4}{dt} = -\frac{dX8}{dt} = \alpha_7 {X_3}^{g73} X_8^{g78} - \alpha_8 {X_4}^{g84}$$

S-system Model

$$\frac{dX1}{dt} = \alpha_1 {X_7}^{g17} X_5^{g15} - \alpha_2{}' {X_1}^{g21'} X_6^{g26'}$$

$$\frac{dX2}{dt} = \alpha_3 {X_1}^{g31} - \alpha_4{}' {X_2}^{g42'}$$

$$\frac{dX3}{dt} = \alpha_5 {X_2}^{g52} - \alpha_6{}' {X_3}^{g63'} X_8^{g68'}$$

$$\frac{dX7}{dt} = -\frac{dX1}{dt} - \frac{dX2}{dt} - \frac{dX3}{dt}$$

$$\frac{dX4}{dt} = -\frac{dX8}{dt} = \alpha_7 {X_3}^{g73} X_8^{g78} - \alpha_8 {X_4}^{g84}$$

$$\alpha_2' = \left.\frac{v_2 + v_3}{{X_1}^{g21'} X_6^{g26'}}\right|_{Steady\ state};$$

$$v_2 = \alpha_2 {X_1}^{g21} X_6^{g26}\big|_{steady\ state}; v_3 = \alpha_3 {X_1}^{g31}\big|_{steady\ state};$$

$$g_{21}' = \frac{v_2 g_{21} + v_3 g_{31}}{v_2 + v_3};\ g_{26}' = \frac{v_2 g_{26}}{v_2 + v_3}$$

$$\alpha_4' = \left.\frac{v_4 + v_5}{{X_2}^{g42'}}\right|_{Steady\ state};$$

$$v_4 = \alpha_4 {X_2}^{g42}\big|_{steady\ state}; v_5 = \alpha_5 {X_2}^{g52}\big|_{steady\ state};$$

$$g_{42}' = \frac{v_4 g_{42} + v_5 g_{52}}{v_4 + v_5}$$

$$\alpha_6' = \left.\frac{v_6 + v_7}{{X_3}^{g63'} X_{10}^{g610'}}\right|_{Steady\ state};$$

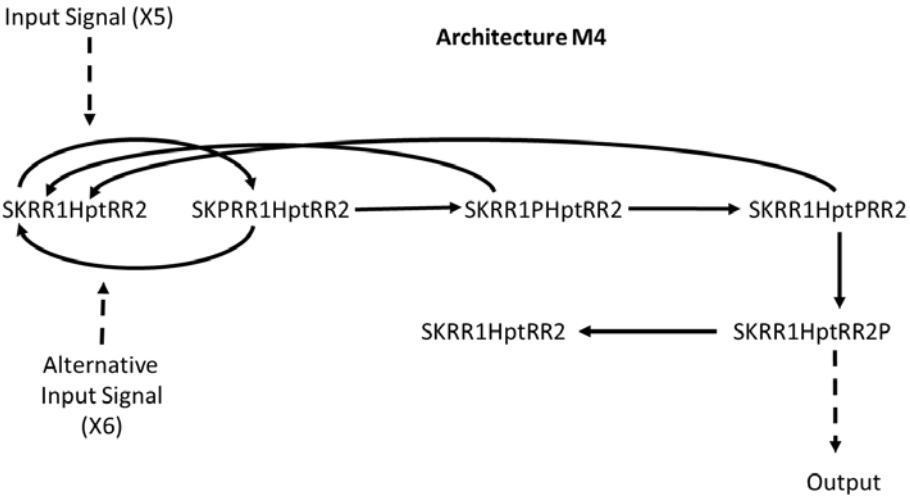
$$v_6 = \alpha_6 {X_3}^{g63}\big|_{steady\ state}; v_7 = \alpha_7 {X_3}^{g73} X_8^{g78}\big|_{steady\ state};$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

$$g'_{63} = \frac{v_6 g_{63} + v_7 g_{73}}{v_6 + v_7};\ g'_{68} = \frac{v_7 g_{78}}{v_6 + v_7}$$

Architecture M4

The conceptual scheme we used to derive the power law mathematical model for architecture M4 is as follows



S-Systems mathematical model notation:
X1 – SKPRR1HptRR2; X2 – SKRR1PHptRR2; X3 – SKRR1HptPRR2; X4 – SKRR1HptRR2P; X5 – Input Signal; X6 – Alternative Input Signal; X7 – SKRR1HptRR2
X1t – SKRR1HptRR2 + SKPRR1HptRR2 + SKRR1PHptRR2 + SKRR1HptPRR2 + SKRR1HptRR2P

GMA Model

$$\frac{dX1}{dt} = \alpha_1 {X_7}^{g17} X_5^{g15} - \alpha_2 {X_1}^{g21} X_6^{g26} - \alpha_3 {X_1}^{g31}$$

$$\frac{dX2}{dt} = \alpha_3 {X_1}^{g31} - \alpha_4 {X_2}^{g42} - \alpha_5 {X_2}^{g52}$$

$$\frac{dX3}{dt} = \alpha_4 {X_2}^{g42} - \alpha_6 {X_3}^{g63} - \alpha_7 {X_3}^{g73}$$

$$\frac{dX4}{dt} = \alpha_7 {X_3}^{g73} - \alpha_8 {X_4}^{g84}$$

$$\frac{dX7}{dt} = -\frac{dX1}{dt} - \frac{dX2}{dt} - \frac{dX3}{dt} - \frac{dX4}{dt}$$

S-system Model

$$\frac{dX1}{dt} = \alpha_1 {X_7}^{g17} X_5^{g15} - \alpha_2{}' {X_1}^{g21\prime} X_6^{g26\prime}$$

$$\frac{dX2}{dt} = \alpha_3 {X_1}^{g31} - \alpha_4{}' {X_2}^{g42\prime}$$

$$\frac{dX3}{dt} = \alpha_5 {X_2}^{g52} - \alpha_6{}' {X_3}^{g63\prime}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

$$\frac{dX4}{dt} = \alpha_7 {X_3}^{g73} - \alpha_8 {X_4}^{g84}$$

$$\frac{dX7}{dt} = -\frac{dX1}{dt} - \frac{dX2}{dt} - \frac{dX3}{dt} - \frac{dX4}{dt}$$

$$\alpha_2' = \left.\frac{v_2 + v_3}{{X_1}^{g21'} X_6^{g26'}}\right|_{Steady\ state} ;$$

$$v_2 = \alpha_2 {X_1}^{g21} X_6^{g26}\big|_{steady\ state}; v_3 = \alpha_3 {X_1}^{g31}\big|_{steady\ state} ;$$

$$g_{21}' = \frac{v_2 g_{21} + v_3 g_{31}}{v_2 + v_3};\ g_{26}' = \frac{v_2 g_{26}}{v_2 + v_3}$$

$$\alpha_4' = \left.\frac{v_4 + v_5}{{X_2}^{g42'}}\right|_{Steady\ state} ;$$

$$v_4 = \alpha_4 {X_2}^{g42}\big|_{steady\ state}; v_5 = \alpha_5 {X_2}^{g52}\big|_{steady\ state} ;$$

$$g_{42}' = \frac{v_4 g_{42} + v_5 g_{52}}{v_4 + v_5}$$

$$\alpha_6' = \left.\frac{v_6 + v_7}{{X_3}^{g63'}}\right|_{Steady\ state} ;$$

$$v_6 = \alpha_6 {X_3}^{g63}\big|_{steady\ state}; v_7 = \alpha_7 {X_3}^{g73}\big|_{steady\ state} ;$$

$$g_{63}' = \frac{v_6 g_{63} + v_7 g_{73}}{v_6 + v_7}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Conceptual schemes and mass action mathematical models

Architecture M1

Reaction	Rate
SK → SKP	$k_1 \times SK$
SKP → SK	$k_2 \times SKP$
SKP + RR1 → SKPRR1	$k_3 \times SKP \times RR1$
SKPRR1 → SKP + RR1	$k_4 \times SKPRR1$
SKPRR1 → SKRR1P	$k_5 \times SKPRR1$
SKRR1P → SK + RR1P	$k_6 \times SKRR1P$
SK + RR1P → SKRR1P	$k_7 \times SK \times RR1P$
SKPRR1 → SKRR1	$k_8 \times SKPRR1$
SKRR1 → SK + RR1	$k_9 \times SKRR1$
SK + RR1 → SKRR1	$k_{10} \times SK \times RR1$
RR1P → RR1	$K_{11} \times RR1P$
RR1P + Hpt →RR1PHpt	$k_{12} \times Hpt \times RR1P$
RR1PHpt → RR1P+Hpt	$k_{13} \times RR1PHpt$
RR1PHpt → RR1HptP	$k_{14} \times RR1PHpt$
RR1HptP → RR1 + HptP	$k_{15} \times RR1HptP$
RR1 + HptP → RR1HptP	$k_{16} \times RR1 \times HptP$
HptP → Hpt	$k_{17} \times HptP$
HptP + RR2 → HptPRR2	$k_{18} \times HptP \times RR2$
HptPRR2 → HptP + RR2	$k_{19} \times HptPRR2$
HptPRR2 → HptRR2P	$k_{20} \times HptPRR2$
HptRR2P → Hpt + RR2P	$k_{21} \times HptRR2P$
Hpt + RR2P → HptRR2P	$k_{22} \times Hpt \times RR2P$
RR2P → RR2	$k_{23} \times RR2P$

Architecture M2

Reaction	Rate
SKRR1 → SKPRR1	$k_1 \times SKRR1$
SKPRR1 → SKRR1	$k_2 \times SKPRR1$
SKPRR1 → SKRR1P	$k_5 \times SKPRR1$
SKPRR1 → SKRR1	$k_8 \times SKPRR1$
SKRR1P → SKRR1	$K_{11} \times SKRR1P$
SKRR1P + Hpt →SKRR1PHpt	$k_{12} \times Hpt \times SKRR1P$
SKRR1PHpt → SKRR1P+Hpt	$k_{13} \times SKRR1PHpt$
SKRR1PHpt → SKRR1HptP	$k_{14} \times SKRR1PHpt$
SKRR1HptP → SKRR1 + HptP	$k_{15} \times SKRR1HptP$
SKRR1 + HptP → SKRR1HptP	$k_{16} \times SKRR1 \times HptP$
HptP → Hpt	$k_{17} \times HptP$
HptP + RR2 → HptPRR2	$k_{18} \times HptP \times RR2$
HptPRR2 → HptP + RR2	$k_{19} \times HptPRR2$
HptPRR2 → HptRR2P	$k_{20} \times HptPRR2$
HptRR2P → Hpt + RR2P	$k_{21} \times HptRR2P$
Hpt + RR2P → HptRR2P	$k_{22} \times Hpt \times RR2P$
RR2P → RR2	$k_{23} \times RR2P$

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Architecture M2'

Reaction	Rate
SKHpt → SKPHpt	$k_1 \times SKHpt$
SKPHpt → SKHpt	$k_2 \times SKPHpt$
SKPHpt + RR1 → SKPHptRR1	$k_3 \times SKPHpt \times RR1$
SKPHptRR1 → SKPHpt + RR1	$k_4 \times SKPHptRR1$
SKPHptRR1 → SKHptRR1P	$k_5 \times SKPHptRR1$
SKHptRR1P → SKHpt + RR1P	$k_6 \times SKHptRR1P$
SKHpt + RR1P → SKHptRR1P	$k_7 \times SKHpt \times RR1P$
SKPHptRR1 → SKHptRR1	$k_8 \times SKPHptRR1$
SKHptRR1 → SKHpt + RR1	$k_9 \times SKHptRR1$
SKHpt + RR1 → SKHptRR1	$k_{10} \times SKHpt \times RR1$
RR1P → RR1	$K_{11} \times RR1P$
RR1P + SKHpt →RR1PSKHpt	$k_{12} \times SKHpt \times RR1P$
RR1PSKHpt → RR1P+SKHpt	$k_{13} \times RR1PSKHpt$
RR1PSKHpt → RR1SKHptP	$k_{14} \times RR1PSKHpt$
RR1SKHptP → RR1 + SKHptP	$k_{15} \times RR1SKHptP$
RR1 + SKHptP → RR1SKHptP	$k_{16} \times RR1 \times SKHptP$
SKHptP → SKHpt	$k_{17} \times SKHptP$
SKHptP + RR2 → SKHptPRR2	$k_{18} \times SKHptP \times RR2$
SKHptPRR2 → SKHptP + RR2	$k_{19} \times SKHptPRR2$
SKHptPRR2 → SKHptRR2P	$k_{20} \times SKHptPRR2$
SKHptRR2P → SKHpt + RR2P	$k_{21} \times SKHptRR2P$
SKHpt + RR2P → SKHptRR2P	$k_{22} \times SKHpt \times RR2P$
RR2P → RR2	$k_{23} \times RR2P$

Architecture M3

Reaction	Rate
SKRR1Hpt → SKPRR1Hpt	$k_1 \times SKRR1Hpt$
SKPRR1Hpt → SKRR1Hpt	$k_2 \times SKPRR1Hpt$
SKPRR1Hpt → SKRR1PHpt	$k_5 \times SKPRR1Hpt$
SKPRR1Hpt → SKRR1Hpt	$k_8 \times SKPRR1Hpt$
SKRR1PHpt → SKRR1Hpt	$K_{11} \times SKRR1PHpt$
SKRR1PHpt → SKRR1HptP	$k_{14} \times SKRR1PHpt$
SKRR1HptP → SKRR1Hpt	$k_{17} \times SKRR1HptP$
SKRR1HptP + RR2 → SKRR1HptPRR2	$k_{18} \times SKRR1HptP \times RR2$
SKRR1HptPRR2 → SKRR1HptP + RR2	$k_{19} \times SKRR1HptPRR2$
SKRR1HptPRR2 → SKRR1HptRR2P	$k_{20} \times SKRR1HptPRR2$
SKRR1HptRR2P → SKRR1Hpt + RR2P	$k_{21} \times SKRR1HptRR2P$
SKRR1Hpt + RR2P → SKRR1HptRR2P	$k_{22} \times SKRR1Hpt \times RR2P$
RR2P → RR2	$k_{23} \times RR2P$

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Architecture M4

Reaction	Rate
SKRR1HptRR2 → SKPRR1HptRR2	k_1 × SKRR1HptRR2
SKPRR1HptRR2 → SKRR1HptRR2	k_2 × SKPRR1HptRR2
SKPRR1HptRR2 → SKRR1PHptRR2	k_5 × SKPRR1HptRR2
SKPRR1HptRR2 → SKRR1HptRR2	k_8 × SKPRR1HptRR2
SKRR1PHptRR2 → SKRR1HptRR2	K_{11} × SKRR1PHptRR2
SKRR1PHptRR2 → SKRR1HptPRR2	k_{14} × SKRR1PHptRR2
SKRR1HptPRR2 → SKRR1HptRR2	k_{17} × SKRR1HptPRR2
SKRR1HptPRR2 → SKRR1HptRR2P	k_{20} × SKRR1HptPRR2
SKRR1HptRR2P → SKRR1HptRR2	k_{23} × SKRR1HptRR2P

Parameter values and protein abundances for the mass action mathematical models

The parameter values for the mass action models were generated by combining all possible values described in Table S4. The protein abundances were approximated to the orders of magnitudes experimentally determined in Table S3 and then also combined using a latin hypercube approach.

Mathematical Models for the Spo0 Phosphorelay

Parameters and abundances for the Spo0 system in *Bacillus subtilis* compiled from [56,99]

Reaction	Rate constants
SK → SKP	0.0019 s^{-1}
SKP → SK	0
SK+RR → SKRR	100
SKRR→SK + RR	200
SKP+RR → SKPRR	200
SKPRR→SKP + RR	100
SKPRR→SKRR	0
SKPRR → SKRRP	0.08 s^{-1}
SKRRP → SKRR	0
SKRRP → SK + RRP	100
SK + RRP → SKRRP	200
RRP → RR	0.05
RRP + Hpt → RRPHpt	100
RRPHpt → RRP + Hpt	2300
RRPHpt →RRHptP	$1.5*10^{-5}$ s^{-1}
RRHptP → RR + HptP	2300
RR + HptP → RRHptP	100
HptP + RR2 → HptPRR2	400
HptPRR2 → HPtP + RR2	100
HptPRR2 → HptRR2P	0.20 s^{-1}
HptRR2P → Hpt + RR2P	100
Hpt + RR2P → HptRR2P	400
RR2P → RR2	0.0001-0.006

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.21.108001; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

PR proteins	Abundances
KinA	610 molecules cell $^{-1}$
Spo0F	123 molecules cell $^{-1}$
Spo0A	192 molecules cell $^{-1}$
Spo0B	267 molecules cell $^{-1}$

Mathematical Models for the Sln1 Phosphorelay

Parameter values and abundances for the Sln1-Ypd1-Ssk1-Skn7 system in *Sacharomyces cerevisiae* compiled from [30,31]

Reaction	Sln1:Ypd1:Ssk1	Sln1:Ypd1:Skn7
SKRR → SKPRR	10000	10000
SKPRR → SKRR	100	100
SKPRR → SKRRP	10000	10000
SKRRP→SKRR	0.001 (0.0005 w/osmolites)	0.001
SKRRP + Hpt → SKRRPHpt	2000 M^{-1} s^{-1}	2000 M^{-1} s^{-1}
SKRRPHpt → SKRRP + Hpt	2800 s^{-1}	2800 s^{-1}
SKRRPHpt →SKRRHptP	29 s^{-1}	29 s^{-1}
SKRRHptP →SKRRPHpt	7.5 s^{-1}	7.5 s^{-1}
SKRRHptP → SKRR + HptP	14000	14000
SKRR + HptP → RRHptP	2000	2000
HptP + RR2 → HptPRR2	2000	2000
HptPRR2 → HPtP + RR2	4800	2800
HptPRR2 → HptRR2P	160 s^{-1}	1.4 s^{-1}
HptRR2P → HptPRR2	0	0.4 s^{-1}
HptRR2P → Hpt + RR2P	0	10000
Hpt + RR2P → HptRR2P	0	2

PR Proteins	Protein abundances
Sln1	696 molecules cell $^{-1}$
Ypd1	6560 molecules cell $^{-1}$
Ssk1	1018 molecules cell $^{-1}$
Skn7	2536 molecules cell $^{-1}$