npj | systems biology and applications

Article

Published in partnership with the Systems Biology Institute

https://doi.org/10.1038/s41540-024-00430-9

Optimal performance objectives in the highly conserved bone morphogenetic protein signaling pathway

Razeen Shaikh [1], Nissa J. Larson [2], Jayden Kam [1], Donny Hanjaya-Putra[3,4], Jeremiah Zartman [4,5], David M. Umulis[2] ✉, Linlin Li[2] ✉ & Gregory T. Reeves [1,6] ✉

Throughout development, complex networks of cell signaling pathways drive cellular decision-making across different tissues and contexts. The transforming growth factor β (TGF-β) pathways, including the BMP/Smad pathway, play crucial roles in determining cellular responses. However, as the Smad pathway is used reiteratively throughout the life cycle of all animals, its systems-level behavior varies from one context to another, despite the pathway connectivity remaining nearly constant. For instance, some cellular systems require a rapid response, while others require high noise filtering. In this paper, we examine how the BMP-Smad pathway balances trade-offs among three such systems-level behaviors, or "Performance Objectives (POs)": response speed, noise amplification, and the sensitivity of pathway output to receptor input. Using a Smad pathway model fit to human cell data, we show that varying non-conserved parameters (NCPs) such as protein concentrations, the Smad pathway can be tuned to emphasize any of the three POs and that the concentration of nuclear phosphatase has the greatest effect on tuning the POs. However, due to competition among the POs, the pathway cannot simultaneously optimize all three, but at best must balance trade-offs among the POs. We applied the multi-objective optimization concept of the Pareto Front, a widely used concept in economics to identify optimal trade-offs among various requirements. We show that the BMP pathway efficiently balances competing POs across species and is largely Pareto optimal. Our findings reveal that varying the concentration of NCPs allows the Smad signaling pathway to generate a diverse range of POs. This insight identifies how signaling pathways can be optimally tuned for each context.

Throughout the process of development, an array of essential communication pathways exert influence over a wide spectrum of cellular destiny determinations and various processes in different tissues and situations, including the Bone Morphogenetic Protein (BMP) pathway, a member of the Transforming Growth Factor β (TGF-β) superfamily of signaling pathways. The BMP pathway regulates a wide variety of cellular responses, including apoptosis, differentiation, homeostasis, stem cell maintenance, and regeneration in animals from flies to humans[1–7]. The BMP module is functionally conserved to a point that human components can replace their respective homologs in *Drosophila*[8–10]. The pathway is activated by ligands of the BMP family, which bind to cognate Type I serine-threonine kinase receptors, promoting the recruitment of the Type II receptors (Fig. 1A). The Type I/Type II receptor complex phosphorylates Smad1/5/8, which then dimerizes and forms a complex with the Co-Smad (Smad4), forming the signaling complex $(PSmad1)_2$/Smad4. This signaling complex (SC) enters the nucleus and activates downstream gene expression[1,11]. The first Smads were discovered in *Drosophila*: Mothers against Dpp (Mad; homolog of Smad1/5/8) and Medea (Med; homologof Smad4)[12].

Despite its high degree of conservation across different species, the BMP/Smad pathway exhibits remarkable diversity in its responses to BMPs in different contexts. For example, the response speed of the Smad pathway varies widely across common model systems (Fig. 1B-H), including the well-studied blastoderm stage of the *Drosophila* embryo (BMP signaling timescale of 30 min[13–16]), *Drosophila* pupal wing formation (timescale of

[1]Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, Texas, TX, USA. [2]Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA. [3]Aerospace and Mechanical Engineering, University of Notre Dame, Notre Dame, IN, USA. [4]Bioengineering Graduate Program, University of Notre Dame, Notre Dame, IN, USA. [5]Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, IN, USA. [6]Faculty of Genetics and Genomics, Texas A&M University, College Station, TX, USA. ✉e-mail: dumulis@purdue.edu; li2212@purdue.edu; gtreeves@tamu.edu

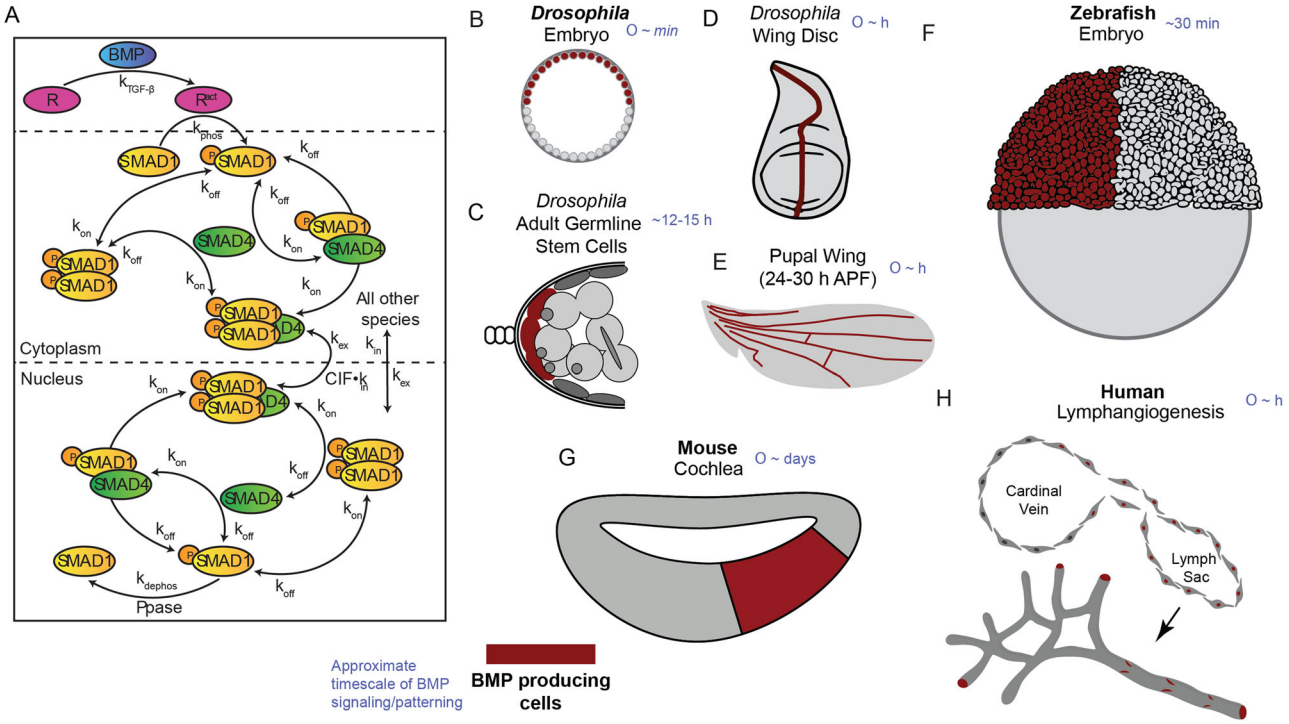


Fig. 1 | Smad signaling is a conserved pathway. A Reaction network diagram of the Smad pathway. TGF-β binds and activates the receptors, thus phosphorylating Smad1. PSmad1 binds to Smad4 and forms dimers which then form trimers. Illustrative figures of model systems in which BMP signaling pathway patterns tissues with BMP producing cells marked in maroon and an approximate timescale over which BMP signaling either is active or patterns the tissue is indicated in blue, (**B**) *Drosophila* Embryo, (**C**) *Drosophila* adult germline stem cells, (**D**) *Drosophila* Wing Disc, (**E**) *Drosophila* Pupal Wing (24–30 h APF), (**F**) Zebrafish Embryo, (**G**) Mouse Cochlea, and (**H**) Human Lymphangiogenesis. (created using Adobe Illustrator).

hours[17]), *Drosophila* germline stem cells (BMP signaling timescale ~12–16 h[18]), the zebrafish embryo (30 min[19,20]), the mammalian cochlea (order of days[21]), and human induced pluripotent stem cells (hiPSCs; response speed on the scale of hours[22–25]).

These differences in response speed highlight the fact that the BMP-Smad pathway exhibits a diversity of behavior depending on the biological context. It is likely that the response speed is in a trade-off relationship with other systems-level behaviors, or "Performance Objectives (POs)"[26,27]. To first order, a minimal set of such POs includes (1) response time, (2) noise filtering, and (3) linear sensitivity of output with respect to extracellular inputs[1–5]. While there are other POs that likely affect the fitness of the organism, such as signal amplification, precision, or information capacity[2,3,6], our three proposed POs are based on fundamental constraints for all communication systems and are shared across the animal kingdom[7–9]. However, due to competition among these POs, a "utopian system" in which all three POs are optimally and simultaneously satisfied is not biologically or physically possible, and compromises must be made to emphasize the proper POs for each biological context. For example, communication systems can sacrifice response speed to reduce noise through time-averaging[28,29]. To simultaneously improve noise filtering and response speed, signaling pathways could operate in the regime of high levels of activated receptor inputs. However, in this regime, linear sensitivity to input levels is sacrificed, as the network would become saturated[28,29]. Therefore, certain cellular systems might demand swift responses to BMP signaling, like *Drosophila* and zebrafish embryos, whereas others may necessitate extensive noise filtration, like human stem cells. This diversity in the behavior of the BMP module cannot derive from rewiring the pathway, or from altering protein biochemical function, as these aspects of the BMP pathway are highly conserved across the animal kingdom[8–11,30]. Instead, we hypothesize the tunability is achieved through differential concentrations of the pathway components, such as Smad1 and Smad4, whose sequences and functions are conserved, but whose concentrations are not.

In this study, we used a Smad pathway model, calibrated to time course data from human cells[31], to analyze how the Smad pathway balances trade-offs among the three POs through variation of non-conserved parameters (NCPs): concentrations of phosphatase and Smad proteins, and the nuclear import rate of activated Smad protein complex. We found that the NCPs allow the Smad signaling pathway to generate a diverse range of POs, and that phosphatase concentration has the largest effect on shifting the balance of POs. To systematically determine which combinations of POs are optimal, we utilized the concept of the Pareto Front, which has been widely employed in economics for multi-objective optimization to identify a collection of designs that offer optimal trade-offs among various requirements[32,33]. The shape of the Pareto front supplies information about the relationship of the objectives. If the Pareto front curves towards the "utopian point" (UP), where all objectives are maximally satisfied, then adequate compromise solutions exist[34]. On the other hand, a concave shape that curves away from the UP suggests that the objectives severely exclude each other, and compromises will be difficult. In the field of biology, the Pareto front approach allows for a comprehensive analysis of the diverse factors influencing biological systems and aids in the identification of solutions that achieve the best compromise between competing objectives. We found that, for most combinations of NCP values, the Smad pathway is Pareto optimal. We analyzed which systems fall on the Pareto front and predicted the relationship between values of NCPs and which POs a given system would emphasize. We conclude that the Smad pathway is highly versatile to achieve a variety of trade-offs among the POs, depending on the needs of each system in their biological context.

Results

Phosphatase, a non-conserved parameter, determines the performance objectives of the system

Our BMP/Smad pathway model was adapted from a similar model constructed for the Smad pathway downstream of the TGF-β pathway[31]. We

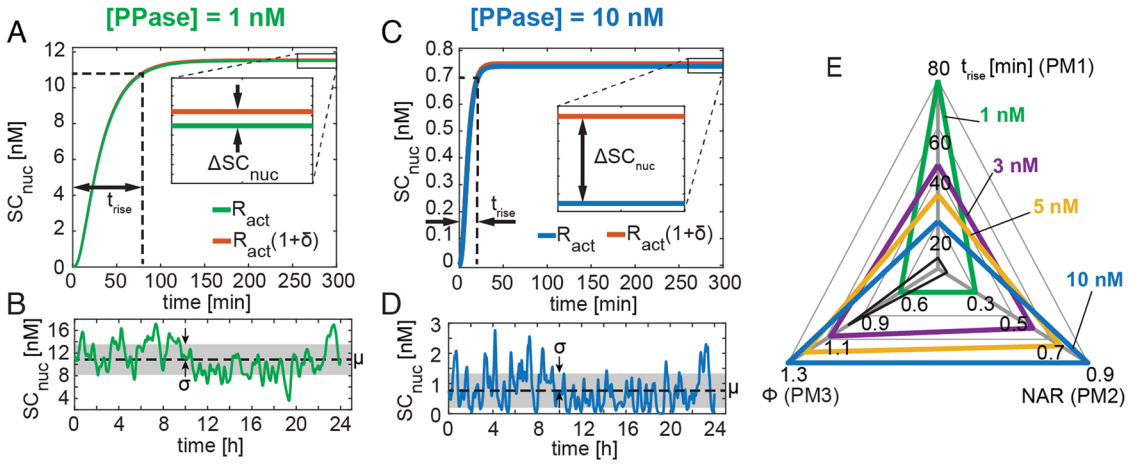


Fig. 2 | The Performance Metrics (PMs) of the Smad signaling pathway depend on the PPase concentration. The simulated concentration profiles for the signaling complex in the nucleus, SC_{nuc}, at a constant activated receptor level and PPase concentration of (**A**) 1 nM (**C**) 10 nM (inset) The differential change in steady state concentration with respect to change in R_{act}. In response to the noise input (S5A) the resulting dynamic profile for SC_{nuc} has a small coefficient of variation at PPase concentration of (**B**) 1 nM (**D**) 10 nM (black dashed line indicates mean and grey rectangle indicates standard deviation. **E** A spider plot of the three PMs: t_{rise} (PM 1), NAR (PM 2), and sensitivity coefficient (φ) (PM 3) at increasing levels of PPase. The "utopian ideal", in which all three PMs are optimized, is shown in black.

used a core set of parameters that were fit to previously published TGF-β/Smad pathway data[31]. With the core parameters in place, we performed simulations of the BMP/Smad pathway (hereafter referred to simply as the Smad pathway) with the pathway initially at rest (active receptor level, $R_{act} = 0$). At time $t = 0$, we increased the active receptor level to its new value, R_{act}, which resulted in an increase in the nuclear concentration of the signaling complex (SC), $(PSmad1)_2$/Smad4. Based on this simulation, we computed the rise time (t_{rise}) (PM1), which serves as an indicator of the system's response speed (PO1). The t_{rise} is defined as the duration it takes for the response variable to reach 95% of the final steady state (Fig. 2A).

To evaluate the noise amplification properties (PO2) of the Smad pathway, we extended the simulation, beginning at the established steady state, but with a time-varying noise R_{act} (Supplementary Fig. 5A, See Supplementary Note 4.1). The variations in R_{act} mimic stochastic simulations of receptor activation[35]. The simulation was continued over a simulation time of 24 hours. During that time, the response variable fluctuated about a mean of μ with a standard deviation of σ (Fig. 2B). Using these simulated fluctuations in the response variable, we calculated the second metric, Noise Attenuation Ratio (NAR, PM2)[36], which tracks the dynamic properties of the response and is defined as the ratio of the coefficient of variation ($CV \equiv \frac{\sigma}{\mu}$) of the response variable to that of the input variable (R_{act}). The final metric, the steady state sensitivity coefficient (ϕ) (PM3), is a measure of the linear fidelity (PO3) of the response variable with respect to the input variable. To calculate this metric, we ran a parallel simulation by perturbing R_{act} by 1%, resulting in a slightly different steady state for the response variable (inset of Fig. 2A). The sensitivity coefficient is defined as the ratio of the fractional change in steady state output to that of the input.

Because phosphatase (PPase) levels often control the dynamics of signaling systems[37–41], we evaluated the three performance objectives (POs) by calculating their respective performance metrics (PMs) at varying levels of PPase concentration (Fig. 2E). We found that increasing the PPase concentration from 1 nM to 10 nM, keeping all other core parameters and initial concentrations unchanged, resulted in a decrease in the rise time (Fig. 2A, C), an increase in NAR (Fig. 2B, D), and an increase in the sensitivity coefficient (inset of Fig. 2A, C). These trends were also maintained for intermediate values of the phosphatase concentration (Fig. 2E). By comparison, the utopian ideal has $t_{rise} = 0$, NAR = 0, and ϕ = 1 (black triangle in Fig. 2E), which according to our mathematical model, cannot be attained and may not biologically or physically possible. In summary, our results indicate that the PPase concentration affects the performance measures (PMs). Therefore, by regulating PPase concentration, the Smad signaling pathway can be fine-tuned to achieve specific desired outcomes.

The effect of non-conserved parameters on the performance objectives

While the Smad pathway topology and component protein sequence are highly conserved across the animal kingdom, we consider four parameters in the model to be "non conserved parameters (NCPs)". The first NCP is the concentration of the PPase (see above). While we do not directly identify the PPase, we assume it is one of the factors that control the pseudo-first order rate constant for PSmad dephosphorylation. The next two NCPs are the concentrations of Smad1, and Smad4, whose sequence and function are conserved, but whose concentrations are not. The fourth NCP in the model is the complex-import factor (CIF), which is defined as the nuclear import rate constant of the SC normalized to the nuclear import rate of the other Smad-containing complexes[31] and is likely controlled by the concentration of specific nuclear transport proteins. Thus, the NCPs are concentrations that either enter the equations as initial conditions (such as for Smad1 and Smad4) or control kinetics (such as for PPase and CIF).

To determine how each of these NCPs affects the three PMs, we performed a parameter screen and calculated the three PMs for 10^4 sets of randomly generated sets of NCPs (see Methods) spanning several orders of magnitude. This extensive analysis allowed us to generate a three-dimensional manifold of all feasible PMs in "PM-space" (Fig. 3A), which revealed that no parameter set resulted in optimal values of all three PMs simultaneously (designated as the utopian point, UP). The projections of the three-dimensional manifold indicate that the PM space is constrained by the model formulation (Fig. 3B–D). The shape of the manifold indicates that there is a large infeasible space, and the Smad pathway has to balance performance trade-offs to attain the desired performance. Notably, these POs do not appear independent but instead are in trade-off relationships, such as between response speed vs. noise amplification (Fig. 3B), and noise amplification vs. sensitivity coefficient (Fig. 3C).

To visualize the effect of each NCP on determining the PMs of the signaling network, we sorted the points from the parameter screen into bins based on the value of each NCP (Supplementary Fig. 12). However, given the high density of points, we replaced each cluster of points by its centroid (see Methods). We found that PPase concentration has the largest effect on the balance of PMs, while CIF has the smallest (Fig. 4A, B). Smad1, Smad4, and their ratio each had a non-monotonic effect on the balance of PMs (Fig. 4C–E).

To determine the combination of NCPs for which the system is "optimal," we calculated the Euclidean distance of each point using equal weights, hereafter referred to as the Equal Weights Distance in the curves plotted in Fig. 4A-E to the Utopian Point (UP, Fig. 4A). The utopian points

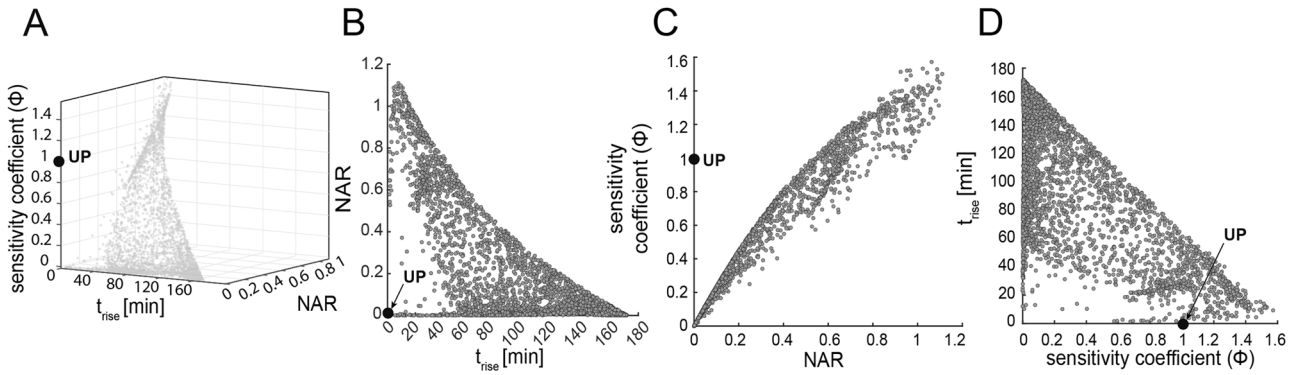


Fig. 3 | The results of the parameter screen in PM-space. A 3D surface representation of the results of the front of the parameter screen in PM-space. Trade-offs are evident as no parameter set results in PMs that approach the utopian point (UP). The rise time (t_{rise})-NAR (**B**), NAR-sensitivity coefficient (φ) (**C**), and sensitivity coefficient (ϕ) – rise time (t_{rise}) (**D**) projection of all PMs generated by sampling NCPs.

are defined as the "ideal" optimal performance point with minimum rise time ($t_{rise} = 0$) and noise ratio (NAR = 0), and optimal sensitivity coefficient ($\phi = 1$). While sensitivity coefficient and NAR are unitless ratios of order one, the rise time is not unitless and can reach values of a few hundred minutes. To account for this difference in scales, we scaled the rise time by the maximum calculated value over all ten clusters, then arbitrarily weighted the three PMs equally (see "Methods"). We found there is an optimal level of PPase that minimizes the distance to the utopian point (Fig. 4F). The centroids of CIF clusters are largely equidistant from the utopian point, except at higher concentrations (clusters 8–10) (Fig. 4F). The Smad1 clusters advance towards the utopian point with increasing Smad1 concentration but stagnate such that any further increase in concentration does not significantly affect the position of the clusters (Fig. 4F). In accord with the "U" shape formed by the curves for Smad4 and the Smad1/Smad4 ratio, there are clear optimal values for these NCPs that minimize the distance to the utopian point (Fig. 4F).

The optimality of performance objectives at different receptor levels

Up to this point, all the simulations described previously were done at a single value of R_{act} (Supplementary Fig. 5A). However, R_{act} levels are expected to vary across species, systems, and space. We generated a total of 60 receptor inputs of varying mean R_{act} and frequency distribution (Supplementary Figs. 6–9) (see "Methods") and calculated the PMs at the 10^4 previously generated NCPs sets. Since R_{act} is based on an extracellular component and can vary across cells in a tissue exposed to a gradient of BMP signaling, we treated it as an input to the model, rather than a model parameter, and varied it separately from the other NCPs.

For PPase, at higher R_{act} levels the centroid of point clusters advances closer to the utopian point and stagnates (Supplementary Figs. 2A, 5A). The centroid of point clusters of CIF have the same nature at all R_{act}, but the distance from the utopian point progressively increases on increasing R_{act} (Supplementary Figs. 2B, 5B). For all R_{act}, increasing total Smad1 shifts the system closer to the utopian point (Supplementary Figs. 2C, 5C) whereas increasing total Smad4 shifts the system further away from the utopian point (Supplementary Figs. 2D, 5D). To balance the competing nature of Smad1 and Smad4, the ratio of $Smad1_{total}/Smad4_{total}$ must be optimally controlled (Supplementary Figs. 2E, 5E). (Fig. 5 summarizes "R_{max}" levels, for "R_{min}" see Supplementary Figs. 1, 3).

The Pareto optimality of the performance objectives

Minimizing the distance of the curves in Fig. 4A–E to the optimal point was based on an equal weights distance metric in which the values of t_{rise} (normalized), NAR, and the sensitivity coefficient were given equal numerical weight. However, as the metrics for the three PMs are not related to each other numerically, there is no inherent justification for equal weighting. To address this concern, we have taken a multi-objective optimization (MOO) approach to determine which points in PM space can be considered optimal. Unlike standard, single-objective optimization, MOO attempts to satisfy multiple, conflicting objectives, and therefore, MOO is the proper method to apply to balance the PMs[42].

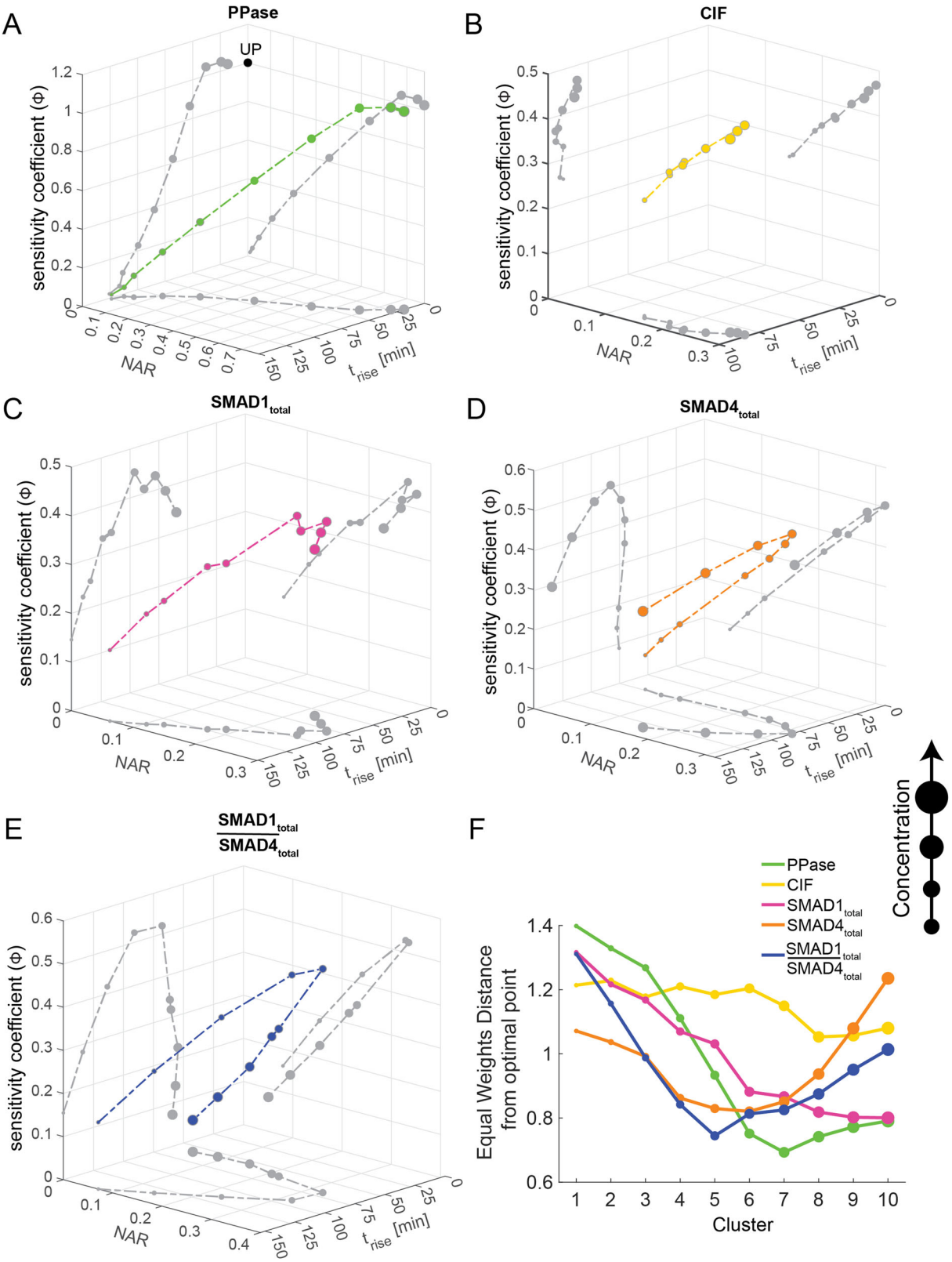
In general, in a MOO approach, one identifies the subset of model results in which improvements to one objective cannot be made without compromising the others. This subset is known as the "Pareto Front," and each of the points on the Pareto Front is considered optimal, from a Pareto standpoint. In practice, points on the Pareto Front are identified as those for which no other point is better in all three objectives. Applying this method to our screen, we found that roughly 70% of results from our screen resided on the Pareto front (Fig. 6). If the optimality criterion is relaxed to allow solutions which are within 10% of the distance from other optimal solutions to also reside on the Pareto front, then ≥99% of all solutions lie on the Pareto front (Supplementary Fig. 14). Even some points that perform very poorly on one PM may still considered Pareto optimal because the trade-offs are such that the other PMs are close to their respective optimal values. Therefore, the Smad pathway is tunable to the needs of each system simply by altering the values of the NCPs.

Mapping performance of biological systems to NCPs

To correlate different system behaviors with values of NCPs, we considered three different potential system behaviors. In the first behavior, System 1 favors linearity and rise time (emphasizing POs 1 and 3) while sacrificing NAR (Fig. 7A). In the second behavior, System 2 favors a small NAR (emphasizing PO2) while sacrificing rise time and linearity (Fig. 7A). In the final behavior, System 3 favors a short response time and high noise filtering (emphasizing POs 1 and 2) while sacrificing linearity (Fig. 7A). In these three different regions of PM-space, being limited to the Pareto Front, the NCPs take on hallmark distributions (Fig. 7B–P). For example, System 1 is characterized by high phosphatase and Smad1 levels, and low Smad4 levels. System 2 has low phosphatase and medium Smad1 and Smad4 levels. Finally, System 3 has medium phosphatase and Smad4 levels, and high Smad1 levels. This relationship between NCPs and PMs would allow us to predict NCP values from experimentally observed systems-level behavior and, perhaps more importantly, to rationally control systems-level behavior by manipulating NCP levels.

The combinatorial role of NCPs in mapping performance objectives along the Pareto plane

To determine the synergistic role of NCPs in mapping performance objectives along the Pareto optimal plane, we simulated the Smad model at a fixed concentration CIF and $Smad4_{total}$, and systematically varied the concentrations of PPase and $Smad1_{total}$, such that PPase varies along each curve and $Smad1_{total}$ varies across the curves (Fig. 8A). We found that varying PPase results in a family of curves that follow the Pareto front, and the section of the Pareto front traced out by a given PPase curve depends on



the concentration of $Smad1_{total}$ (Fig. 8A). We observed a similar behavior when we fixed $Smad1_{total}$ and CIF and varied $Smad4_{total}$ and PPase (Fig. 8B), although the extent to which $Smad4_{total}$ allows the PPase curves to sweep the Pareto optimal surface depends on the concentration of $Smad1_{total}$ (Supplementary Fig. 15). The results indicate that PPase is a major driver of the location of the system in PM-space while NCPs such as $Smad1_{total}$ and $Smad4_{total}$ allow the system's PMs to span the entire Pareto plane (Fig. 8C). Since CIF has only a small effect on the performance of a system (Figs. 4B, 5B, 7E–G) we did not consider it in this analysis. Note that driving PPase levels too high causes the system to move off the plane (Fig. 8A, B); however, for some values of $Smad1_{total}$ and $Smad4_{total}$, further increases in PPase levels bring the system back on the Pareto surface (Supplementary Fig. 15).

Fig. 4 | NCPs determine the position of the system in three-dimensional PM-space. A–E The centroid of ten point cloud clusters into which the PMs were sorted based on NCP concentration. The projections of the curve formed by connecting the centroids is as shown on the t_{rise}-NAR, t_{rise}-sensitivity coefficient (ϕ), NAR-sensitivity coefficient (ϕ) planes (shown in grey). Increasing circle diameter denotes increasing concentration. **A** An increase in the PPase concentration shifts the PMs of the system closer to the utopian point "UP". "UP" is the "ideal" optimal performance point with coordinates [t_{rise}, NAR, ϕ] = [0, 0, 1] (shown as a black-filled circle). **B** Increasing CIF concentration has a small effect in determining the position of the centroid of the point cloud clusters. **C** Increasing total Smad1 concentration continuously advances the centroid of the point cloud clusters toward the optimal point. After a certain level, any further increase in Smad1 has a small effect in repositioning the PMs. **D** The centroid of point cloud clusters forms a "U" shape such that increasing total Smad4 moves the system closer to the optimal point and then after a certain point any further increase in Smad4 moves the system away from the optimal point. **E** The $Smad1_{total}/Smad4_{total}$ has a mirror "U" trend with lower $Smad1_{total}/Smad4_{total}$ values being further away from the optimal than higher $Smad1_{total}/Smad4_{total}$. **F** The equal weights distance of the point cloud cluster from the optimal point for all four NCPs and $Smad1_{total}/Smad4_{total}$.

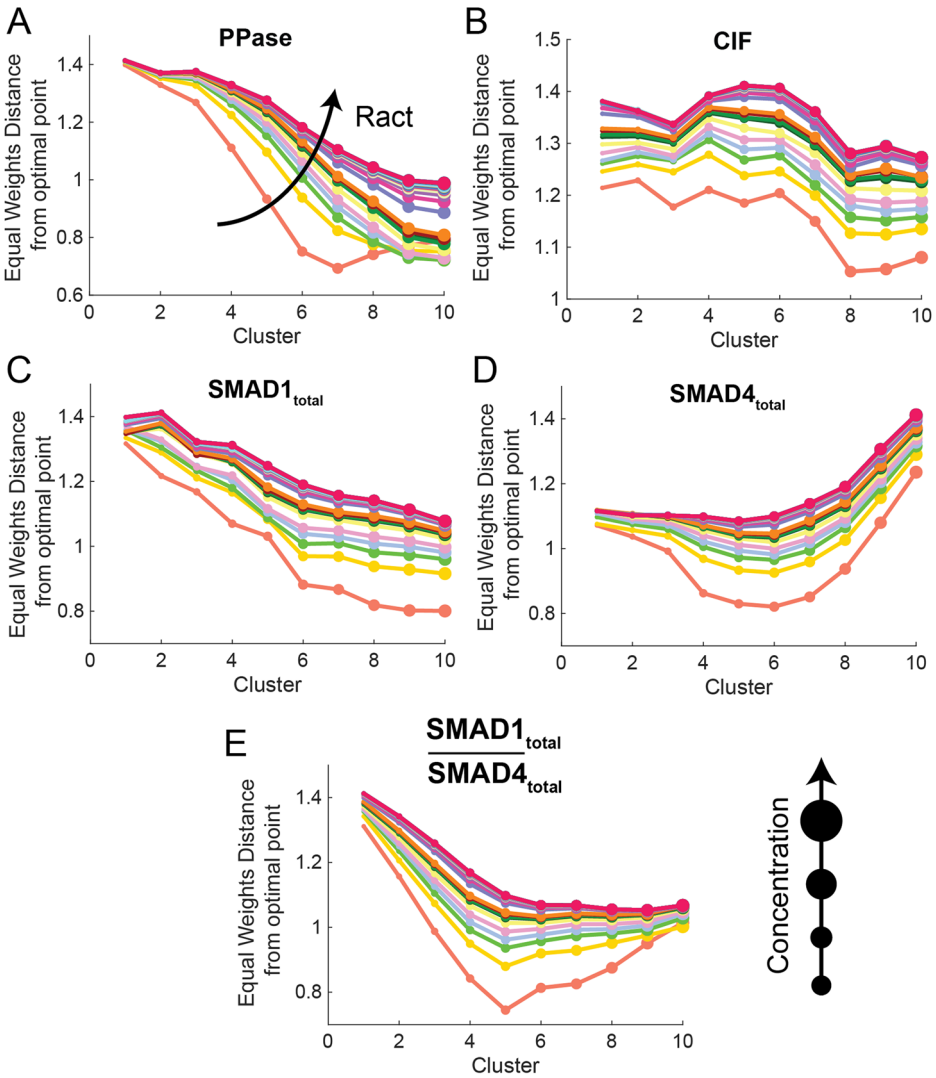


Fig. 5 | Equal Weights Distance of point cloud clusters. A–E The equal weights distance from the optimal point of the point cloud clusters at 30 Rmax (see Methods) inputs. The size of the circle represents increasing NCP concentrations. The equal weights distance from the optimal point of the point cloud clusters of (**A**) PPase (**B**) CIF (**C**) $Smad1_{total}$ (**D**) $Smad4_{total}$ (**E**) ratio of $Smad1_{total}/Smad4_{total}$ at 30 R_{max} inputs.

Phosphatase levels are variable across biological systems

Our results have shown that nuclear phosphatase, which dephosphorylates PSmad, has the greatest effect of the NCPs on the balance of trade-offs among the POs. To ascertain the variability in phosphatase levels across different species, we used available RNA-Seq data from *Drosophila* embryos, zebrafish embryos, and human cell data. Several phosphatases have been previously linked to the Smad pathway, such as PPM1A[43–45], PPM1H[46], MTMR4[47], SCP1/2/3[48], SCP4/CTDSPL2[49], Dullard, Pyruvate dehydrogenase phosphatase (PDP)[50], PP2A[51], etc. Notably, we identified that PPM1A, MTMR4, and Dullard were consistently expressed in *Drosophila*, zebrafish, and human cells. We focused on the available published RNA-Seq data for zebrafish embryos[52], *Drosophila* embryos[53], and human aorta cells[54] specifically examining the expression of the phosphatase Dullard (ctdnep1a in zebrafish and ctdnep1 in humans). The current data reflected that Dullard mRNA exhibited significant expression during the developmental stages of both zebrafish and *Drosophila* embryos, where the BMP/Smad pathway plays a crucial role in shaping the major body axis for these two species. These RNA-Seq datasets arise from different experimental groups, making direct comparisons across species on a universal scale challenging[55]. For a broader perspective on expression levels throughout different species, we conducted a comparison of Dullard expression relative to two highly conserved housekeeping genes, GAPDH (glyceraldehyde-3-phosphate dehydrogenase) and UBC (Ubiquitin C) (Data summary available in Supplementary Note 6). Our findings indicate that the relative expression level of Dullard is higher in *Drosophila* embryos than in zebrafish embryos as measured against both GAPHD and UBC. The Human Aorta data show a low expression level of Dullard compared to both the Drosophila embryo and the zebrafish embryo data in reference to GAPDH, but slightly higher

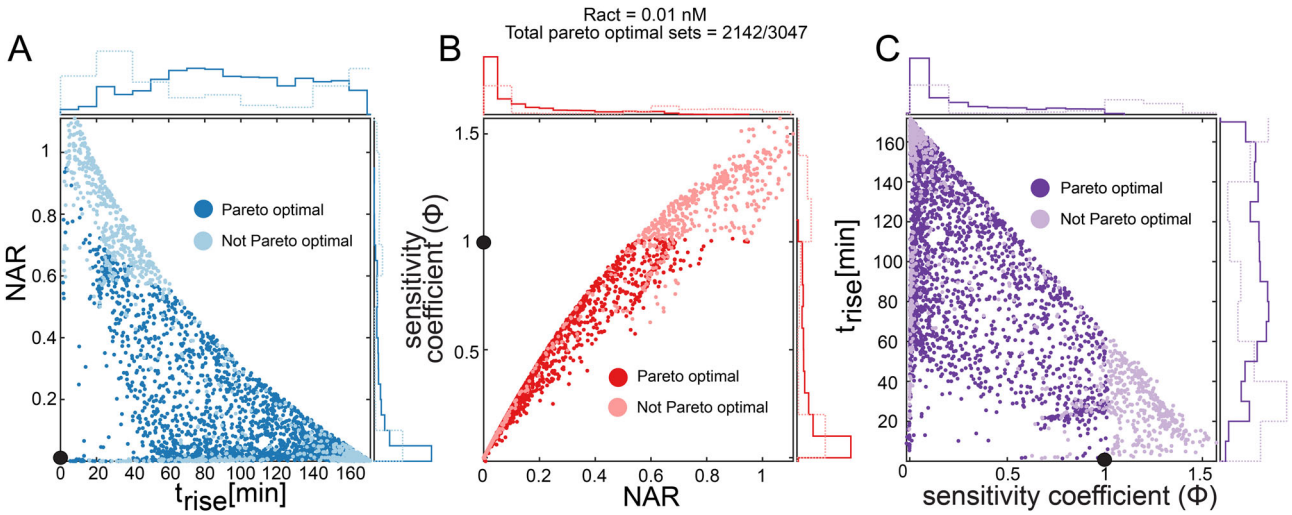


Fig. 6 | Pareto optimality of POs. Projections in the t_{rise}-NAR (**A**), NAR-sensitivity coefficient (ϕ), (**B**), and The sensitivity coefficient (ϕ) − t_{rise}, (**C**) planes of the three-dimensional PM-space. Darker filled circles represent Pareto optimal POs, whereas lighter circles represent POs that are not on the Pareto front. The histogram with solid lines indicates the distribution of Pareto optimal POs and the histogram with dashed line indicates the distribution of POs that are not Pareto optimal. indicated with black-filled circles.

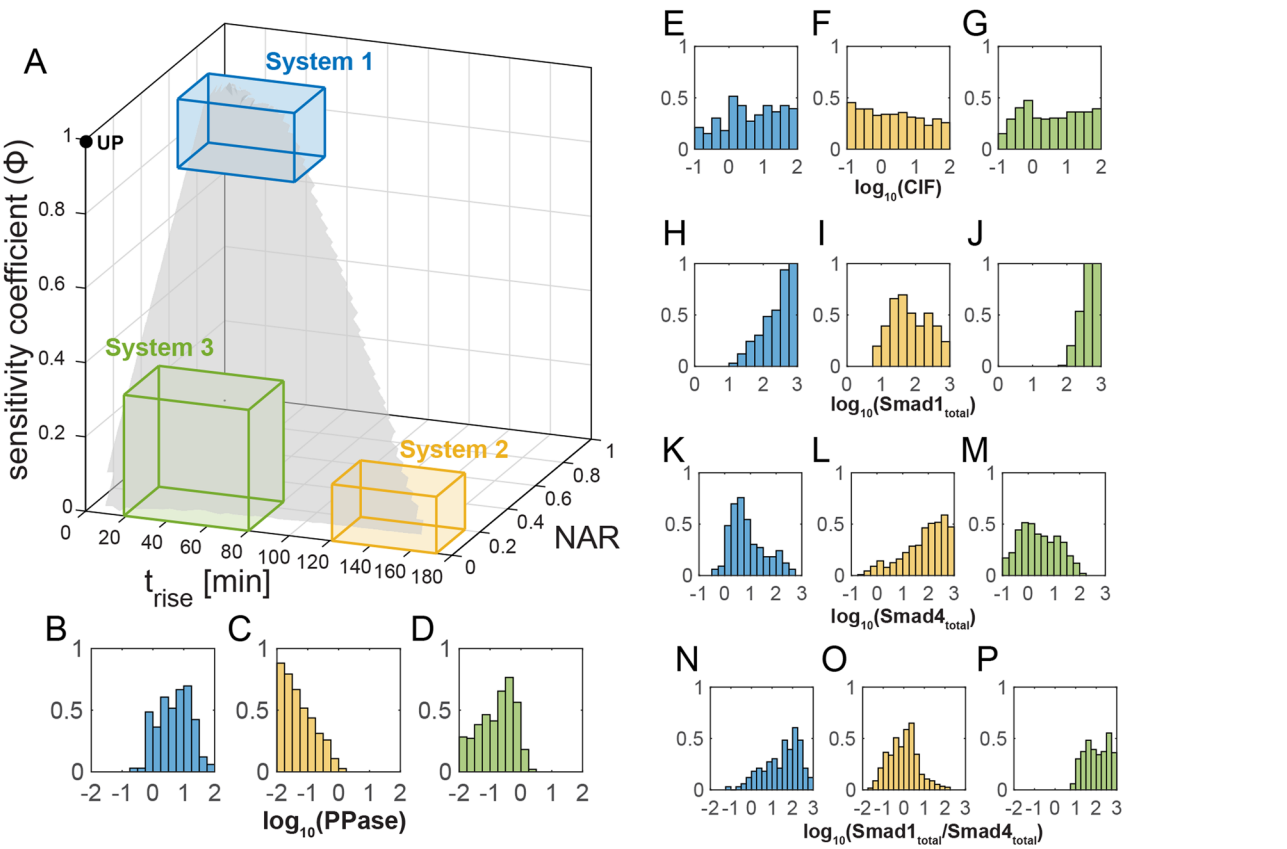


Fig. 7 | Mapping the performance of biological systems to NCPs. A We arbitrarily chose three systems on the PM surface such that system 1 favors linearity, system 2 favors NAR, and system 3 favors both rise time and NAR. The three systems lie on different regions on the surface fit through all the Pareto optimal PMs. **B**–**P** The distribution of each NCP, normalized by a probability distribution function, for the three systems in (**A**). **B**–**D** PPase: System 1 prefers high values of PPase, whereas System 2 favors low values of PPase, and System 3 favors intermediate values of PPase. **E**–**G** CIF**:** There is no observed trend in the distribution of the concentration of CIF across the three systems. **H**–**J** Smad1$_{total}$: In general, the distribution of Smad1$_{total}$ is centered around the highest concentrations for System 3, followed by System 1 and 2. **K**–**M** Smad4$_{total}$: The distribution of Smad4$_{total}$ is centered around the highest concentration for System 2, followed by System 1 and 3. **N**–**P** Smad1$_{total}$/Smad4$_{total}$: In general, System 1 prefers high, System 2 intermediate, and System 3 very high concentrations of Smad1$_{total}$/Smad4$_{total}$.

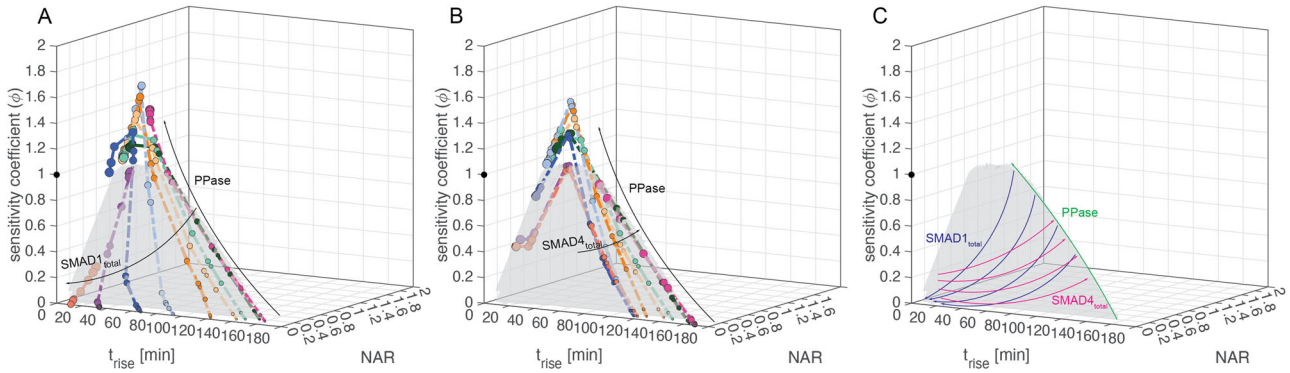


Fig. 8 | The combinatorial role of NCPs in mapping performance objectives along the Pareto plane. Each curve traces an increasing concentration of PPase, visualized with increasing marker size, at a fixed (**A**) $Smad1_{total}$ and (**B**) $Smad4_{total}$ which is varied across curves as indicated. The colors in (**A**) and (**B**) represent that the same $Smad1_{total}/Smad4_{total}$ was maintained while performing the two sets of simulations. **C** Illustrative trends by which each NCP moves the system along Pareto plane.

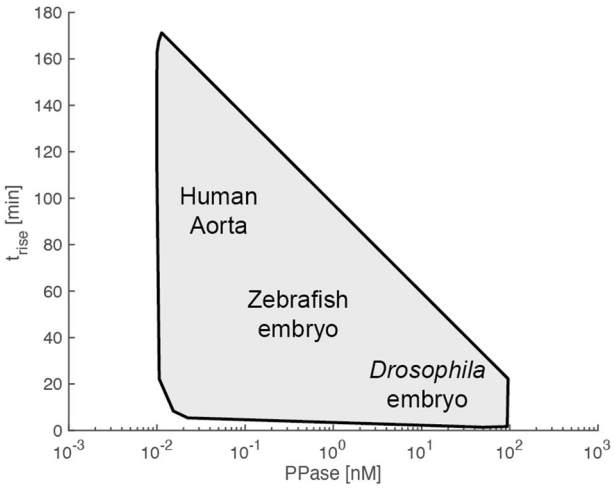


Fig. 9 | PPase levels across tissues and BMP signaling response time. The relationship between BMP signaling response time and relative PPase levels across various tissues can be mapped onto model-predicted space.

relative to the zebrafish embryo in refernce to UBC. This relative expression difference could be the result of the significantly higher expression of UBC in zebrafish embryos. To illustrate the significance of these findings, we constructed a convex hull from all Pareto optimal response time at their respective PPase levels, and annotated the three biological systems; Human Aorta, Zebrafish embryo, and *Drosophila* embryo based on relative PPase levels from RNA-Seq data (Table S10, ratio to GAPDH) and approximate BMP signaling response time (Fig. 1). Our findings are consistent with our prior modeling outcomes, and a higher level of PPase maintains POs by achieving a balance between lower noise filtration and heightened sensitivity, leading to quicker response times (Fig. 9).

Discussion

Highly conserved cell-cell signaling pathways, such as the BMP pathway, operate within multiple biological systems, and as such, must be adaptable and versatile to achieve a variety of goals specific to each organism. A fundamental question is how the general morphogen pathway has adapted to the specific needs of a given system, and what mechanisms are at its core in fine-tuning system performance. In the case of BMP-Smad pathway, despite the central and highly conserved role that BMP signaling plays in tissue patterning, the dynamics of signal transmission from the BMP input to the PSmad output varies widely across taxa or developmental stages. For example, BMP signaling requires a rapid response in *Drosophila* and zebrafish embryo development, which occurs on the order of minutes to a few hours. On the other hand, in *Drosophila* pupal vein formation and hiPSC differentiation, BMP signaling persists for hours and days, implying that noise filtering may be more important than a rapid response[25]. These observations suggest that the shared BMP signaling module may be tuned to emphasize different POs depending on the context. In certain scenarios, an "ideal" system with optimal POs would respond rapidly, filter noise, and act as a linear sensor of BMP concentration. These response features play a crucial role in determining the fitness of an organism when it faces challenges such as biological noise and perturbations. However, real communication systems are subject to constraints and must navigate trade-offs between various optimal solutions. Balancing these competing factors becomes essential for the functional adaptation and robustness of the BMP-Smad pathway.

Here we investigated the effect of non-conserved parameters (NCPs) on the performance objectives (POs) of the Smad nucleo-cytoplasmic dynamic model. The investigation involved the evaluation of three specific PMs: response time (PM1), noise-amplification-ratio (NAR) (PM2), and sensitivity coefficient (PM3). To assess these objectives, we evaluated the POs by characterizing the response of the system to the localization of $(PSmad1)_2/Smad4$ in the nucleus upon pathway activation. The results revealed an intriguing relationship between the concentration of one NCP, PPase, and the POs. As the concentration of PPase increased, the response time of the system decreased. This suggests that a higher concentration of PPase led to a more rapid response, indicating its role in modulating the response timing of the Smad nucleo-cytoplasmic dynamics. On the other hand, the results also unveiled that an elevated concentration of PPase correlated with an increased noise-amplification capability. This implies that PPase plays a vital role in filtering out extraneous noise, enhancing the system's ability to maintain accurate signaling despite external perturbations. Additionally, as the concentration of PPase was increased, the sensitivity coefficient of the system also increased. This suggests that PPase concentration influences the system's sensitivity to variations in the input signals, potentially allowing for more precise and nuanced responses to changes in the cellular environment. PPase plays an essential role in regulating the TGF-β/BMP signaling pathway by dephosphorylating key components to modulate its activity. Based on our preliminary analysis of the available published RNA-Seq data from *Drosophila* and zebrafish embryos, as well as human cell data, the phosphatase expression levels vary across different species, the relative expression level of Dullard is higher in *Drosophila* embryos than in zebrafish embryos as measured against both GAPHD and UBC. Our findings are consistent with our prior modeling outcomes, and a higher level of PPase maintains POs by achieving a balance between lower noise filtration and heightened sensitivity, leading to quicker response times. These observations align with the developmental requirements of *Drosophila* embryos, which necessitate rapid responsiveness during their developmental stages. The findings underscore the significance of protein phosphatases in modulating TGF-β/BMP signaling and highlight the need for further exploration into their role.

While PPase levels play the largest role in shaping the balance of trade-offs among the POs we analyzed here, it is only a singular parameter, and variation of PPase alone can, at best, trace out a single curve in PO-space. Therefore, the other NCPs, in particular, Smad1 and Smad4 concentrations, can fine-tune the balance of trade-offs to allow the system to sample the entire Pareto surface. Interestingly, our model predicts that Smad1 and Smad4 have largely opposite effects on the POs. Our model predictions of the relationship between NCPs and POs could be tested experimentally in a variety of BMP systems across multiple animal species.

In this work, we focused on only three POs, given they are requirements for communications systems[26,27,36,56]. However, the BMP system may have other competing POs that must also be balanced depending on the context. For example, signal amplification, precision, and information processing are plausible POs that could also be considered in the future[26,27,57].

This study presents a fresh perspective on the regulation of BMP/Smad signaling by emphasizing the significance of considering multiple systems-level objectives in determining signaling pathway function across different species. This research also sheds light on the flexibility of the BMP/Smad pathway to achieve different outcomes in different contexts, and further exploration is warranted to determine how cells regulate the concentrations of pathway components to tune the POs according to the needs of the cellular system. Our results, which are based on engineering principles of trade-offs and compromise among competing objectives, are likely to generalize to other biological pathways and have implications not only for developmental biology but also for disease-related contexts including regenerative medicine.

Methods

Simulating the Smad model

For this study, we modified the Smad signal transduction model reported in Schmierer et al. [31]. In this model, the ligand TGF-β binds to the receptor resulting in the phosphorylation of Smad2, a receptor-regulated Smad (R-Smad). The phosphorylated-Smad2 (PSmad2) forms homomeric or heteromeric complexes with Smad4, a common mediator Smad (Co-Smad) which then translocates to the nucleus to regulate gene expression. We implemented this mathematical model to simulate the BMP pathway and model Smad1 dynamics as its network structure is similar to TGF- β pathway. The Smad complexes localize to the nucleus, which is modeled as a multiplication factor, complex-import factor (CIF).

It has been reported in the literature[58–60] that the active signaling complex is a trimer consisting of two molecules of PSmad1 and one molecule of Smad4 (see Supplementary Table 3 for complete model reactions). We accordingly modified the model to account for interactions between dimers that result in the formation of the trimer—$(PSmad1)_2$/Smad4 (see Supplementary Table 4 for model equations). We fit this modified model to the nuclear EGFP-Smad2 concentration data reported in Schmierer et al. [31] (Supplementary Fig. 4).

In our modified model (referred to as the Smad model; See Supplementary Note 3), before the activation of the pathway, only monomeric-Smad1 is present in the nucleus, and after the pathway is activated due to TGF-β signaling all Smad species such as PSmad1, PSmad1/Smad4, PSmad1/PSmad1, and $(PSmad1)_2$/Smad4 are present in the nucleus. This leads to the accumulation of Smad1 species in the nucleus when the TGF-β signaling is active. All the Smad species (monomeric and dimeric) can dynamically move between the cytoplasm and the nucleus. An externally supplied inhibitor that deactivates the receptors turns off this signaling network. In our Smad model, the complex-import factor (CIF), which is a multiplication factor to the import rate of the Smad complex to denote enhanced nuclear translocation of the Smad complexes, only multiplies the import rate constant of the trimeric Smad complex: $(PSmad1)_2$/Smad4.

Simulation conditions. The initial conditions of all protein complexes were set to 0 (Supplementary Table 5).

To calculate the POs at randomly generated NCPs we simulated the Smad model (Supplementary Table 9) for 24 h (R_{max}) and 48 h (R_{min}).

Smad1 and Smad4 initial conditions. Smad1 distribution in initial conditions:

$$K_{eq} = \frac{k_{in}}{k_{ex}} \tag{1}$$

$$Smad1_{tot}(cyt.) = \frac{Smad1_{tot}}{1 + K_{eq}} \tag{2}$$

$$Smad1_{tot}(nuc.) = Smad1_{tot} - Smad1_{tot}(cyt.) \tag{3}$$

Smad4 is equally distributed in the cytoplasm and the nucleus before the activation of the pathway.

CIF and PPase were passed to the model without modification.

Parameters for Smad model. We implemented Improved Stochastic Ranking Evolutionary Strategy plus (ISRES+)[61] to estimate the 10 variable parameters in the model (Supplementary Table 2) by minimizing the sum of squared errors (SSE) between the experimental data and the model predictions, normalized by standard deviation. The Smad model (Supplementary Table 3) was fit to nuclear EGFP-Smad1 concentrations in HaCat cells in the presence of TGF-β signaling and later when an inhibitor of the Smad pathway was externally added (see Supplementary Table 5 for initial conditions). ISRES+ is a (μ-λ)-based evolutionary algorithm employed to solve global optimization problems using stochastic ranking. It additionally has gradient-based algorithms embedded to improve the search strategy, called lin-step and newton-step. The algorithm was run for 100 generations with 150 individuals, with a recombination rate of 0.85. Lin-step contributed one individual and Newton-step contributed two individuals to every generation. The complete description of hyperparameters and bounds on the model parameters supplied to ISRES+ are in the Supplementary Tables 7 and 8. The best individual of all generations, i.e., the parameter set with the best fitness score was chosen for all the simulations in this study (Supplementary Table 6). The model fit with the estimated parameters is shown in Supplementary Fig. 4. Additional details are in the Supplementary Text Note 3.

Receptor concentrations. The time-varying value of the activated receptor (R_{act}) is the number of active receptor complexes predicted by a stochastic model of the extracellular BMP signaling pathway. These activated receptor level inputs used to calculate the PM-2 were obtained using a model from Larson et al.[35]. This model uses the GillesPy python package[62] to simulate the cell surface BMP receptor network as it interacts with given extracellular BMP ligand concentrations. There were two sets of starting parameters for the number of receptor components. Based on literature TGF-β receptor estimations from[63,64], we used a value of 14,000 TGF-β receptors per cell split between type-I and type-II[65]. suggests that at some stages only ten percent of these may be present on the cell membrane. To this end, the simulations were run a second time with "R_{min}" being ten percent of "R_{max}". For each of the two starting receptor levels, extracellular BMP ligand concentrations ranging from 0.001 to 3 nM were used to initialize the simulation. Following 24–48 h of simulation, all species population values were tracked over 24 h to collect noise information at the steady state (see Supplementary Note 4.1 for frequency analysis on stochastic receptor inputs).

The R_{act} values for "R_{max}" range from 0.01 nM to 3.5 nM, and for "R_{min}" range from 0.001 to 0.1 nM (Figure S5).

To calculate the PM-1 and PM-3, we adapt (details in the Supplementary Text Note 4) our model and simulate at starting receptor levels as a constant input.

Sampling non-conserved protein concentrations

We generated 10^4 sets of three non-conserved protein concentrations (Smad4, CIF, and PPase) using Latin-hypercube design with a sample

Table 1 | Bounds on the non-conserved parameters

	Min (nM)	Max (nM)
CIF	10^{-1}	10^{2}
PPase	10^{-2}	10^{2}
$Smad1_{tot}$	10^{-1}	10^{3}
$Smad4_{tot}$	10^{-1}	10^{3}

density of 10^4. We uniformly sample the Smad1/Smad4 ratio in the logspace (Table 1).

The following were the range we generated to sample within:

The simulation results with nuclear $(PSmad1)_2$/Smad4 concentration less than 0.1 nM are not shown in the figures since these concentrations are too low to be biologically relevant.

Calculating the three performance objectives

The three performance objectives (POs) are:

Rise time (t_{rise}). The rise time is defined as the duration of a response to reach 95% of its steady state value. We calculated the rise time using a MATLAB function stepinfo from the control systems toolbox, on the dynamic profile of total nuclear $(PSmad1)_2$/Smad4. The steady state $(PSmad1)_2$/Smad4 concentration was externally supplied as an input to stepinfo, which was calculated by solving the system of ODEs using Newton's method with the final condition from ode23tb in MATLAB as an initial guess. The "RiseTimeLimits" were defined as [0 0.95].

Noise amplification ratio (NAR). The NAR is defined as the ratio of coefficient of variation of the response signal to the input signal. We used the steady state of nuclear $(PSmad1)_2$/Smad4 obtained in PO1 as an initial condition to simulate continuous model again using ode23tb with a "MaxStep" of 0.5 with a stochastic receptor input. We simulate the model for the same timespan as PO1 i.e., 24 h for "Rmax" and 48 h for "Rmin". The NAR is then calculated as the ratio CV of $(PSmad1)_2$/Smad4 over R_{act}:

$$\mathrm{NAR} = \frac{\mathrm{CV}((\mathrm{PSmad1})_2/\mathrm{Smad4})}{\mathrm{CV}(\mathrm{R_{act}})} \tag{4}$$

Here,

$$\mathrm{CV(x)} = \frac{\sigma(\mathrm{x})}{\mu(\mathrm{x})} \tag{5}$$

Sensitivity coefficient (ϕ). We calculated the sensitivity coefficient at the steady state concentration of $(PSmad1)_2$/Smad4 obtained by solving the system of ODEs using Newton's method with the final condition from ode23tb in MATLAB as an initial guess. The sensitivity coefficient defined as:

$$\phi = \frac{d\ln\left(\frac{(\mathrm{PSmad1})_2}{\mathrm{Smad4}}\right)}{d\ln(\mathrm{Ract})} \tag{6}$$

Here,

$$\phi = \frac{\mathrm{Ract}}{(\mathrm{PSmad1})_2/\mathrm{Smad4}} \times \frac{(\mathrm{PSmad1})_2/\mathrm{Smad4}_\delta - (\mathrm{PSmad1})_2/\mathrm{Smad4}}{\mathrm{Ract}_\delta - \mathrm{Ract}} \tag{7}$$

Here, $\mathrm{Ract}_\delta = 1.01 \times R_{act}$ and $(PSmad1)_2/Smad4_\delta$ ($SC_{nuc\delta}$) is the steady state concentration of $(PSmad1)_2$/Smad4 (SC_{nuc}) for Ract_δ.

Pareto optimality

We define an optimal point that has coordinates such that $t_{rise} = 0$, NAR = 0, and ϕ = 1.

The optimization problem was formulated as:

$$\text{minimize } F = f\left(\frac{t_{rise}}{\max(t_{rise})}, \mathrm{NAR}, \mathrm{abs}(\log_{10}(abs(\phi)))\right) \tag{8}$$

$$t_{rise} = f(PPase, CIF, Smad1, Smad4) \tag{9}$$

$$\mathrm{NAR} = f(PPase, CIF, Smad1, Smad4) \tag{10}$$

$$\phi = f(PPase, CIF, Smad1, Smad4) \tag{11}$$

$$PPase, CIF, Smad1, Smad4 \geq 0 \tag{12}$$

We determined Pareto optimality by selecting non-dominated solutions[66,67], which are defined as solutions for which no other member of the solution set is closer to the UP in all three PMs. This was done by calculating the distance between every point in the solution set from the utopian point. These distances were calculated by normalizing the POs to ensure they lie in the same range. The response speed, t_{rise} (PM 1) was normalized by the maximum t_{rise} in the random screen. The abs $\left(\log_{10} abs(\phi)\right)$ value of ϕ was used to ensure ϕ = 1 is optimal.

Centroid of point clouds

The POs were sorted into ten discrete clusters to visualize the effect of varying NCP. We calculated the centroid of each cluster by taking the average of the three POs. The same procedure is used to calculate the centroid of point cloud clusters at all R_{act} concentrations.

The Euclidean distance (d) was calculated as follows:

$$\mathrm{d} = \sqrt{\sum_{k=1}^{3} (\underline{\mathrm{PM}}_{\mathrm{NCP_i(k)}} - \mathrm{OP}_{\mathrm{PM(k)}})^2} \tag{13}$$

$$\mathrm{NCP_i} \in (\mathrm{Ppase}, \mathrm{CIF}, \mathrm{Smad1_{tot}}, \mathrm{Smad4_{tot}}) \tag{14}$$

Surface fitting and generating distribution of NCPs

To visualize the PM-space in three dimensions and correlate system behavior using a data-driven approach, we fit a curve through all the Pareto optimal solutions using the "fit" function in MATLAB using "cubicinterp" method and setting "ExtrapolationMethod" to "none". The resulting surface is as shown in Fig. 7A. To generate the distribution of NCPs in distinct regions on the PM surface, we draw a cuboid in three dimensions using the interactive MATLAB function "drawcuboid", and reverse calculate the distribution of NCPs that allows each system behavior. The histograms were normalized by probability distribution function.

The combinatorial role of NCPs

The simulations in Fig. 8A were performed at a fixed CIF concentration of 50 nM and $Smad4_{total}$ equal to 50 nM and in Fig. 8B were performed at a fixed CIF concentration of 50 nM and $Smad1_{total}$ equal to 50 nM. The concentration of PPase along each curve in Fig. 8 ranges from 0.01 nM to 100 nM. Across each curve, the Smad1/Smad4 is in the range 10^{-2}–10^4.

The same conditions were used to simulate Supplementary Fig. 15, except the concentration of Smad1total was fixed to values in the range {50, 100, 500, 1000, 5000}.

Phosphatase levels across biological systems

The illustration in Fig. 9 is a convex hull constructed with Pareto optimal response time along increase PPase levels. It was computed by performing a Delaunay triangulation on the 2-D response time v/s PPase concentration points, which was then used to identify the vertices of the convex hull on MATLAB. We then annotated the three biological systems: Human Aorta, Zebrafish embryo, and Drosophila embryo based on relative PPase levels

from RNA-Seq data (Table S10, ratio to GAPDH) and approximate BMP signalingresponse time (Fig. 1).

Data availability

The data and code for this manuscript are uploaded here: https://github.com/gtreeves/balancing-pos-shaikh2024.

Received: 13 February 2024; Accepted: 21 August 2024;
Published online: 14 September 2024

References

1. Shi, Y. & Massagué, J. Mechanisms of TGF-β signaling from. *Cell Membr. Nucl. cell* **113**, 685–700 (2003).
2. Bayat, V., Jaiswal, M. & Bellen, H. J. The BMP signaling pathway at the Drosophila neuromuscular junction and its links to neurodegenerative diseases. *Curr. Opin. Neurobiol.* **21**, 182–188 (2011).
3. Davies, E. L. & Fuller, M. T. Regulation of self-renewal and differentiation in adult stem cell lineages: lessons from the drosophila male germ line. *Cold Spring Harb. Symposia Quant. Biol.* **73**, 137–145 (2008).
4. Shaikh, S., Ravenndranath, R., Banerjee, M., Joseph, A. & Jahgirdar, P. Evidence for transforming growth factor-beta 3 gene polymorphism in non-syndromic cleft lip and palate patients from Indian sub-continent. *Med. Oral Patol. Oral y Cir. Bucal*, **17**, e197-e200 (2012).
5. Finnson, K. W., Chi, Y., Bou-Gharios, G., Leask, A. & Philip, A. TGF-b signaling in cartilage homeostasis and osteoarthritis. *Front Biosci.* **4**, 251e268 (2012).
6. Samanta, D. & Datta, P. K. Alterations in the Smad pathway in human cancers. *Front. Biosci.* **17**, 1281 (2012).
7. Restrepo, S., Zartman, J., Jeremiah & Basler, K. Coordination of patterning and growth by the morphogen DPP. *Curr. Biol.* **24**, R245–R255 (2014).
8. Padgett, R. W., Wozney, J. M. & Gelbart, W. M. Human BMP sequences can confer normal dorsal-ventral patterning in the Drosophila embryo. *Proc. Natl Acad. Sci.* **90**, 2905–2909 (1993).
9. Holley, S. A. et al. A conserved system for dorsal-ventral patterning in insects and vertebrates involving sog and chordin. *Nature* **376**, 249–253 (1995).
10. Holley, S. A. et al. The Xenopus dorsalizing factor noggin ventralizes drosophila embryos by preventing DPP from activating its receptor. *Cell* **86**, 607–617 (1996).
11. Bier, E. & De Robertis, E. M. BMP gradients: a paradigm for morphogen-mediated developmental patterning. *Science* **348**, aaa5838 (2015).
12. Raftery, L. A., Twombly, V., Wharton, K. & Gelbart, W. M. Genetic screens to identify elements of the decapentaplegic signaling pathway in Drosophila. *Genetics* **139**, 241–254 (1995).
13. Raftery, L. A. & Sutherland, D. J. Gradients and thresholds: BMP response gradients unveiled in Drosophila embryos. *TRENDS Genet.* **19**, 701–708 (2003).
14. O'Connor, M. B., Umulis, D., Othmer, H. G. & Blair, S. S. Shaping BMP morphogen gradients in the Drosophila embryo and pupal wing. *Development* **133**, 183–193 (2006).
15. Gavin-Smyth, J., Wang, Y.-C., Butler, I., Ferguson, L. & Ferguson, E. L. A genetic network conferring canalization to a bistable patterning system in Drosophila. *Curr. Biol.* **23**, 2296–2302 (2013).
16. Al Asafen, H., Beseli, A., Chen, H. Y., Hiremath, S., Williams, C. M. & Reeves, G. T. Dynamics of BMP signaling and stable gene expression in the early Drosophila embryo. *Biol. Open.* **13**, bio061646 (2024).
17. Yan, S. J. et al. Bistability coordinates activation of the EGFR and DPP pathways in Drosophila vein differentiation. *Mol. Syst. Biol.* **5**, 278 (2009).
18. Villa-Fombuena, G., Lobo-Pecellin, M., Marin-Menguiano, M., Rojas-Rios, P. & Gonzalez-Reyes, A. Live imaging of the Drosophila ovarian niche shows spectrosome and centrosome dynamics during asymmetric germline stem cell division. *Development* **148**, dev199716 (2021).
19. Tucker, J. A., Mintzer, K. A. & Mullins, M. C. The BMP signaling gradient patterns dorsoventral tissues in a temporally progressive manner along the anteroposterior axis. *Dev. Cell* **14**, 108–119 (2008).
20. Ramel, M.-C. & Hill, C. S. The ventral to dorsal BMP activity gradient in the early zebrafish embryo is determined by graded expression of BMP ligands. *Dev. Biol.* **378**, 170–182 (2013).
21. Thompson, M. J., Young, C. A., Munnamalai, V. & Umulis, D. M. Early radial positional information in the cochlea is optimized by a precise linear BMP gradient and enhanced by SOX2. *Sci. Rep.* **13** https://doi.org/10.1038/s41598-023-34725-4 (2023).
22. Subileau, M. et al. Bone morphogenetic protein 9 regulates early lymphatic-specified endothelial cell expansion during mouse embryonic stem cell differentiation. *Stem Cell Rep.* **12**, 98–111 (2019).
23. Derynck, R. & Akhurst, R. J. BMP-9 balances endothelial cell fate. *Proc. Natl Acad. Sci.* **110**, 18746–18747 (2013).
24. Ponomarev, L. C., Ksiazkiewicz, J., Staring, M. W., Luttun, A. & Zwijsen, A. The BMP pathway in blood vessel and lymphatic vessel biology. *Int. J. Mol. Sci.* **22**, 6364 (2021).
25. Alderfer, L., Wei, A. & Hanjaya-Putra, D. Lymphatic tissue engineering and regeneration. *J. Biol. Eng.* **12** https://doi.org/10.1186/s13036-018-0122-7 (2018).
26. Lander, A. D., Lo, W.-C., Nie, Q. & Wan, F. Y. M. The Measure of success: constraints, objectives, and tradeoffs in morphogen-mediated patterning. *Cold Spring Harb. Perspect. Biol.* **1**, a002022–a002022 (2009).
27. Lo, W.-C., Zhou, S., Wan, F. Y.-M., Lander, A. D. & Nie, Q. Robust and precise morphogen-mediated patterning: trade-offs, constraints and mechanisms. *J. R. Soc. Interface* **12**, 20141041 (2015).
28. Qiao, L., Zhao, W., Tang, C., Nie, Q. & Zhang, L. Network topologies that can achieve dual function of adaptation and noise attenuation. *Cell Syst.* **9**, 271–285.e277 (2019).
29. Hornung, G. & Barkai, N. Noise propagation and signaling sensitivity in biological networks: a role for positive feedback. *PLoS Comput. Biol.* **4**, e8 (2008).
30. Lochab, A. K. & Extavour, C. G. Bone morphogenetic protein (BMP) signaling in animal reproductive system development and function. *Dev. Biol.* **427**, 258–269 (2017).
31. Schmierer, B., Tournier, A. L., Bates, P. A. & Hill, C. S. Mathematical modeling identifies Smad nucleocytoplasmic shuttling as a dynamic signal-interpreting system. *Proc. Natl Acad. Sci. USA* **105**, 6608–6613 (2008).
32. Shoval, O. et al. Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space. *Science* **336**, 1157–1160 (2012).
33. Schuetz, R., Zamboni, N., Zampieri, M., Heinemann, M. & Sauer, U. Multidimensional optimality of microbial metabolism. *Science* **336**, 601–604 (2012).
34. Mertens, J., Stenger, R. & Barkle, G. F. Multiobjective inverse modeling for soil parameter estimation and model verification. *Vadose Zone J.* **5**, 917–933 (2006).
35. Larson, N. J., Madamanchi, A., Karim, M. S., Li, L. & Umulis, D. M. ([Unpublished Manuscript]).
36. Karim, M. S., Buzzard, G. T. & Umulis, D. M. Secreted, receptor-associated bone morphogenetic protein regulators reduce stochastic noise intrinsic to many extracellular morphogen distributions. *J. R. Soc. Interface* **9**, 1073–1083 (2012).
37. Locasale, J. W., Shaw, A. S. & Chakraborty, A. K. Scaffold proteins confer diverse regulatory properties to protein kinase cascades. *Proc. Natl Acad. Sci.* **104**, 13307–13312 (2007).
38. Mishra, A. et al. A protein phosphatase network controls the temporal and spatial dynamics of differentiation commitment in human epidermis. *Elife* **6**, e27356 (2017).

39. Landry, B. P., Palanki, R., Dyulgyarov, N., Hartsough, L. A. & Tabor, J. J. Phosphatase activity tunes two-component system sensor detection threshold. *Nat. Commun.* **9** https://doi.org/10.1038/s41467-018-03929-y (2018).
40. Chen, W. W. et al. Input–output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. *Mol. Syst. Biol.* **5**, 239 (2009).
41. Bruce, D. L. & Sapkota, G. P. Phosphatases in SMAD regulation. *FEBS Lett.* **586**, 1897–1905 (2012).
42. Boada, Y., Reynoso-Meza, G., Picó, J. & Vignoni, A. Multi-objective optimization framework to obtain model-based guidelines for tuning biological synthetic devices: an adaptive network case. *BMC Syst. Biol.* **10** https://doi.org/10.1186/s12918-016-0269-0 (2016).
43. Lin, X. et al. PPM1A functions as a Smad phosphatase to terminate TGFβ signaling. *Cell* **125**, 915–928 (2006).
44. Duan, X., Liang, Y.-Y., Feng, X.-H. & Lin, X. Protein serine/threonine phosphatase PPM1A dephosphorylates Smad1 in the bone morphogenetic protein signaling pathway. *J. Biol. Chem.* **281**, 36526–36532 (2006).
45. Kokabu, S. et al. Protein phosphatase magnesium-dependent 1A–mediated inhibition of BMP signaling is independent of Smad dephosphorylation. *J. Bone Miner. Res.* **25**, 653–660 (2010).
46. Shen, T. et al. Specific control of BMP signaling and mesenchymal differentiation by cytoplasmic phosphatase PPM1H. *Cell Res.* **24**, 727–741 (2014).
47. Yu, J. et al. Myotubularin-related protein 4 (MTMR4) attenuates BMP/Dpp signaling by dephosphorylation of Smad proteins. *J. Biol. Chem.* **288**, 79–88 (2013).
48. Knockaert, M., Sapkota, G., Alarcón, C., Massagué, J. & Brivanlou, A. H. Unique players in the BMP pathway: small C-terminal domain phosphatases dephosphorylate Smad1 to attenuate BMP signaling. *Proc. Natl Acad. Sci.* **103**, 11940–11945 (2006).
49. Zhao, Y. et al. C-terminal domain (CTD) small phosphatase-like 2 modulates the canonical bone morphogenetic protein (BMP) signaling and mesenchymal differentiation via Smad dephosphorylation. *J. Biol. Chem.* **289**, 26441–26450 (2014).
50. Chen, H. B., Shen, J., Ip, Y. T. & Xu, L. Identification of phosphatases for Smad in the BMP/DPP pathway. *Genes Dev.* **20**, 648–653 (2006).
51. Bengtsson, L. et al. PP2A regulates BMP signalling by interacting with BMP receptor complexes and by dephosphorylating both the C-terminus and the linker region of Smad1. *J. Cell Sci.* **122**, 1248–1257 (2009).
52. White, R. J. et al. A high-resolution mRNA expression time course of embryonic development in zebrafish. *Elife* **6** https://doi.org/10.7554/eLife.30860 (2017).
53. Graveley, B. R. et al. The developmental transcriptome of Drosophila melanogaster. *Nature* **471**, 473–479 (2011).
54. Lonsdale, J. et al. The genotype-tissue expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
55. Deshpande, D. et al. RNA-seq data science: From raw data to effective interpretation. *Front. Genet.* **14** https://doi.org/10.3389/fgene.2023.997383 (2023).
56. Reeves, G. T. The engineering principles of combining a transcriptional incoherent feedforward loop with negative feedback. *J. Biol. Eng.* **13** https://doi.org/10.1186/s13036-019-0190-3 (2019).
57. Dubuis, J. O., Tkačik, G., Wieschaus, E. F., Gregor, T. & Bialek, W. Positional information, in bits. *Proc. Natl Acad. Sci.* **110**, 16301–16308 (2013).
58. Schmierer, B. & Hill, C. S. TGF beta-SMAD signal transduction: molecular specificity and functional flexibility. *Nat. Rev. Mol. Cell Biol.* **8**, 970–982 (2007).
59. Massagué, J., Seoane, J. & Wotton, D. Smad transcription factors. *Genes Dev.* **19**, 2783–2810 (2005).
60. Tiago, G., Pau, M. M., Lidia, R., Eric A., Tiago, N. C., Maria, J. M. Conformational landscape of multidomain SMAD proteins. *Comput. Struct. Biotechnol. J.* **19**, 5210-5224 (2021).
61. Bandodkar, P., Shaikh, R. & Reeves, G. T. ISRES+: an improved evolutionary strategy for function minimization to estimate the free parameters of systems biology models. *Bioinformatics* **39** https://doi.org/10.1093/bioinformatics/btad403 (2023).
62. Abel, J. H., Drawert, B., Hellander, A. & Petzold, L. R. GillesPy: a Python package for stochastic model building and simulation. *IEEE Life Sci. Lett.* **2**, 35–38 (2016).
63. Wakefield, L. M., Smith, D. M., Masui, T., Harris, C. C. & Sporn, M. B. Distribution and modulation of the cellular receptor for transforming growth factor-beta. *J. Cell Biol.* **105**, 965–975 (1987).
64. Nicklas, Da. S. L. Computational modelling of Smad-mediated negative feedback and crosstalk in the TGF-β superfamily network. *J. R. Soc. Interface* **10**, 20130363 (2013).
65. Chung, S.-W. et al. Quantitative modeling and analysis of the transforming growth factor β signaling pathway. *Biophys. J.* **96**, 1733–1750 (2009).
66. Patanè, A., Santoro, A., Costanza, J., Carapezza, G. & Nicosia, G. Pareto optimal design for synthetic biology. *IEEE Trans. Biomed. Circuits Syst.* **9**, 555–571 (2015).
67. Otero-Muras, I. & Banga, J. R. Multicriteria global optimization for biocircuit design. *BMC Syst. Biol.* **8**, 113 (2014).

Acknowledgements

This work is based upon efforts supported by the EMBRIO Institute, contract #2120200, a National Science Foundation (NSF) Biology Integration Institute. This research was supported in part by the NIH grant R01GM132501 awarded to D.U. and by NSF grant 2313692 awarded to G.T.R. and by NSF grant 2047903 awarded to D.H.-P. Portions of this research were conducted with the advanced computing resources provided by Texas A&M High Performance Research Computing.

Author contributions

D.U., G.T.R., J.Z., D.H-P., L.L. conceptualized the study. R.S., G.T.R., L.L., and N.L. developed the methodology. R.S. developed the model and performed the simulations (except sine wave analysis), formal analysis, and visualizations. L.L. compiled and analyzed the RNA-Seq data. J.K. simulated the model with sine wave inputs at individual frequency components. R.S., L.L., and G.T.R. wrote the original draft. R.S., G.T.R., L.L., D.U., J.Z., D.H.-P., N.L. reviewed and edited the manuscript. G.T.R, N.L., R.S., L.L, and D.H.-P. contributed to illustrations in Fig. 1. The work was supervised by G.T.R., D.U., L.L, J.Z., D.H.-P. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41540-024-00430-9.

Correspondence and requests for materials should be addressed to David M. Umulis, Linlin Li or Gregory T. Reeves.