Linearity in Cell Signaling Pathways

Thesis by
Harry Nunns

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2019
Defended January 11, 2019

ABSTRACT

Accurate cellular communication is of paramount importance for the development, growth, and maintenance of multi-cellular organisms. Communication between cells is carried out by a highly conserved set of signaling pathways, whose dysregulation can lead to many diseases. The molecular details of these signaling pathways are now well-characterized, allowing researchers to investigate the emergent properties that arise from the complex signaling networks. These properties often arise from counter-intuitive or paradoxical mechanisms, meaning that systems-level analysis is necessary. Importantly, mathematical models have been constructed for many pathways that capture measured reaction rates and protein levels. These mathematical models successfully recapitulate dynamic responses of each pathway. Here, I investigated the input-output response of the Wnt, MAPK/ERK, and Tgfβ pathways using analytical and numerical treatment of mathematical models. Using this approach, I found that the distinct architectures of the three signaling pathways lead to a convergent behavior, linear input-output response. Specifically, mathematical analysis reveals that a futile cycle in the Wnt pathway, a kinase cascade coupled to feedback in the ERK pathway, and nucleocytoplasmic shuttling in the Tgfβ pathways all yield linear signal transmission. I then verified this finding experimentally in the Wnt and ERK pathways. For the Wnt pathway, direct measurements of the input-output response reveal that β-catenin is linear with respect to Wnt co-receptor LRP5/6 activity up until receptor saturation. For the ERK pathway, direct measurements indicate a linear relationship between phosphorylated ERK1/2 and the concentration of EGF ligand, up until saturation of ERK1/2. Finally, mathematical modeling reveals that linear response in the Wnt pathway, in conjunction with a recently identified cis-regulatory motif, is sufficient to explain gene expression buffering to perturbations. Therefore, this thesis demonstrates how linearity emerges across three dissimilar architectures, and introduces a novel benefit for linear signal transmission in biology.

PUBLISHED CONTENT AND CONTRIBUTIONS

[1] Kibeom Kim, Jaehyoung Cho, Thomas S Hilzinger, Harry Nunns, Andrew Liu, Bryan E Ryba, and Lea Goentoro. Two-element transcriptional regulation in the canonical wnt pathway. *Current Biology*, 27(15):2357–2364, 2017. doi: https://doi.org/10.1016/j.molcel.2009.11.018.
H.N. performed the mathematical analysis for the paper.

[2] Harry Nunns and Lea Goentoro. Signaling pathways as linear transmitters. *eLife*, 7:e33617, 2018. doi: https://doi.org/10.7554/eLife.33617.001.
H.N. developed all the results and figures for the paper.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Chapter 1

INTRODUCTION

Signaling pathways are critical to metazoan development and maintenance

Development and maintenance of multi-cellular organisms depend on the coordinated growth, movement, and death of large numbers of cells. After fertilization, a single egg rapidly divides and differentiates into an increasingly complex embryo; in humans, normal development produces hundreds of cell types and trillions of cells, each guided to the correct location at the correct time [8, 9]. To achieve this outcome, cells must communicate with nearby cells (juxtacrine signaling) and over longer distances (paracrine signaling). This coordination enables the emergence of many layers of tissue patterning, from precise mosaic patterning between cells to patterning across an entire embryonic axis [19, 34]. In this way, cellular communication drives the complex body plans of all multi-cellular organisms [39]. After development, cellular communication is important for organ homeostasis, in regulating cell turnover by controlling growth and death responses, and preventing the occurrence of cancers and other harmful activities [14]. The remarkably complex processes of cellular regulation require effective systems for cell-cell communication.

The intricate process of communication is orchestrated by signaling pathways: networks of multiple proteins within the cell that sense, interpret, and transmit signals from outside. Signaling pathways govern an array of cellular responses including gene transcription and regulation of metabolism and cell-cycle. Signaling pathways also initiate cell differentiation, pluripotency, proliferation, apoptosis, migration, adhesion, cytoskeletal reorganization, and polarity[39]. With the wide range of functions that must be achieved in many dissimilar multi-cellular organisms, it would not be surprising if hundreds of individualized signaling pathways developed independently to satisfy a diversity of needs. Yet fewer than 20 classes of metazoan signaling pathways have been identified [19], all of which evolved prior to the diversification of modern body plans. Of these 20, seven signal pathways are responsible for most developmental processes – the Hedgehog, Wnt, Tgfβ, ERK, Notch, JAK/STAT, and nuclear hormone pathways [40]. As a result, these signaling pathways must be highly versatile and perform a wide range of functions across many developmental contexts.

Signaling pathways transmit ligand-receptor action to cellular response

The structure of the seven developmental signaling pathways can be broadly generalized to a few key steps and components. These signaling pathways function by transmitting some extracellular signal such as an extracellular protein ligand or chemical into a response within the cell.

This process begins with ligand-specific receptors, which undergo conformational changes and oligomerization following ligand-recognition [31]. There are often multiple receptor-ligand pairs that activate each pathway. In general, receptors span the cell membrane, though some receptors, as in the nuclear hormone receptor pathways, reside within the cell and respond to membrane-permeable factors. Once bound to ligand, receptors begin a series of biochemical reactions within the cell [19].

With the exception of the nuclear hormone pathway, ligand-receptor binding is followed by recruitment of protein factors to the plasma membrane. This recruitment instigates changes to the state and location of various proteins through sequestration, nuclear transport, cleavage, chemical modification, and protein binding. Finally, a transcriptional regulator accumulates in the nucleus. The transcriptional regulator then binds directly to DNA as a transcription factor, or synergizes with another transcription factor to regulate gene expression [19]. In this thesis, I refer to the network of protein interactions that transmits signal from ligand-receptor dynamics into gene transcription as the “core pathway” (Figure 1.1). As the number of ligands and receptors vary significantly across organisms, this thesis focuses on analysis of the core pathway.

The complexity of the core pathway varies significantly between pathways. In the Notch pathway, Notch acts as both the transmembrane receptor and, after being cleaved and imported to the nucleus, as the transcriptional regulator [3]. Similarly, in the Tgfβ and JAK/STAT pathways, the transcriptional regulator (Smad and STAT, respectively), are activated directly at the receptor and then imported [41, 50]. In contrast, the ERK pathway transduces signal through a kinase cascade, with several levels of protein activation separating the receptor from the transcriptional regulator [32].

While the molecular details of each signaling pathway are well understood, it is still unclear what information about ligand-receptor dynamics is carried into the cell.

In order to understand how cells coordinate developmental events, we must first understand how signals are sensed, interpreted, and transmitted through the core signaling pathway.

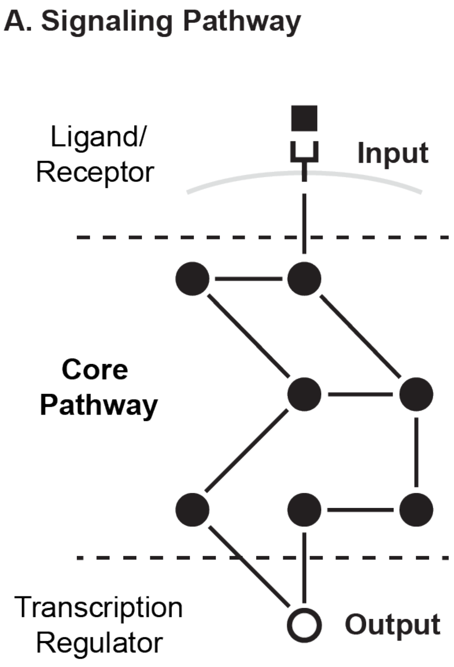


Figure 1.1: **(A)** Signaling pathways transmit inputs from ligand-receptor interaction to a change in output, the level of transcriptional regulator (white circle).

Input-output response is key to understanding signaling pathway function

Signaling pathways have been well-characterized over the last few decades, and the molecular interactions that govern each pathway are largely identified. In addition, quantitative assays of reaction rates and protein levels have enabled the construction of detailed models that capture pathway dynamics [49]. Consequently, the time is ripe to study the input-output response of these signaling pathways. The analyses performed in this thesis will focus on the core pathway as defined in Figure 1.1, with the input as ligand-receptor activation, and the output as the level of transcriptional regulator. Input-output behavior captures how the output of the core pathway responds to various inputs, e.g. sinusoidal inputs (for frequency response) or steady-state inputs (for static response). This allows us to abstract the behavior of a signaling pathway away from the underlying molecular details [15].

Studying input-output response serves several purposes. First, input-output response provides insight into signal processing performed by pathways. Rather than produce output in faithful representation of the input, signal processing involves the manipulation of the input signal. For instance, signal processing includes amplification, filtering, and applying threshholds to the input signal [44]. Each of these signal processing capabilities may confer benefits to cell signaling. In the next section, I examine specific examples of input-output response, the mechanism by which they arise, and the benefits they may confer.

Second, studying input-output response enables a modular approach to tackling large networks. Modularity underlies many biological processes, and in particular enables the study of signaling pathways in isolation from other cellular processes. Once we characterize the input-output response of a pathway, however, we can examine its behavior within larger inter-connected systems [15]. In Chapter 4, I utilize this modular approach to examine a transcriptional network downstream of the Wnt pathway. In this case, the output of the core Wnt pathway connects to the input of the transcriptional network.

In the past several decades, researchers have studied input-output behavior within many signaling pathways, uncovering a variety of responses to signal. These include responses within the receptor module, through the core pathway, and at the gene transcriptional level. Below, I discuss several examples of input-output response in cell signaling, as well as the mechanisms by which the input-output responses are generated and the benefits they may confer. These responses include logarithmic computation, ultrasensitivity, hysteresis, and non-monotonic response to input. Notably, all of these responses involve nonlinear processing of input and indicate that a broad array of signaling strategies are employed in biology. By contrast, in Chapters 2-3, I demonstrate linear input-output response in three distinct signaling pathways, suggesting that linearity is a pervasive behavior of cell signaling.

Many nonlinear input-output responses are found in in cell signaling

Logarithmic computations can arise from protein allostery. For an allosteric protein, the protein's active site is regulated by a distal effector [36]. Allostery is widespread, appearing in signaling, metabolism, and cell-cycle regulation[10]. The thermodynamic basis for conformational selection in allosteric proteins can be captured by the Moneaud-Wyman-Changeaux (MWC) model [36]. In this model, a protein can switch between two or more states through conformational change. For protein kinases, this conformation change often occurs at the substrate binding site. Conformational selection is achieved through binding of one or more effectors, which reinforce the protein's existing conformation. For instance, in the *E. coli* chemotaxis network, the aspartate receptor Tar can switch the sensor protein CheA from an active kinase to inactive. In particular, the inactive and active states are in thermodynamic equilibrium, while chemoattractant binding reinforces the inactive state. Thus, aspartate can provide conformational selection of CheA, with bound receptor complex less likely to transition to the active state than unbound receptor complex [30].

How does the architecture of MWC proteins affect input-output response? The conformational selection provided by one or more effectors responds nonlinearly, since each effector further decreases the probability of conformational switching. As shown by Olsman and Goentoro, this architecture is capable of producing logarithmic sensitivity to ligand [37]. As a result, the MWC model produces linear changes in receptor activity in response to logarithmic changes in effector concentration. Interestingly, this relates to Weber's law of sensory detection that describes systems that are sensitive to relative changes in signal [37]. Indeed, logarithmic sensing by the allosteric Tar receptor produces Weber's law detection in the bacterial chemotaxis network [46].

Cascades of protein kinases, in which each kinase activates a subsequent kinase, can produce highly ultrasensitive response to input. Ultrasensitivity is defined by Goldbeter and Koshland as an input-output response that requires less than an 81-fold increase in input to drive the output from 10% to 90% of maximum [21]. Generally, an ultrasensitive response is one that is sigmoidal, having an S-shaped response profile. Sigmoidal response arises through cooperativity in allosteric proteins, enzyme saturation for two-state kinase/phosphatase substrates, and multi-site distributive phosphorylation (one site phosphorylated per enzyme-substrate collision). Protein cascades amplify the sigmoidal response of each layer, producing extremely sharp response to stimulus [26].

Ultrasensitivity was demonstrated for the ERK pathway response to progesterone in *Xenopus* oocytes [16]. The ERK pathway is activated through a kinase cascade consisting MAPKKK (e.g. Raf-1, B-Raf, Mos), MAPKK (e.g. MEK1, MEK2), and MAPK (e.g. ERK1, ERK2). Huang and Ferrell demonstrated through modeling and experiments that that saturation of MEK and ERK phosphatases, combined with distributive multi-site phosphorylation, produces ultrasensitive response to stimulus [26]. Indeed, in single *Xenopus* oocytes, the hill coefficient of ERK pathway response was experimentally determined at $n_h > 42$ [16]. The highly ultrasensitive response of the ERK pathway serves two purposes. First, within the responsive (steep) region of the input-output curve, small changes in signal are amplified to large changes in output. This effect is enhanced by the concentration of each kinase within the ERK cascade: the first kinase is expressed at nanomolar concentration, while the latter kinases are expressed at micromolar concentrations, causing a 1000-fold increase in signal amplification [43]. Second, for inputs outside the responsive region of the input-output curve, ultrasensitivity produces all-or-none response. Small inputs produce no output, while large inputs produce maximal output, yielding a sharp, threshhold response to input [16].

Positive feedback, in which system's output feeds back and reinforces the state of system, can generate hysteretic response to input [33]. Positive feedback can be achieved either through a cycle of activation, or a cycle of inactivation (double-negative feedback). This topology can create a bistable response to input, meaning that an input can produce two distinct, stable outputs. The output not only depends on the current input, but the history of the system as well. As a result, if an increase in input shifts the system from a low-output to high-output state, a similar decrease in input may not be sufficient to bring the system back to the low-output state. A much larger decrease in input may be needed to return the system to its original low-output state. This input lag is referred to as hysteresis, and was first used to describe the behavior of magnetic materials [12].

In the Sonic Hedgehog (Ssh) pathway, there is evidence of hysteretic response to Ssh ligand. In developing mouse limbs, both current and previous exposure to Ssh determined digit identity, indicating a history-dependence to Ssh signaling [23]. Similarly, Ssh target genes continue to be expressed even after decrease in ligand level [6, 27]. Bistability and hysteresis were initially characterized as arising from a positive feedback from Gli transcription factors activating their own expression. This was thought to reinforce the activity of Gli transcription factors even after Ssh signaling had been removed. However, further evidence supports a model in which hysteresis in Ssh signaling arises from the gene regulatory network governing Ssh target genes, downstream of the Ssh pathway [6, 11]. In this model, there is mutual repression between genes expressed under high Ssh signaling, such as *Nkx2.2*, and genes expressed in the absence of Ssh, such as *Pax6*. This double-negative produces an effective positive feedback in Ssh target gene expression, and a hysteretic transcriptional response to Ssh. In this context, bistability can generate sharp response to Ssh input, and hysteresis provides a memory of past signaling events [11].

Lastly, the incoherent feed-forward loop (IFFL), a paradoxical system that both activates and inhibits its output, can produce non-monotonic response to input [28]. The IFFL is a recurrent motif in cell biology, appearing in signaling pathways, transcriptional networks, and immune regulation [1]. For instance, in the galactose system in *E. coli*, cAMP-responsive CRP has two roles in regulating the transcription of the *galE* gene. First, it directly binds to the promoter region and recruits RNA polymerase. Second, it upregulates a repressor *galS*, which represses the transcription of the *galE* gene. Thus, CRP has both a positive and negative effect on *galE* transcription. This IFFL architecture can produce numerous input-output responses, depending on parameters and timescales [24].

In the case of the *galE* transcriptional circuit, Kaplan et al. demonstrated that the IFFL increases *galE* expression over a range of low cAMP levels, while decreasing *galE* expression for higher cAMP levels. This is a type of non-monotonic response, in which the output increases over one interval of input while decreasing over another interval. In the gal IFFL circuit, non-monotonic response to input is generated by differential binding to the *galE* and *galS* promoters. When cAMP-activated CRP is at low levels, CRP binds strongly to the *galE* promoter and only weakly to the *galS* promoter, leading to *galE* expression. However, higher levels of activated CRP drive binding onto the *galS* promoter, upregulating the repressor of *galE* and decreasing *galE* expression. Thus, the IFFL motif can act as a bi-phasic amplitude filter, filtering out low and high inputs and only responding to intermediate cAMP inputs [28].

Do signaling pathways exhibit linear input-output response?

While nonlinear input-output response is utilized in many signaling contexts, there is evidence that linear signal transmission may be a widespread behavior of cell signaling. Linear input-output response entails that each frequency component of the input signal be faithfully transmitted into output, such that only the amplitude and phase are affected. For static inputs, this means that the output is equal to the input scaled by some multiple, or gain G. There are several reasons to expect linear response in cell signaling.

One benefit of linearity is superposition, a technique employed in engineered linear systems. By definition, a linear system satisfies the properties of homogeneity and additivity. Homogeneity is satisfied if scaling an input results in a correspondingly scaled output. Additivity is satisfied if the output for two combined inputs is the summed output for the two inputs alone. Therefore, a linear system can transmit multiple signals simultaneously without distortion, a property referred to as superposition [44]. Superposition is what enables a car radio to tune to a single radio station, and for many houses to share the same telephone line. By the same merits we find for linearity in engineered systems, we might expect to find linearity in biological systems as well. I propose that linearity could increase the number of signaling events transmitted through the limited set of pathways in cells. Multiple ligand-receptor complexes often feed into one pathway (e.g. all growth factors into the ERK pathway [29]), and so superposition could allow several ligands to activate the pathway at once, with gene regulatory networks then deconvolving the pathway output into the response from each ligand alone.

Second, input-output linearity increases the information capacity of a signaling pathway as compared to nonlinear input-output responses [51]. Noise is an inescapable aspect of signal transmission and limits the number of inputs that can be distinguished by measuring output. For nonlinear input-output responses, the pathway compresses inputs where the response slope is shallow and amplifies noise when the response slope is steep, further reducing the number of distinguishable inputs. In contrast, information loss due to noise is minimized for linear input-output response, maximizing the number of distinguishable inputs. In this context, linearity is also referred to as "dose response alignment," meaning that the dose response curves of bound receptor and downstream output are aligned [51].

Dose response alignment was observed in several hormone signaling pathways [2, 13], though the mechanisms for dose-alignment are unknown. Dose response alignment was observed for the MAPK Fus3 in the yeast pheromone response system as well, using a feedback mechanism not conserved in metazoan MAPK/ERK signaling [4, 51]. There is also evidence of linearity between ligand and the time-integrated activity of the epidermal growth factor and erythropoietin receptor modules [7, 38].

Motivated by this, I studied multiple metazoan signaling pathways to determine whether linearity is a more widespread strategy of cell signaling. Specifically, I chose the Wnt, ERK and Tgfβ pathways for two reasons: one, these signaling pathways represent three of the seven classes of early developmental signaling pathways, and two, their core pathways are characterized by detailed mathematical models. Importantly, these models have predicted a wide range of pathway behaviors over the years (e.g. Wnt refs. [20, 25, 35]; ERK refs. [16, 18, 26, 43, 45]; Tgfβ refs.

[5, 17, 22, 42, 47, 48]). I used these models to analytically derive the pathways' input-output responses. I then validated key findings experimentally using quantitative Western blots.

In Chapter 2, I demonstrate that despite the distinct architectures of the Wnt, ERK, and Tgfβ pathways, these three highly-conserved developmental signaling pathways operate using linear signal transmission. I show that linearity arises in the Wnt pathway due to a futile cycle of β-catenin synthesis and degradation, a kinase cascade coupled to negative feedback in the ERK pathway, and rapid nucleocytoplasmic shutting of Smad complex in the Tgfβ pathway. In Chapter 3, I validate these findings with direct measurements of input-output response in the Wnt and ERK pathway using quantitative Western blot. In the Wnt pathway, experiments indicate that β-catenin accumulates linearly with active Wnt co-receptor LRP and that this linearity persists throughout the dynamic range of LRP activation. In the ERK pathway, experiments indicate that doubly-phosphorylated ERK increases linearly with EGF dose, until saturation of doubly-phosphorylated ERK. I also demonstrate that specific perturbations can reduce the range of linear response of each pathway. Finally, in Chapter 4, I explore an additional benefit of linearity in the context of the Wnt pathway. Modeling the Wnt pathway, I find that linearity produces robust fold-changes in β-catenin with respect to parameter variations. Therefore, this thesis demonstrates that linearity is a widespread property of cell signaling, and introduces a novel benefit for linear signal transmission in biology.

BIBLIOGRAPHY

[1] Uri Alon. *An introduction to systems biology: design principles of biological circuits*. CRC press, 2006. ISBN 1584886420.

[2] Syed M Amir, Terry F Carraway, Leonard D Kohn, and Roger J Winand. The binding of thyrotropin to isolated bovine thyroid plasma membranes. *Journal of Biological Chemistry*, 248(11):4092–4100, 1973.

[3] Emma R Andersson, Rickard Sandberg, and Urban Lendahl. Notch signaling: simplicity in design, versatility in function. *Development*, 138(17):3593–3612, 2011.

[4] Steven S. Andrews, William J. Peria, C. Yu Richard, Alejandro Colman-Lerner, and Roger Brent. Push-pull and feedback mechanisms can align signaling system outputs with inputs. *Cell systems*, 3(5):444–455. e2, 2016. ISSN 2405-4712.

[5] Geoffroy Andrieux, Laurent Fattet, Michel Le Borgne, Ruth Rimokh, and Nathalie Théret. Dynamic regulation of tgf-b signaling by tif1γ: a computational approach. *PloS one*, 7(3):e33761, 2012. ISSN 1932-6203.

[6] Nikolaos Balaskas, Ana Ribeiro, Jasmina Panovska, Eric Dessaud, Noriaki Sasai, Karen M Page, James Briscoe, and Vanessa Ribes. Gene regulatory logic for reading the sonic hedgehog signaling gradient in the vertebrate neural tube. *Cell*, 148(1):273–284, 2012.

[7] Verena Becker, Marcel Schilling, Julie Bachmann, Ute Baumann, Andreas Raue, Thomas Maiwald, Jens Timmer, and Ursula Klingmüller. Covering a broad dynamic range: information processing at the erythropoietin receptor. *Science*, 328(5984):1404–1408, 2010. ISSN 0036-8075.

[8] Eva Bianconi, Allison Piovesan, Federica Facchin, Alina Beraudi, Raffaella Casadei, Flavia Frabetti, Lorenza Vitale, Maria Chiara Pelleri, Simone Tassani, Francesco Piva, et al. An estimation of the number of cells in the human body. *Annals of human biology*, 40(6):463–471, 2013.

[9] Sean B Carroll. Chance and necessity: the evolution of morphological complexity and diversity. *Nature*, 409(6823):1102, 2001.

[10] Jean-Pierre Changeux. Allostery and the monod-wyman-changeux model after 50 years. *Annual review of biophysics*, 41:103–133, 2012.

[11] Michael Cohen, Karen M Page, Ruben Perez-Carrasco, Chris P Barnes, and James Briscoe. A theoretical framework for the regulation of shh morphogen-controlled gene expression. *Development*, 141(20):3868–3878, 2014.

[12] Michael Cohen, Karen M Page, Ruben Perez-Carrasco, Chris P Barnes, and James Briscoe. Mathematical models help explain experimental data. response to ‘transcriptional interpretation of shh morphogen signaling: computational modeling validates empirically established models’. *Development*, 143(10): 1640–1643, 2016.

[13] Pedro Cuatrecasas. Insulin-receptor interactions in adipose tissue cells: direct measurement and properties. *Proceedings of the National Academy of Sciences*, 68(6):1264–1268, 1971.

[14] Klaus-Michael Debatin. Apoptosis pathways in cancer and cancer therapy. *Cancer Immunology, Immunotherapy*, 53(3):153–159, 2004.

[15] Domitilla Del Vecchio and Richard M Murray. *Biomolecular feedback systems*. Princeton University Press Princeton, NJ, 2015.

[16] James E. Ferrell and Eric M. Machleder. The biochemical basis of an all-or-none cell fate switch in xenopus oocytes. *Science*, 280(5365):895–898, 1998. ISSN 0036-8075.

[17] Christopher L. Frick, Clare Yarka, Harry Nunns, and Lea Goentoro. Sensing relative signal in the tgf-β/smad pathway. *Proceedings of the National Academy of Sciences*, page 201611428, 2017. ISSN 0027-8424.

[18] Raphaela Fritsche-Guenther, Franziska Witzel, Anja Sieber, Ricarda Herr, Nadine Schmidt, Sandra Braun, Tilman Brummer, Christine Sers, and Nils Blüthgen. Strong negative feedback from erk to raf confers robustness to mapk signalling. *Molecular systems biology*, 7(1):489, 2011. ISSN 1744-4292.

[19] John Gerhart. 1998 warkany lecture: signaling pathways in development. *Teratology*, 60(4):226–239, 1999. ISSN 0040-3709.

[20] Lea Goentoro and Marc W. Kirschner. Evidence that fold-change, and not absolute level, of β-catenin dictates wnt signaling. *Molecular cell*, 36(5): 872–884, 2009. ISSN 1097-2765.

[21] Albert Goldbeter and Daniel E Koshland. An amplified sensitivity arising from covalent modification in biological systems. *Proceedings of the National Academy of Sciences*, 78(11):6840–6844, 1981.

[22] Vinicio González-Pérez, Bernhard Schmierer, Caroline S. Hill, and Richard P. Sear. Studying smad2 intranuclear diffusion dynamics by mathematical modelling of frap experiments. *Integrative Biology*, 3(3):197–207, 2011.

[23] Brian D Harfe, Paul J Scherz, Sahar Nissim, Hua Tian, Andrew P McMahon, and Clifford J Tabin. Evidence for an expansion-based temporal shh gradient in specifying vertebrate digit identities. *Cell*, 118(4):517–528, 2004.

[24] Yuval Hart and Uri Alon. The utility of paradoxical components in biological circuits. *Molecular cell*, 49(2):213–221, 2013.

[25] Ana R. Hernández, Allon M. Klein, and Marc W. Kirschner. Kinetic responses of β-catenin specify the sites of wnt control. *Science*, 338(6112):1337–1340, 2012. ISSN 0036-8075.

[26] Chi-Ying Huang and James E. Ferrell. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Sciences*, 93 (19):10078–10083, 1996. ISSN 0027-8424.

[27] Jan Philipp Junker, Kevin A Peterson, Yuichi Nishi, Junhao Mao, Andrew P McMahon, and Alexander van Oudenaarden. A predictive model of bifunctional transcription factor signaling during embryonic tissue patterning. *Developmental cell*, 31(4):448–460, 2014.

[28] Shai Kaplan, Anat Bren, Erez Dekel, and Uri Alon. The incoherent feed-forward loop can generate non-monotonic input functions for genes. *Molecular systems biology*, 4(1):203, 2008.

[29] Menachem Katz, Ido Amit, and Yosef Yarden. Regulation of mapks by growth factors and receptor tyrosine kinases. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1773(8):1161–1176, 2007.

[30] Juan E Keymer, Robert G Endres, Monica Skoge, Yigal Meir, and Ned S Wingreen. Chemosensing in escherichia coli: two regimes of two-state receptors. *Proceedings of the National Academy of Sciences*, 103(6):1786–1791, 2006.

[31] Edda Klipp and Wolfram Liebermeister. Mathematical modeling of intracellular signaling pathways. *BMC neuroscience*, 7(1):S10, 2006.

[32] Walter Kolch. Coordinating erk/mapk signalling through scaffolds and inhibitors. *Nature reviews Molecular cell biology*, 6(11):827–837, 2005. ISSN 1471-0072.

[33] Karen Lai, Matthew J Robertson, and David V Schaffer. The sonic hedgehog signaling system as a bistable genetic switch. *Biophysical Journal*, 86(5): 2748–2757, 2004.

[34] Pamela J Lanford, Yu Lan, Rulang Jiang, Claire Lindsell, Gerry Weinmaster, Thomas Gridley, and Matthew W Kelley. Notch signalling pathway mediates hair cell development in mammalian cochlea. *Nature genetics*, 21(3):289, 1999.

[35] Ethan Lee, Adrian Salic, Roland Krüger, Reinhart Heinrich, and Marc W. Kirschner. The roles of apc and axin derived from experimental and theoretical analysis of the wnt pathway. *PLoS Biol*, 1(1):e10, 2003. ISSN 1545-7885.

[36] Jacques Monod, Jeffries Wyman, and Jean-Pierre Changeux. On the nature of allosteric transitions: a plausible model. *Biol*, 12:228–113, 1965.

[37] Noah Olsman and Lea Goentoro. Allosteric proteins as logarithmic sensors. *Proceedings of the National Academy of Sciences*, 113(30):E4423–E4430, 2016. ISSN 0027-8424.

[38] Diego A. Oyarzún, Jo L. Bramhall, Fernando López-Caamal, Frances M. Richards, Duncan I. Jodrell, and Ben-Fillippo Krippendorff. The egfr demonstrates linear signal transmission. *Integrative Biology*, 6(8):736–742, 2014.

[39] Norbert Perrimon, Chrysoula Pitsouli, and Ben-Zion Shilo. Signaling mechanisms controlling cell fate and embryonic patterning. *Cold Spring Harbor perspectives in biology*, 4(8):a005975, 2012.

[40] André Pires-daSilva and Ralf J. Sommer. The evolution of signalling pathways in animal development. *Nature Reviews Genetics*, 4(1):39–49, 2003. ISSN 1471-0056.

[41] Jason S Rawlings, Kristin M Rosler, and Douglas A Harrison. The jak/stat signaling pathway. *Journal of cell science*, 117(8):1281–1283, 2004.

[42] Bernhard Schmierer, Alexander L. Tournier, Paul A. Bates, and Caroline S. Hill. Mathematical modeling identifies smad nucleocytoplasmic shuttling as a dynamic signal-interpreting system. *Proceedings of the National Academy of Sciences*, 105(18):6608–6613, 2008. ISSN 0027-8424.

[43] Birgit Schoeberl, Claudia Eichler-Jonsson, Ernst Dieter Gilles, and Gertraud Müller. Computational modeling of the dynamics of the map kinase cascade activated by surface and internalized egf receptors. *Nature biotechnology*, 20 (4):370–375, 2002. ISSN 1087-0156.

[44] Steven Smith. *Digital signal processing: a practical guide for engineers and scientists*. Elsevier, 2013.

[45] O. E. Sturm, R. Orton, J. Grindlay, M. Birtwistle, V. Vyshemirsky, D. Gilbert, M. Calder, A. Pitt, B. Kholodenko, and W. Kolch. The mammalian mapk/erk pathway exhibits properties of a negative feedback amplifier. *Science signaling*, 3(153):ra90, 2010. ISSN 1937-9145.

[46] Yuhai Tu, Thomas S Shimizu, and Howard C Berg. Modeling the chemotactic response of escherichia coli to time-varying stimuli. *Proceedings of the National Academy of Sciences*, 105(39):14855–14860, 2008.

[47] P. Vizán, D. S. Miller, I. Gori, D. Das, B. Schmierer, and C. S. Hill. Controlling long-term signaling: receptor dynamics determine attenuation and refractory behavior of the tgf-β pathway. *Science signaling*, 6(305):ra106, 2013. ISSN 1937-9145.

[48] Junjie Wang, Lisa Tucker-Kellogg, Inn Chuan Ng, Ruirui Jia, P. S. Thiagarajan, Jacob K. White, and Hanry Yu. The self-limiting dynamics of tgf-β signaling in silico and in vitro, with negative feedback through ppm1a upregulation. *PLoS computational biology*, 10(6):e1003573, 2014. ISSN 1553-7358.

[49] Joseph A Wayman and Jeffrey D Varner. Biological systems modeling of metabolic and signaling networks. *Current Opinion in Chemical Engineering*, 2(4):365–372, 2013.

[50] Mary Y. Wu and Caroline S. Hill. Tgf-β superfamily signaling in embryonic development and homeostasis. *Developmental cell*, 16(3):329–343, 2009. ISSN 1534-5807.

[51] C. Richard Yu, C. Gustavo Pesce, Alejandro Colman-Lerner, Larry Lok, David Pincus, Eduard Serra, Mark Holl, Kirsten Benjamin, Andrew Gordon, and Roger Brent. Negative feedback that improves information transmission in yeast signalling. *Nature*, 456(7223):755–761, 2008. ISSN 0028-0836.

Chapter 2

LINEARITY IN CELL SIGNALING PATHWAYS: ANALYTICS

2.1 Introduction

Insights into the versatility of signaling pathways may be gleaned from pathway architectures. As discussed in Chapter 1, studies over the past several decades have revealed many signaling capabilities that arise from distinct pathway architectures, e.g., logarithmic response [35], all-or-none response [13, 21], hysteresis [5, 7, 36], and non-monotonic response [23]. Alternatively, analysis of pathway architectures may also reveal shared signaling capabilities that emerge from distinct architectures, pointing to a fundamental property that pathways have converged upon despite their separate evolutionary trajectories. In this study, we sought to identify shared properties between conserved signaling pathways.

To approach this question, we use established math models of signaling pathways. As the number of known signaling pathway components and interactions continues to grow, our intuition becomes less capable of understanding the myriad interactions that occur with each signaling event. Conceptual models of linear signaling cascades become quickly muddied by feedbacks, cross-talk, retroactivity, and other nonlinear effects [9, 14, 46]. To this end, mathematical models are a necessary tool in making progress in the study of signaling pathways [2]. Mathematical models can incorporate large sets of known interactions and provide testable predictions for the unknown. In our case, we can use analytical tools to approach these pathways to uncover dominant architectures within the pathway and how this influences pathway behavior.

We examined three signaling pathways, the canonical Wnt, ERK, and Tgfβ pathways, since the architectures of the three pathways are captured by mathematical models that have been refined by years of experiments [28, 40, 44]. These models capture the salient features of each pathway and include biochemical details such as synthesis, degradation, binding, dissociation, and post-translational modifications. In all the models, biochemical parameters have been directly measured or fitted to kinetic measurements from cell, embryo, or extract systems. Although by no means complete, the mathematical models have track records of success in predicting systems-level behaviors across multiple biological systems. For instance, the

Wnt model [28] predicted robustness in fold-change response [17] and the differential roles of the two scaffolds in the pathway [28]; the ERK model [12, 21, 41, 44] captures the ultrasensitivity in the phosphorylation cascade [21]; and the Tgfβ model [40] reveals the roles of nucleocytoplasmic shuttling in transducing the duration and intensity of ligand stimulation [40].

The Wnt, ERK, and Tgfβ pathways transmit input using different core transmission architecture (Figure 2.1A-C). In the Wnt pathway, signal transmission is characterized by a futile cycle of synthesis and rapid degradation [20, 24, 37]. Ligand-receptor input diminishes the degradation arm of this cycle, leading to accumulation of β-catenin output [24, 34, 43]. In the ERK pathway, signal transmission is characterized by a cascade of phosphorylation events coupled to feedbacks, leading to an increase in phosphorylated ERK output [4, 25, 27, 50]. Finally, signal transmission in the Tgfβ pathway is characterized by continual nucleocytoplasmic protein shuttling [22, 32, 33, 39, 48]. Ligand-receptor input effectively increases the rate of nuclear import, leading to an increase in output, known as the nuclear Smad complex [40].

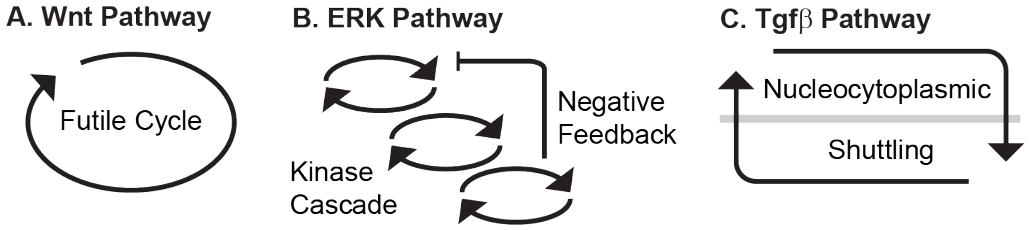


Figure 2.1: The core pathway for each metazoan signaling pathway is defined by distinct architectural features. In the Wnt pathway **(A)**, the output is regulated by a futile cycle of continual synthesis and rapid degradation. In the ERK pathway **(B)**, the output is regulated by a kinase cascade coupled to negative feedback. In the Tgfβ pathway **(C)**, the output is regulated through continual nucleocytoplasmic shuttling.

We studied these mathematical models to identify what, if any, behaviors converge across pathways. The Wnt [28], ERK [44], and Tgfβ [40] models consist of 7, 26, and 10 coupled, nonlinear ODEs, respectively, with 22, 46, and 13 parameters. Because of their large sizes, they are typically solved numerically to simulate experimental observations and generate new predictions. However, for the questions posed here, we found that numerical simulations are not sufficient. Rather, we needed analytics to uncover exactly how the pathway behaviors depend on the underlying biochemical processes.

Analytic treatment has often offered deeper insight than simulations alone - for instance, this approach was used to identify integral feedback in the bacterial chemotaxis network [49], and robustness in the bacterial osmolarity network [42] and in the Wnt pathway [17]. While an analytical solution was previously derived for the Wnt pathway [17], analytical treatment of the Tgfβ and ERK pathways has not been attempted due to the complex, nonlinear equations involved. To address this problem, we employed various analytical techniques, including graph theory-based variable elimination and dimensional analysis, to derive analytical or semi-analytical solutions to the steady-state output of each pathway. Our analysis, along with subsequent experimental verification, reveals a striking convergence across the Wnt, Tgfβ, and ERK pathways: cells operate in the parameter regime where the complex, nonlinear interactions in each pathway give rise to linear signal transmission. Below, we describe our analysis of each pathway and the unifying behavior that emerges from all three pathways.

2.2 Results

Mathematical analysis reveals linear signal transmission in the canonical Wnt pathway

In the Wnt pathway, cells sense ligand-receptor input by monitoring β-catenin protein [6, 24, 30, 34, 43]. β-catenin is continually synthesized and rapidly degraded by a large destruction complex, comprised of multiple proteins including APC, Axin, and GSK3β. The destruction complex binds and phosphorylates β-catenin, tagging it for degradation by the ubiquitin/proteosome machinery [24, 43]. Wnt ligands, through binding to Frizzled and LRP receptors, inhibit the destruction complex, leading to accumulation of β-catenin. β-catenin then regulates the expression of broad target genes [34, 43].

The model of the Wnt pathway (Figure 2.2A) was published in 2003 by a collaboration between the Kirschner and Heinrich labs [28]. The Wnt model consists of 7 nonlinear differential equations and 22 parameters. Applying dimensional analysis, we previously derived the analytical solution to β-catenin concentration at steady-state [17]:

$$[\beta cat]_{ss} = K_{17} \cdot \frac{1-\gamma+\alpha/u}{2}\left(\sqrt{(1+\frac{4\gamma}{(1-\gamma+\alpha/u)^2}} - 1\right) \tag{2.1}$$

$$\alpha = \frac{k_4 k_6 k_9 v_{14} \cdot GSK3_{tot} \cdot APC_{tot}}{k_5 k_{-6} K_7 K_8 k_{13} k_{15}} \tag{2.2}$$

$$\gamma = \frac{v_{12}}{k_{13}K_{17}} \tag{2.3}$$

where the input function $u = u(Wnt)$ is the rate of inhibition of the destruction complex via Dishevelled, a function of ligand-receptor activation. As illustrated in Figure 2.2A, K_is are equilibrium dissociation constants, k_is are rate constants, and v_is are synthesis rates. α and γ in Equation 2.1 are dimensionless parameter groups defined in Eqs. 2.2 and 2.3: α characterizes β-catenin degradation by the destruction complex, and γ characterizes the extent to which β-catenin binds to APC independently of the destruction complex.

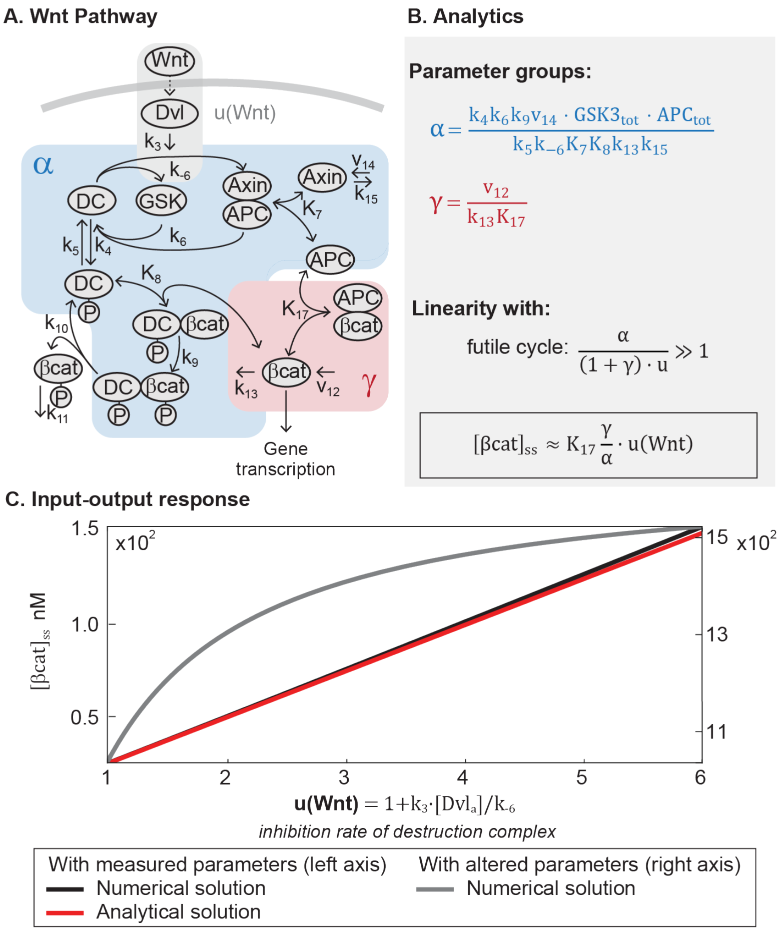


Figure 2.2: Mathematical analysis reveals linear signal transmission in the Wnt pathway. **(A)** Network diagram of the Wnt pathway. **(B)** The parameter groups and input-output relationship that were derived in this study. Parameter groups and input function are color-coded to the corresponding reactions in the network diagram. Parameters that do not appear in the parameter groups drop out due to irreversible reaction steps (such as $k_1 0$ and $k_1 1$). **(C)** The input-output relationship of the Wnt pathway. The output is β-catenin, and the input function $u(Wnt)$ is the rate by which Dishevelled inhibits the destruction complex upon Wnt ligand activation, where k_3 and k_{-6} are defined in the figure, and $[Dvl]_a$ is the concentration Wnt-activated Dishevelled.

Equation 2.1 demonstrates that, in general, β-catenin concentration is a nonlinear function of the input u. Many parameters of the model were directly measured in Xenopus extracts, and the remaining calculated from measurements in the same system. In this study, we examined how the analytical solution (Equation 2.1) behaves with these measured parameters. The measured parameters indicate that $\alpha \sim 66$, $\gamma \sim 1.4$, and for maximal stimulation, $u \sim 6.0$. The large α reflects how β-catenin stability is primarily dictated by the destruction complex, i.e., $\alpha/u \gg 1$

means that non-Axin dependent degradation is minimal, and $\alpha/u \gg \gamma$ means that the positive feedback from sequestration by APC is minimal. Indeed, the rapid action of the destruction complex in the Wnt pathway is a recurring observation across biological systems [20, 24, 37]. With $\alpha/u \gg 1 + \gamma$, Equation 2.1 simplifies to

$$[\beta cat]_{ss} \approx K_{17} \frac{\gamma}{\alpha} u \tag{2.4}$$

with detailed derivations presented in the SI. Therefore, within physiologically relevant parameter values, the steady-state β-catenin concentration becomes a linear function of the input u (red line, Figure 2.2C). The linear input-output relationship holds for the entire dynamic range of the model, until the system saturates at maximal stimulation ($u \sim 6.0$). We confirmed that the numerical solution of the full model matches the analytical solution in Equation 2.4 (blue line, Figure 2.2C), and that the response becomes nonlinear when α is decreased, breaking the requirement $\alpha/u \gg 1 + \gamma$ (grey line, Figure 2.2C).

Mathematical analysis reveals linear signal transmission in the MAPK/ERK pathway

The unexpected linearity that emerges from the model of the Wnt pathway prompted us to wonder if such simplicity may be found in other pathways. Strikingly, we observed the same linearity in the ERK and Tgfβ pathways. In the ERK pathway (Figure 2.3A), ligand-receptor input is transmitted via a cascade of protein phosphorylation [25, 50]. In particular, ligand-receptor interactions activate Ras, which leads to membrane recruitment and phosphorylation of Raf. Phosphorylated Raf subsequently doubly-phosphorylates MEK, which in turn doubly-phosphorylates ERK [25].

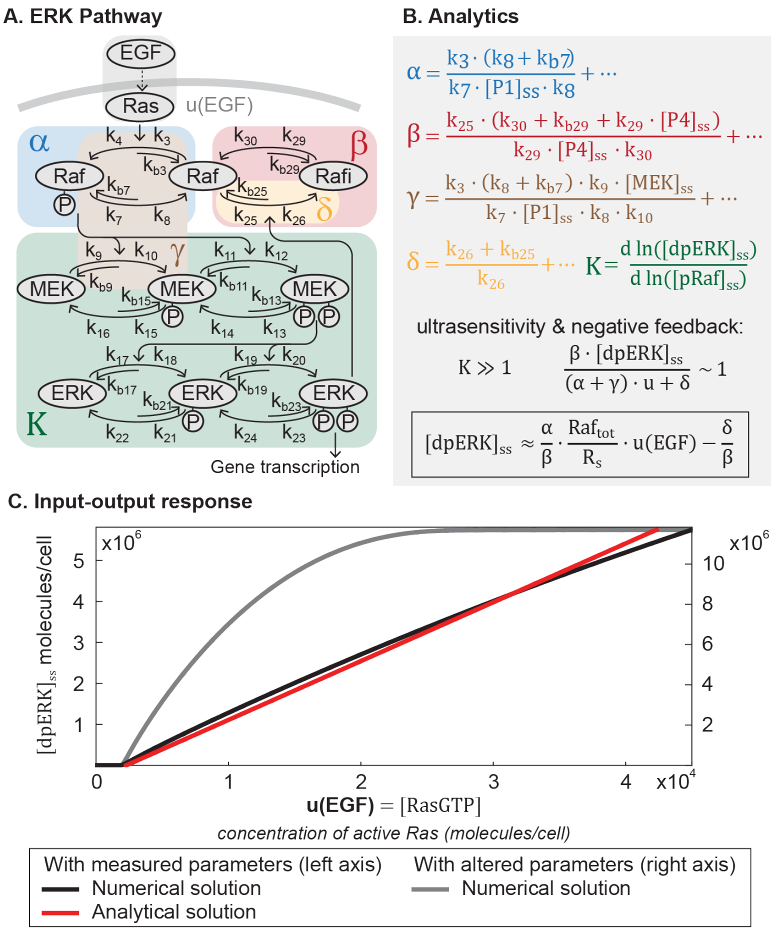


Figure 2.3: Mathematical analysis reveals linear signal transmission in the ERK pathway. **(A)** Network diagram of the ERK pathway. **(B)** The parameter groups and input-output relationship that were derived in this study. Parameter groups and input function are color-coded to the corresponding reactions in the network diagram. Parameters that do not appear in the parameter groups are negligible (as indicated by ellipses). **(C)** The input-output relationship of the ERK pathway. The output is dpERK, and the input function u is the concentration of EGF-activated Ras (Ras-GTP).

Doubly-phosphorylated ERK (dpERK) is a transcriptional regulator that affects a broad array of genes [50]. The multi-step topology of the kinase cascade, combined with distributive phosphorylation of each kinase, gives rise to ultrasensitivity – first demonstrated in the seminal work by the Ferrell lab [13, 21]. In other contexts, the pathway also exhibits a graded response [1, 8, 31, 47] that is thought to arise from the incorporation of negative feedbacks [27, 38], one of which is the inhibition of Raf by dpERK through hyper-phosphorylation of serine residues [10, 18, 44].

The ERK model [44] is the product of more than two decades of refinement

[13, 15, 21, 41, 44]. The model, which captures ultrasensitivity and Raf feedback, consists of 26 differential equations and 46 parameters. To derive an analytical expression for the ERK pathway, we used a variable elimination technique developed for networks of mass action kinetics [11]. The technique utilizes an algebraic framework, linear elimination of variables, and mass conservation laws to parameterize steady-state in terms of core variables (described in SI). We derived an analytical relationship between the steady-state output of the pathway (dpERK) and the input to the phosphorylation cascade u:

$$[dpERK]_{ss} = \frac{\alpha}{\beta} \cdot \left(\frac{Raf_{tot}}{[pRaf]_{ss}} - 1 - \frac{\gamma}{\alpha} \right) \cdot u - \frac{\delta}{\beta} \tag{2.5}$$

$$\alpha = \frac{k_3 \cdot (k_8 + k_b 7)}{k_7 \cdot [P1]_{ss} \cdot k_8} + \dots \tag{2.6}$$

$$\beta = \frac{k_{25} \cdot (k_{30} + k_{b29} + k_{29} \cdot [P4]_{ss})}{k_{29} \cdot [P4]_{ss} \cdot k_{30}} + \dots \tag{2.7}$$

$$\gamma = \frac{k_3 \cdot (k_8 + k_{b7}) \cdot k_9 \cdot [MEK]_{ss}}{k_7 \cdot [P1]_{ss} \cdot k_8 \cdot k_{10}} + \dots \tag{2.8}$$

$$\delta = \frac{k_{26} + k_{b25}}{k_{26}} + \dots \tag{2.9}$$

Detailed derivations of Equation 2.5 are presented in the SI. The input $u = u(EGF)$ in Equation 2.5 is the concentration of active Ras, which is activated via GTP loading at the ligand-receptor complex [25]. The parameter groups α, β, γ, and δ in Equation 2.5 are defined in Eqs. 2.6-2.9, where the ellipses indicate additional small terms (expanded in SI). The relative magnitudes of α, β, γ, and δ indicate how the Raf pool partitions during signaling (Eqs. S3.3, 11-13). The dimensionless group $\alpha \cdot u$ relates to the amount of free, phosphorylated Raf (α, blue-shaded in Figure 2.3A), $\beta \cdot [dpERK]_{ss}$ describes the amount of Raf inhibited through negative feedback by dpERK (β, red-shaded in Figure 2.3A), δ relates to the amount of unphosphorylated (δ, blue-shaded in Figure 2.3A), and $\gamma \cdot u$ relates to the amount of phosphorylated Raf bound to other proteins (e.g., to MEK, brown-shaded in Figure 2.3A). Equation 2.5 is not a closed solution, as it includes the term $[pRaf]_{ss}$, and there are variables included in parameter groups α, β, γ. We confirmed that the parameter groups remain constant over the course of signaling (within 10%, Figure S2.1), justifying treating the latter variables as parameters.

Next, we considered how the analytical expression (Equation 2.5) behaves within a specific parameter regime observed in experiments. First, experiments in several mammalian cell systems have shown that feedback is strong, such that a significant fraction of the Raf pool is inhibited [10, 15]. This means that $\beta \cdot [dpERK]_{ss} \sim (\alpha+\gamma)\cdot u+\delta$. Second, as has been observed in multiple contexts [13, 21, 41, 44], ERK phosphorylation is ultrasensitive to the amount of pRaf (the ultrasensitive cascade is shaded green in Figure 2.3A). Denoting K as the relative change of $[dpERK]_{ss}$ with respect to $[pRaf]_{ss}$, ultrasensitivity entails that $K \gg 1$. In this range, small changes in pRaf level have very large effects on dpERK level (e.g., in model simulations, a 30% change in pRaf level results in a 900% change in dpERK level, Figure S2.1). We find analytically that in the parameter regime where $\beta\cdot[dpERK]_{ss} \sim (\alpha+\gamma)\cdot u+\delta$ and $K \gg 1$, the negative feedback holds the level of pRaf constant ($[pRaf]_{ss} \approx R_s$, details in SI). With these two features, strong negative feedback and ultrasensitivity, dpERK becomes a linear function of the input u:

$$[dpERK]_{ss} \approx \frac{\alpha}{\beta} \cdot \frac{Raf_{tot}}{R_s} \cdot u - \frac{\delta}{\beta} \tag{2.10}$$

The full derivation is given in the SI, and includes a toy model to illustrate the intuition for how ultrasensitivity combines with negative feedback to produce linearity. Equation 2.10 is plotted in Figure 2.3C (red line). We confirmed that the numerical solution of the full model matches the analytics in Equation 2.10 and becomes nonlinear when the negative feedback is weakened (grey line, Figure 2.3C). Although the analytical expression describes up until 50% of ERK activation, we verified numerically that the predicted linearity extends to 93% of ERK activation (Figure S2.2).

The linearity derived here applies across different dynamic ERK responses. The model we analyzed gives a sustained dpERK response. In some contexts, however, the ERK pathway shows a pulsatile response, which has been attributed to receptor desensitization [41]. Using a larger model that includes details of receptor desensitization [41], we numerically verified that the linearity holds for pulsatile responses, i.e. the peak level of dpERK increases linearly with the peak level of u (Figure S2.1).

Mathematical analysis reveals linear signal transmission in the Tgf-β pathway

Finally, we examined signal transduction within the Tgfβ pathway (Figure 2.4A). In the Tgfβ pathway, input from ligand-receptor interactions is transmitted by the Smad proteins. There are several classes of Smad proteins, including the receptor-regulated Smads (R-Smads) and the common Smad (co-Smad or Smad4) [32]. Ligand-activated receptors phosphorylate R-Smads. Phosphorylated R-Smads bind to the co-Smad and shuttle into the nucleus and regulate broad target genes. In the nucleus, the Smad complex dissociates and R-Smads are constitutively dephosphorylated and shuttled out to the cytoplasm, where the cycle of phosphorylation and complex formation begins again [40]. This dynamic translocation in and out of the nucleus forms a continual nucleocytoplasmic shuttling of Smads, a known integral feature of the Tgfβ pathway [22, 33, 39, 48].

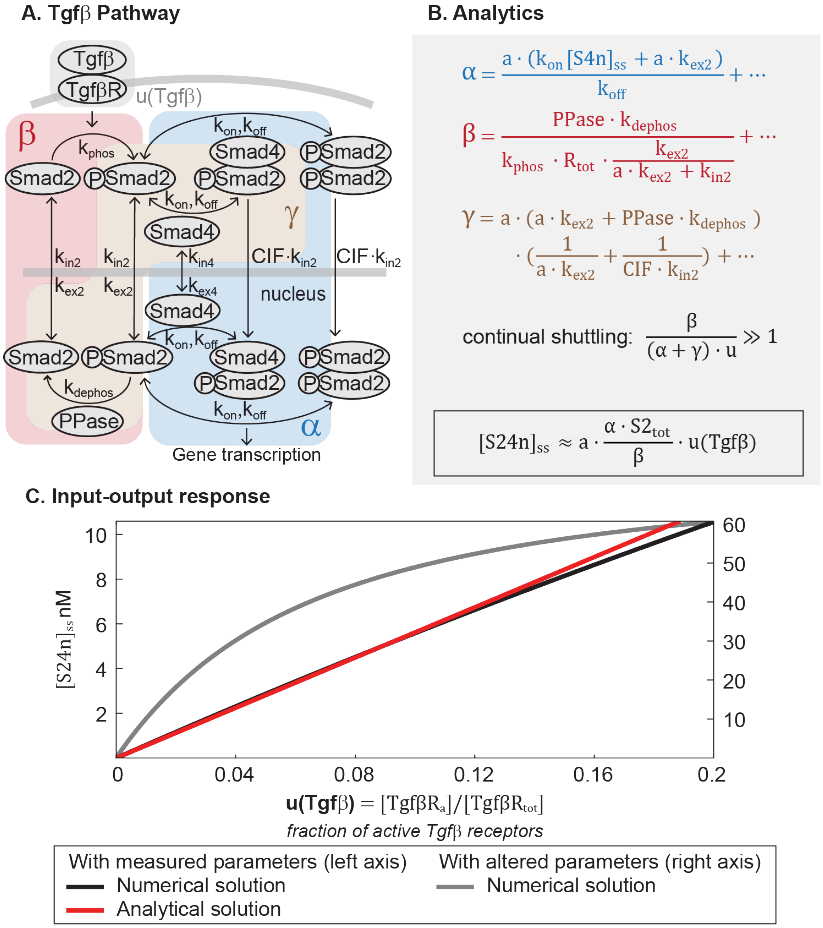


Figure 2.4: Mathematical analysis reveals linear signal transmission in the Tgf-β pathway. **(A)** Network diagram of the Tgfβ signaling pathway, modified from ref. [40]. **(B)** The parameter groups and input-output relationship that were analytically derived in this study. Parameter groups and input function are color-coded to the corresponding reactions in the network diagram. Parameters that do not appear in the parameter groups are negligible (as indicated by ellipses). **(C)** The input-output relationship of the Tgfβ pathway. The output is nuclear Smad complex, and the input function *u* is the fraction of Tgfβ -activated receptors.

The Tgfβ model [40] was published in 2008 by the Hill lab, and consists of 10 differential equations and 13 parameters. Even though the model was fitted to R-

Smad2 data, the general architecture of signal transmission is conserved across all five R-Smads [32, 40]. Using the variable elimination technique described before [11], we derived an analytical expression of the steady-state concentration of Smad complex in the nucleus:

$$[S24n]_{ss} = a \cdot \frac{\alpha \cdot u}{(\alpha + \gamma) \cdot u + \beta} \cdot S2_{tot} \tag{2.11}$$

$$\alpha = \frac{a \cdot (k_{on}[S4n]_{ss} + a \cdot k_{ex2})}{k_{off}} + \dots \tag{2.12}$$

$$\beta = \frac{PPase \cdot k_{dephos}}{k_{phos} \cdot R_{tot} \cdot \frac{k_{ex2}}{a \cdot k_{ex2} + k_{in2}}} + \dots \tag{2.13}$$

$$\gamma = a \cdot \left(a \cdot k_{ex2} + PPase \cdot k_{dephos}\right) \left(\frac{1}{a \cdot k_{ex2}} + \frac{1}{CIF \cdot k_{in2}}\right) + \dots \tag{2.14}$$

In Equation 2.11, the input function $u = u(Tgf\beta)$ is the active fraction of Tgfβ receptors. The parameter a is the nucleocytoplasmic volume ratio. The dimensionless parameter groups α, β, and γ in Equation 2.11 are defined in Equation 2.12-2.14, where the ellipses indicate additional small terms (expanded in SI). α, β, and γ describe how the Smad2 pool partitions during signaling (Eqs. S4.2, 8, 9): $\alpha \cdot u$ relates to the amount of nuclear Smad complex (α, blue-shaded in Figure 2.4A, captures the parameters related to complex formation and translocation to the nucleus), β relates to the amount of free, unphosphorylated Smad2 (β, red-shaded in Figure 2.4A, captures the parameters related to complex dissociation and translocation to the cytoplasm), and $\gamma \cdot u$ loosely relates to the remaining Smad2 pool (γ is brown-shaded in Figure 2.4A). Phosphorylated Smad2 quickly forms complex [26], so β essentially corresponds to total monomeric Smad2. Finally, Equation 2.11 is not a closed solution, since variable $[S4n]_{ss}$ appears in α. We numerically tested that it is constant within 2% for non-saturating inputs (Figure S2.3), justifying treating it as a parameter.

As in the Wnt and ERK pathways, the analytical expression for nuclear Smad complex (Equation 2.11) allows us to see that the behavior dramatically simplifies with parameters observed in experiment. We consider the case for non-saturating inputs ($u \sim .1$). Protein concentrations in the Tgfβ model were measured in human keratinocyte cells and the rate constants fitted to kinetic data measured in the cells [40]. With the measured parameters, we find that $\beta \sim 46$, $\alpha \cdot u \sim 1.5$, and $\gamma \cdot u \sim 0.7$. In this parameter regime, once Smad2 is imported to the nucleus, it is rapidly dephosphorylated and exported. Dynamic Smad2 translocation maintains monomeric Smad2 in excess to Smad complex ($\beta \gg (\alpha + \gamma) \cdot u$). and forms the continual nucleocytoplasmic shuttling that is characteristic of the Tgfβ pathway. Even under maximal Tgfβ stimulation, it has been estimated that phosphorylated Smad2 comprises only 36% of the Smad2 pool [16, 39]. With $\beta \gg (\alpha + \gamma) \cdot u$, the first term in the denominator of Equation 2.11 is small, and concentration of nuclear Smad complex becomes a linear function of input:

$$[S24n]_{ss} \approx a \cdot \frac{\alpha \cdot S2_{tot}}{\beta} \cdot u \tag{2.15}$$

Equation 2.15 is plotted in Figure 2.4C (red line), and we confirmed that numerical simulations recapitulate Equation 2.15 (blue line, Figure 2.4C). Although the analytical solution is valid only for small values of u, we numerically verified that the predicted linearity holds for the entire range of input u (from 0 to 1, Figure S2.2). We confirmed that the pathway becomes nonlinear when the R-Smad phosphatase is inhibited such that $\beta \sim (\alpha + \gamma) \cdot u$ (grey line, Figure 2.4C). While the model analyzed here gives a sustained Smad response, we verified numerically that the linearity holds for a larger model that includes receptor desensitization and gives a pulsatile Smad response (Figure S2.3) [45].

Linearity in the Wnt, MAPK/ERK, and Tgfβ pathways occurs across wide parameter range

Our study suggests that the canonical Wnt pathway, the ERK pathway, and the Tgfβ pathway have converged upon a shared strategy of linear signal transmission. We would then like to explore how wide a range of parameters this linearity occurs across. To explore this, we varied several parameter groups for each signaling pathway: α and γ for the Wnt pathway, α, β, and K for the ERK pathway, and α and β for the Tgfβ pathway. In the ERK and Tgfβ pathways, the remaining parameter groups are not able to be varied independently, since they share many parameters with other groups. After varying the parameter groups, we numerically calculated the input-output response and assessed its linearity using the L1-norm. We find that linearity occurs through a considerable range of parameters (Figure 2.5).

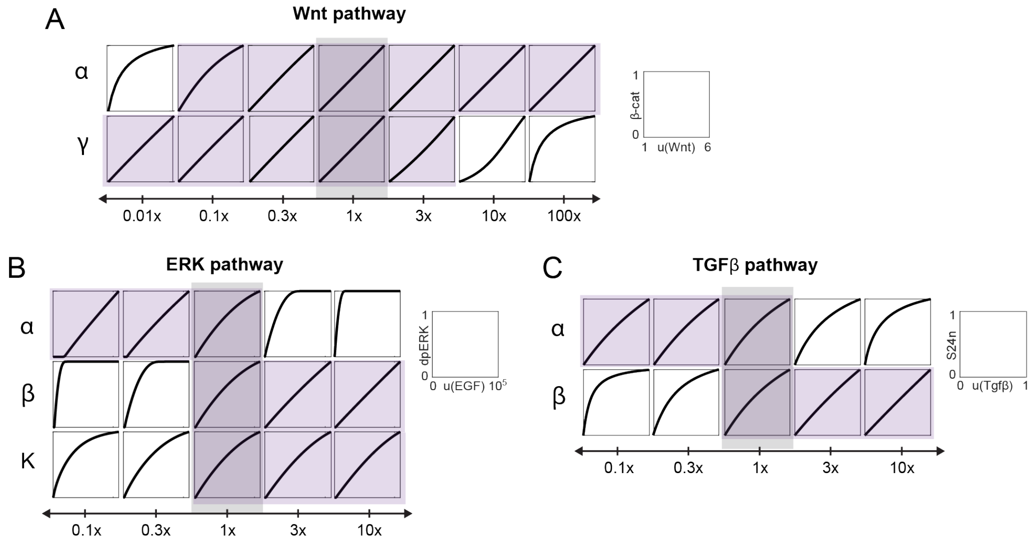


Figure 2.5: Linear signal transmission occurs over a range of parameters in the model. In this analysis, the parameter groups in each model were varied as indicated e.g., 3x is 3-fold increase, 0.3x is 3-fold decrease. 1x corresponds to the measured parameters. Plotted in each box is the input-output relationship, numerically simulated over the full dynamic range of the models, i.e., 1-6 for $u(Wnt)$, 0-10^5 for $u(EGF)$, and 0-1 for $u(Tgf\beta)$. For simplicity, all outputs are normalized from 0 to 1. Grey shade: the unperturbed state. Purple shade: linear input-output response, as defined by L-1 norm < 0.1.

Next, we validated the relationship between linearity and the requirements for linearity derived in the previous sections. We restated each requirement in terms of a scaling term S, such that linearity occurs when $S >> 1$ for the Wnt and Tgfβ pathways, and $S > 1$ for the ERK pathway. For the Wnt pathway, linearity occurs where $\alpha/u \gg 1_\gamma$, giving us the scaling term

$$S = \frac{\alpha}{(1+\gamma)\cdot u} \tag{2.16}$$

For the ERK pathway, the feedback must be strong enough to maintain the response in the linear regime. We can express this in terms of the scaling term:

$$S = \frac{\beta[dpERK]_{ss} + \delta}{\alpha \cdot u} \cdot \frac{R_s}{Raf_{tot}} \tag{2.17}$$

We also tested the requirement of ultrasensitivity ($K >> 1$) for linearity, by varying parameters in the kinase cascade to increase or decrease the sensitivity of the ERK cascade. For the Tgfβ pathway, linearity requires that $\beta \gg (\alpha+\gamma)\cdot u$. We rewrite this in terms of the scaling term:

$$S = \frac{\beta}{(\alpha+\gamma)\cdot u} \tag{2.18}$$

We then varied each scaling term by altering the value of the underlying parameter groups. As in the previous section, we assessed the linearity of the input-output response using the L1-norm. We find that linearity occurs over a large range of S (Figure 2.6). Interestingly, the ERK and Tgfβ pathway operate close to the edge of this linear regime. This could indicate that these pathways are capable of switching to nonlinear response in different cellular contexts.

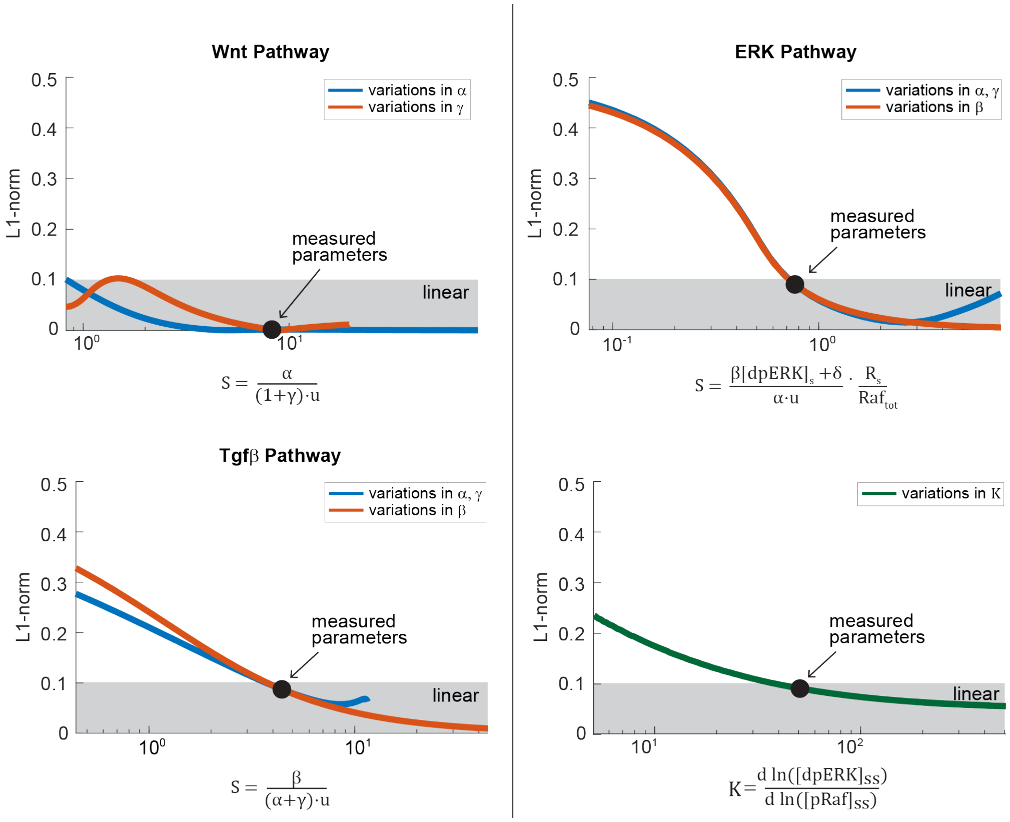


Figure 2.6: The requirements for linear signal transmission in the Wnt, Tgfβ, and ERK pathways. In each plot, we varied S, defined in the equation shown on the x-axis, and simulated the input-output curve over the dynamic range of the model. The parameters in the equations are as defined in the main text. For the ERK and Tgfβ pathways, α and γ are linked in such a way that they could not easily be varied independently. Linearity was assessed using the L-1 norm, which ranges from 0 to 0.5, with L-1 norm < 0.1 indicating linearity. L1-norm analysis was performed over the full dynamic range of the system, i.e., $u(Wnt)$ = 1- 6, $u(ERK)$ = 0 to 110,000 molecules of Ras-GTP, which gave 90% activation of $[dpERK]_{S}s$ in unperturbed cells, and $u(Tgf\beta)$ = 0 to 1.

2.3 Discussion

In this chapter, we examined models of three signaling pathways to determine their input-output response. Surprisingly, we found that despite their dissimilar architectures, all three pathways converge upon a shared signaling strategy: linear signal transmission. We also showed that linearity holds for steady-state response as well as transient response in the ERK and Tgfβ pathways, and that the linearity holds for a wide range of parameters.

How general is linear signal transmission? We have identified linearity in three pathways – we would like to know whether linearity arises in other signaling pathways as well. Specifically, we analyze the NF-κB since there is an established math model that captures the biochemical reactions of the pathway [3, 19, 29]. However, we cannot perform similar steady-state analysis as for the other pathways, due to the dynamic nature of the pathway: notably, the NF-κB pathway exhibits strong oscillatory response to stimulus due to a slow transcriptional feedback [19]. Instead, we employ numerical simulations of a well-established NF-κB model [3] over the range of nuclear NF-κB translocation observed in human epithelial cells [29]. From this, we indeed find that the peak of the nuclear NF-κB pulse correlates linearly with ligand concentration (Figure 2.7).

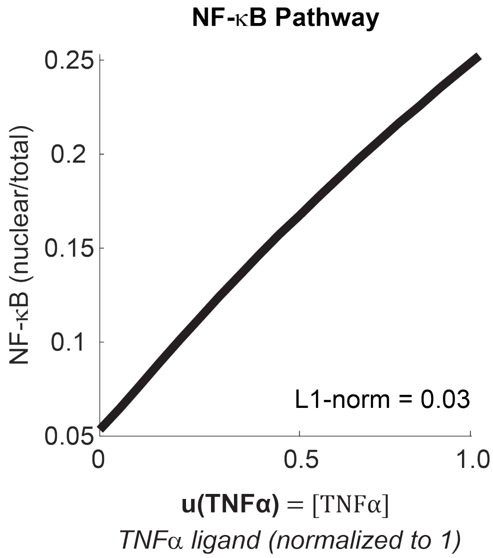


Figure 2.7: Numerical simulation of the input-output relationship of the NF-κB pathway. We used the model first built by Hoffman et al. in 2002 [19] and later revised by Ashall et al. in 2009 [3]. The parameters in the model have been measured or fitted to single-cell dynamics in multiple cell types. We simulated the model here over a physiologically-observed dynamic range, i.e., Lee et al., 2014 [29] observed in HeLa cells that at saturating ligand dose (10 ng/mL TNFα, set to 1 in the model), 25% of NF-κB pool is nuclear. Linearity is assessed using the L1-norm, where L1-norm <1 indicates linear relationship (see Method).

While mathematical analysis provides strong insight into mechanisms of cell signaling, further experimental evidence can improve our understanding of linear signal transmission. Models are useful in providing qualitative predictions, but it is also important to perform quantitative measurements of cell signaling behavior. Therefore, in the next chapter, we validate linearity findings experimentally.

2.4 Supplementary Figures

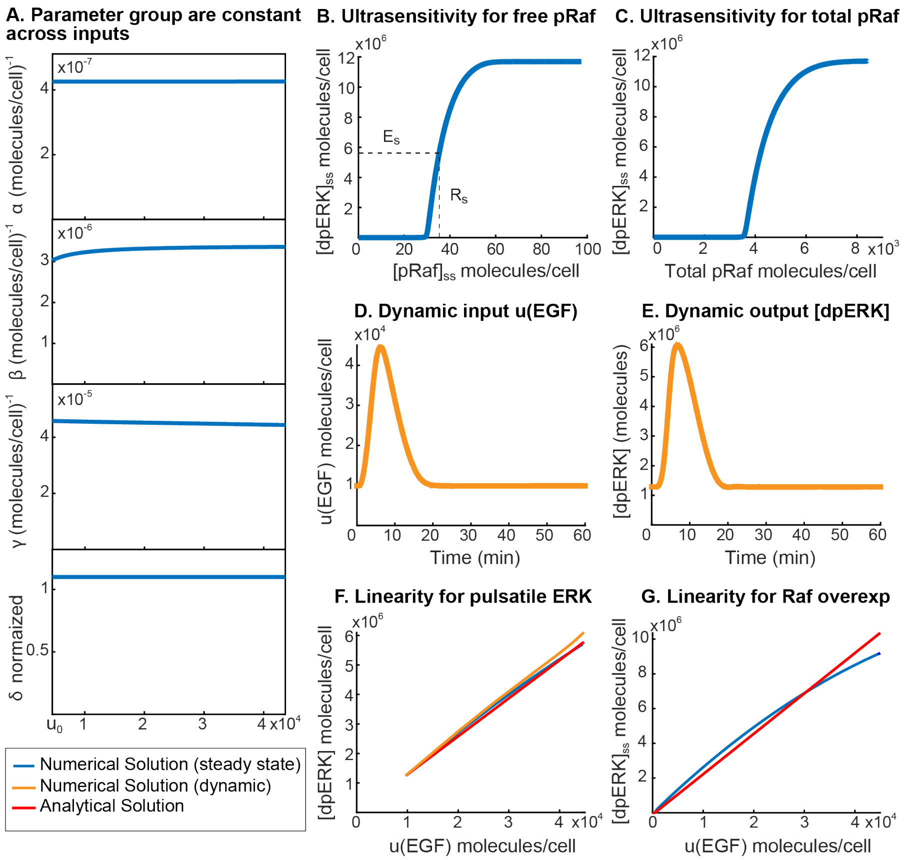


Figure S2.1: Model simulations for the ERK pathway. **(A)** Parameter groups in the ERK model are constant to within 10%, over the physiologically relevant range of u considered here, justifying the inclusion of variables into the parameter groups. **(B-C)** The dpERK output is an ultrasensitive function of both free and total phosphorylated Raf. The values E_s and R_s are illustrated in **(B)** and are defined in SI Section 3. **(D-F)** Numerical simulation of pulsatile response in the ERK pathway. **(D)** A pulse of input, RasGTP, is generated by EGF addition in an ERK model that includes details of receptor desensitization [41]. Basal activity of Ras is included to ensure constitutive negative feedback [15]. **(E)** dpERK output also exhibits a pulsatile response, peaking within 10 minutes. **(F)** We plot the peak dpERK output against peak input for a range of physiologically relevant $u(EGF)$ doses, and find that it matches our steady-state predictions for linear input-output response. **(G)** Five-fold Raf overexpression does not break the linear input-output response.

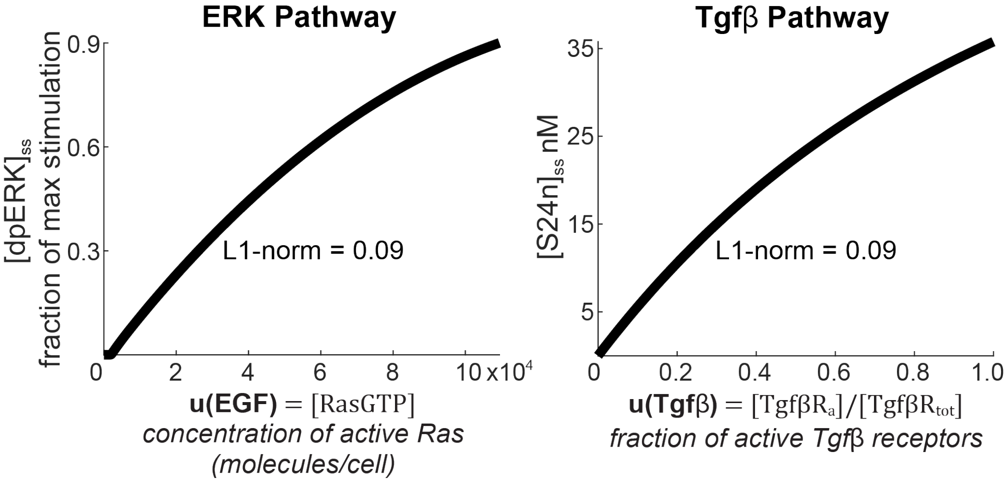


Figure S2.2: The predicted linearity extends throughout the dynamic range of the ERK and Tgfβ pathways. **(A-B)** Numerical simulation of the ERK and Tgfβ models. **(A)** The ERK model shows linear input-output relationship up to 93% of dpERK activation. **(B)** The Tgfβ pathway shows linear input-output relationship throughout the entire input range (from 0 to 1). Linearity was analyzed using the L1-norm or least absolute deviation (see Method). The blue range indicates where L1-norm was computed.

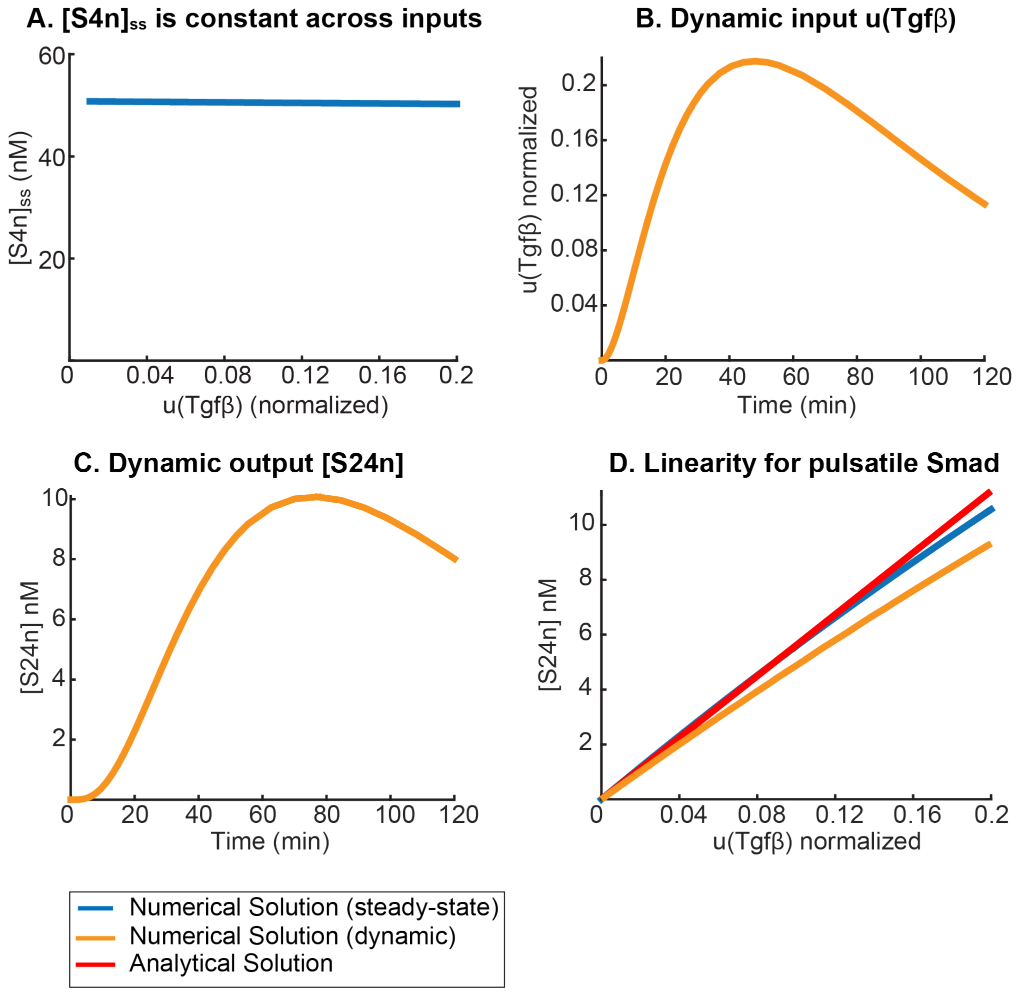


Figure S2.3: Model simulations for the Tgfβ pathway. **(A)** Nuclear Smad4 concentration is constant to within 2%, over a physiologically relevant range of $u(Tgf\beta)$ considered here, justifying its inclusion into parameter group $alpha$. **(B-D)** Numerical simulation of pulsatile response in the Tgfβ pathway. **(B)** A pulse of input, active Tgfβ receptor, is generated by Tgfβ addition in a model that includes details of receptor desensitization [45]. **(C)** S24n output also exhibits a pulsatile response. **(D)** We plot the peak S24n output against peak input and find that it matches our steady-state predictions for linear input-output response.

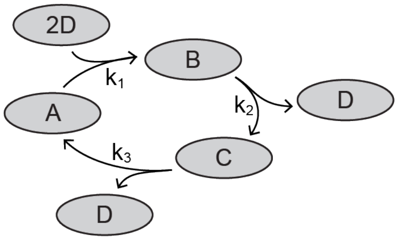
2.5 Supporting Information

Variable Elimination

We use a variable elimination technique from Feliu et al. [11] to derive analytic expressions for the steady-states of the Tgfβ and ERK pathways. This technique was developed to handle the complexity of large chemical reaction networks. By eliminating variables from the steady-state solution, we can express the steady-state of the system in terms of a smaller subset of variables. This is a useful tool for analyzing the Tgfβ and ERK models, as the steady-state solution consists of a large set of variables, each with a polynomial equation describing its steady-state.

The technique works as follows: if we can identify a cut set within the reaction network, we can reduce the system to a set of first-order homogeneous equations with respect to that cut. This set of equations can then be solved using linear algebra.

A cut is a set of species such that for every reaction involving those species, there is exactly one reactant and one product that falls within that cut. For example, let us consider a network of 4 interacting species, A, B, C, and D.



$$
\begin{aligned}
A+2D &\xrightarrow{k_1} B\\
B &\xrightarrow{k_2} C+D\\
C &\xrightarrow{k_3} A+D
\end{aligned}
$$

In this network, there is a cut $\{A, B, C\}$ that contains exactly one product and one reactant for each reaction. We have highlighted this cut in the reaction set:

$$
\begin{aligned}
{\color{red}A}+2D &\xrightarrow{k_1} {\color{red}B}\\
{\color{red}B} &\xrightarrow{k_2} {\color{red}C}+D\\
{\color{red}C} &\xrightarrow{k_3} {\color{red}A}+D
\end{aligned}
$$

The species D cannot belong in the cut, since it appears twice as a reactant in the first reaction. The behavior of this network is described by four differential equations,

$$[A] = -k_1[A] \cdot [D]^2 + k_3[C] = 0 \tag{2.19}$$

$$[B] = k_1[A] \cdot [D]^2 - k_2[B] = 0 \tag{2.20}$$

$$[C] = k_2[B] - k_3[C] = 0 \tag{2.21}$$

$$[D] = -2k_1[A] \cdot [D]^2 + k_2[B] + k_3[C] = 0 \tag{2.22}$$

which are set to zero at steady-state, and two additional conservation equations:

$$T_1 = [A] + [B] + [C] \tag{2.23}$$

$$T_2 = 2[B] + [C] + [D] \tag{2.24}$$

The variable elimination technique allows us to reduce the steady-state system of equations by four (three equations for the cut set, and one conservation equation). We do this by expressing each member of the cut set as a dependent variable of D, shown below. We utilize the fact that the differential equations for A, B, and C are first-order and homogeneous with respect to our cut, and rewrite them in matrix form. We use the subscript “ss” to denote steady-state:

$$\begin{pmatrix} -k_1[D]_{ss}^2 & 0 & k_3 \\ k_1[D]_{ss}^2 & -k_2 & 0 \\ 0 & k_2 & -k_3 \end{pmatrix} \begin{pmatrix} [A]_{ss} \\ [B]_{ss} \\ [C]_{ss} \end{pmatrix} = 0 \tag{2.25}$$

Feliu and Wiuf [11] provide a proof of why a cut set guarantees that we can rewrite the corresponding equations in matrix form. It can be understood intuitively from the fact that a cut contains exactly one reactant of each reaction, and therefore each rate is first-order with respect to the cut. Homogeneity also follows from this, since there are no rate terms that do not include members of the cut. For a complex model, there is no guarantee that we can derive closed-form analytical solutions for steady-state. The matrix formulation and variable elimination technique immediately provides us with a set of solvable variables. The solution to the matrix equation above is:

$$[A]_{ss} = c \cdot k_2 k_3 \tag{2.26}$$

$$[B]_{ss} = c \cdot k_1 k_3 D_{ss}^2 \tag{2.27}$$

$$[C]_{ss} = c \cdot k_1 k_2 D_{ss}^2 \tag{2.28}$$

c is a scaling factor not constrained by the matrix equation. With the use of the conservation equation S1.5, we can calculate c and express the steady state of all three species solely in terms of the parameters of the network, and $[D]_{ss}$. For instance, the solution for $[C]_{ss}$ is below.

$$[C]_{ss} = \frac{k_1 k_2 [D]_{ss}^2}{k_2 k_3 + k_1 [D]_{ss}^2 \cdot (k_2 + k_3)} T_1 \tag{2.29}$$

The solutions for $[A]_{ss}$, $[B]_{ss}$, and $[C]_{ss}$ derived from the variable elimination technique still depend on $[D]_{ss}$. If we plug in the solutions for the cut species, we can obtain polynomial equations for the remaining species (in this case $[D]_{ss}$), but closed form expressions are not necessarily obtainable. In all the cases analyzed in this paper, variables that appear in the analytical solutions for the cut set happen to be approximately constant across a wide range of input values, as they are present in excess relative to other species. Finally, each parameter group is physically meaningful. For instance, $k_2 k_3$, $k_1 k_3 [D]_{ss}^2$, and $k_1 k_2 [D]_{ss}^2$ represent the un-normalized fraction of T_1 that exists as A, B, and C, respectively. The normalization factor for these fractions is c/T_1, or in this case, simply the sum of all parameter groups. This provides an intuitive way of analyzing how parameter groups affect the overall distribution of T_1. For instance, increasing the value of k_1 will increase the amount of T_1 that exists as B and C, while necessarily decreasing the amount of A (assuming $[D]_{ss}$ does not change significantly)

Wnt Model

We analyzed a mathematical model of the canonical Wnt pathway built by Lee et al. [28]. The model is illustrated in Figure 2.2A, and consists of 7 ODEs and 22 parameters.

Solving the Wnt model at steady-state

We previously derived an expression for β-catenin in steady-state [17]:

$$[\beta cat]_{ss} = K_{17} \frac{1 - \gamma + \alpha / u(Wnt)}{2} \left(\sqrt{1 + \frac{4\gamma}{(1 - \gamma + \alpha / u(Wnt))^2}} - 1 \right) \tag{2.30}$$

where the parameters are dimensionless groups of the binding rate constants and protein concentrations:

$$\alpha = \frac{k_4 \cdot k_6 \cdot k_9 \cdot v_1 4 \cdot GSK3_{tot} \cdot APC_{tot}}{k_5 \cdot k_{-6} cdot K_7 \cdot K_8 \cdot k_{13} \cdot k_{15}} \tag{2.31}$$

$$\gamma = \frac{v_{12}}{k_{13} \cdot K_{17}} \tag{2.32}$$

$$u(Wnt) = 1 + \frac{k_3 \cdot Dvl_{tot}}{k_{-6}} \cdot \frac{k_1 \cdot Wnt}{k_2 + k_1 \cdot Wnt} \tag{2.33}$$

The input function $u = u(Wnt)$ corresponds to the rate at which Wnt stimulation inhibits the destruction complex, normalized by k_{-6}. The value of Wnt ranges from 0 to 1 in the model. Please refer to Goentoro et al. [17] for the physical intuition of each parameter group.

Derivation of Linear Behavior

We calculate the value of the parameter groups, as well as the value of the input function at saturating Wnt stimulation:

$$\alpha = 66 \tag{2.34}$$

$$\gamma = 1.4 \tag{2.35}$$

$$u(Wnt = 1) = 6.0 \tag{2.36}$$

Within the parameter regime measured in cells, the analytical expression for β-catenin dramatically simplifies. We can perform the following first-order Taylor expansion:

$$\sqrt{1+\epsilon} \approx 1 + \frac{1}{2}\epsilon, \qquad \epsilon \ll 1 \tag{2.37}$$

$$\epsilon = \frac{4\gamma}{(1 - \gamma + \alpha/u)^2} \tag{2.38}$$

This holds true for $\alpha/u \gg \gamma$. Furthermore, we can make the approximation $1 - \gamma + \alpha/u \approx \alpha/u$ as long as $\alpha/u \gg 1$ also holds. We can encompass these two inequalities within $\alpha/u \gg 1 + \gamma$. The equation simplifies to:

$$[\beta cat]_{ss} \approx K_{17}\frac{\gamma}{\alpha}u \tag{2.39}$$

ERK Model

We analyzed a mathematical model built by Huang et al. [21] and revised by Sturm et al. [44]. The model is illustrated in Figure 2.3A and contains 26 ODEs and 46 parameters. We changed two parameters from the original model. k_{25} characterizes the negative feedback from dpERK to unphosphorylated Raf, and k_{27} characterizes the negative feedback from dpERK to phosphorylated Raf. In Sturm et al. [44], the values of these parameters were estimated, rather than measured. Experimental measurements indicate that dpERK mostly interacts with Raf, and that this feedback causes strong repression of Raf [10]. We therefore increased the value of k_{25} and set k_{27} to zero.

Solving the ERK model at steady-state

In the ERK pathway, doubly-phosphorylated ERK is produced by the Raf/MEK/ERK cascade of phosphorylation,

$$[dpERK]_{ss} = g([pRaf]_{ss}) \tag{2.40}$$

There is a negative feedback within the pathway, such that,

$$[pRaf]_{ss} = f(u, [dpERK]_{ss}) \tag{2.41}$$

where u is the input function, the concentration of RasGTP (a function of ligand dose). We first focus on deriving the negative feedback function in Equation S3.2. Using the variable elimination techniques in section “Variable Elimination”, we identify the following cut set:

$\{Raf, Raf : RasGTP, pRaf, pRaf : P1, MEK : pRaf, pMEK : pRaf, Raf : ppERK, Rafi, Rafi : P4\}$

This allows us to express the steady-state concentration of pRaf as a function of parameters, and the remaining species in the ERK pathway. Specifically, members of this cut interact directly with, and have dependencies on, the following set:

$\{P1, MEK, pMEK, dpERK, P4\}$

With this, we derive the expression for $[pRaf]_{ss}$,

$$[pRaf]_{ss} = \frac{\alpha \cdot u}{\beta \cdot [dpERK]_{ss} + (\alpha + \gamma) \cdot u + \delta} \cdot Raf_{tot} \tag{2.42}$$

where the parameter groups are:

$$\alpha = \frac{k_3 \cdot (k_8 + k_b 7)}{k_7 \cdot [P1]_{ss} \cdot k_8} + \ldots \tag{2.43}$$

$$\beta = \frac{k_{25} \cdot (k_{30} + k_{b29} + k_{29} \cdot [P4]_{ss})}{k_{29} \cdot [P4]_{ss} \cdot k_{30}} + \ldots \tag{2.44}$$

$$\gamma = \frac{k_3 \cdot (k_8 + k_{b7}) \cdot (k_9 \cdot [MEK]_{ss})}{k_7 \cdot [P1]_{ss} \cdot k_8 \cdot k_{10})} + \ldots \tag{2.45}$$

$$\delta = \frac{k_{26} + k_{b25}}{k_{26}} + \ldots \tag{2.46}$$

The ellipses indicate additional small terms (i.e., <10% of the previous terms, numerically calculated using the model parameters and $u = 4.5e4$). All the calculations for this paper use these truncated parameter groups. The complete parameter groups are written below:

$$\begin{aligned}\alpha =&(k_3 \cdot (k_8 + k_{b7}) \cdot (k_{10} + k_{b9}) \cdot (k_{12} + k_{b11}) \cdot (k_{26} + k_{b25}))/([P1]_{ss} \cdot k_7 \cdot k_8 \\ &\cdot k_{10} \cdot k_{12} \cdot k_{26})\end{aligned}$$

$$\begin{aligned}\beta =&(k_{25} \cdot (k_4 + k_{b3}) \cdot (k_{10} + k_{b9}) \cdot (k_{12} + k_{b11}) \cdot (k_{26} \cdot k_{30} + k_{26} \cdot k_{b29} + [P4]_{ss} \\ &\cdot k_{26} \cdot k_{29} + [P4]_{ss} \cdot k_{29} \cdot k_{30}))/([P4]_{ss} \cdot k_4 \cdot k_{10} \cdot k_{12} \cdot k_{26} \cdot k_{29} \cdot k_{30})\end{aligned}$$

$$\begin{aligned}\gamma =&(k_3 \cdot (k_{26} + k_{b25}) \cdot (k_4 \cdot k_8 \cdot k_{10} \cdot k_{12} + k_4 \cdot k_8 \cdot k_{10} \cdot k_{b11} + k_4 \cdot k_8 \cdot k_{12} \cdot k_{b9} \\ &+ k_4 \cdot k_{10} \cdot k_{12} \cdot k_{b7} + k_4 \cdot k_8 \cdot k_{b9} \cdot k_{b11} + k_4 \cdot k_{10} \cdot k_{b7} \cdot k_{b11} + k_4 \cdot k_{12} \cdot k_{b7} \\ &\cdot k_{b9} + k_4 \cdot k_{b7} \cdot k_{b9} \cdot k_{b11} + [MEK]_{ss} \cdot k_4 \cdot k_8 \cdot k_9 \cdot \cot k_{12} + [MEK]_{ss} \cdot k_4 \\ &\cdot k_8 \cdot k_9 \cdot k_{b11} + [MEK]_{ss} \cdot k_4 \cdot k_9 \cdot k_{12} \cdot k_{b7} + [MEK]_{ss} \cdot k_4 \cdot k_9 \cdot k_{b7} \cdot k_{b11} \\ &+ [P1]_{ss} \cdot k_4 \cdot k_7 \cdot k_{10} \cdot k_{12} + [P1]_{ss} \cdot k_7 \cdot k_8 \cdot k_{10} \cdot k_{12} + [P1]_{ss} \cdot k_4 \cdot k_7 \cdot k_{10} \\ &\cdot k_{b11} + [P1]_{ss} \cdot k_4 \cdot k_7 \cdot k_{12} \cdot k_{b9} + [P1]_{ss} \cdot k_7 \cdot k_8 \cdot k_{10} \cdot k_{b11} + [P1]_{ss} \cdot k_7 \\ &\cdot k_8 \cdot k_{12} \cdot k_{b9} + [P1]_{ss} \cdot k_4 \cdot k_7 \cdot k_{b9} \cdot k_{b11} + [P1]_{ss} \cdot k_7 \cdot k_8 \cdot k_{b9} \cdot k_{b11} + k_4 \\ &\cdot k_8 \cdot k_{10} \cdot k_{11} \cdot [pMEK]_{ss} + k_4 \cdot k_8 \cdot k_{11} \cdot k_{b9} \cdot [pMEK]_{ss} + k_4 \cdot k_{10} \cdot k_{11} \\ &\cdot k_{b7} \cdot [pMEK]_{ss} + k_4 \cdot k_{11} \cdot k_{b7} \cdot k_{b9} \cdot [pMEK]_{ss}))/([P1]_{ss} \cdot k_4 \cdot k_7 \cdot k_8 \\ &\cdot k_{10} \cdot k_{12} \cdot k_{26})\end{aligned}$$

$$\delta =((k_4 + k_{b3}) \cdot (k_{10} + k_{b9}) \cdot (k_{12} + k_{b11}) \cdot (k_{26} + k_{b25}))/(k_4 \cdot k_{10} \cdot k_{12} \cdot k_{26})$$

Physical significance of parameter groups

Next, we would like to develop an intuition for the physical significance of these parameter groups. As discussed above, $\alpha \cdot u$ relates to the amount of free, phosphorylated Raf since $\alpha \cdot u/((\alpha + \gamma) \cdot u + \beta[dpERK]_{ss} + \delta)$ is the fraction of Raf present as pRaf. Thus, as $\alpha \cdot u$ increases relative to $\gamma \cdot u + \beta[dpERK]_{ss} + \delta$, the amount of pRaf also increases.

We can define three subpopulations of Raf: Raf inhibited by dpERK, $[R_i]$; Raf activated by RasGTP (input), $[R_a]$; and unphosphorylated Raf $[R_n]$. Specifically:

$$[R_i] = [Raf : dpERK] + [Rafi] + [Raf : P4] \tag{2.47}$$

$$[R_a] = [pRaf] + [pRaf : P1] + [MEK : pRaf] + [pMEK : pRaf] \tag{2.48}$$

$$+ [Raf : RasGTP] \tag{2.49}$$

$$[R_n] = [Raf] \tag{2.50}$$

We can calculate the steady-state of each subpopulation as:

$$[R_i]_{ss} = \frac{\beta \cdot [dpERK]_{ss}}{(\alpha + \gamma) \cdot u + \beta[dpERK]_{ss} + \delta} Raf_{tot} \tag{2.51}$$

$$[R_a]_{ss} = \frac{\gamma \cdot u}{(\alpha + \gamma) \cdot u + \beta[dpERK]_{ss} + \delta} Raf_{tot} + [pRaf]_{ss} \tag{2.52}$$

$$[R_n]_{ss} = \frac{\delta}{(\alpha + \gamma) \cdot u + \beta[dpERK]_{ss} + \delta} \cdot Raf_{tot} \tag{2.53}$$

Thus, in the same sense that $\alpha \cdot u$ relates to the amount of free phosphorylated Raf, $\beta \cdot [dpERK]_{ss}$ relates to the amount of inhibited Raf, $\gamma \cdot u$ relates to the amount of phosphorylated Raf bound to other proteins (not free), and δ relates to the amount of unphosphorylated Raf.

Derivation of linear behavior

Now that we have derived the negative feedback function from Equation S3.2, we examine Equation S3.1. The relationship $[dpERK]_{ss} = g([pRaf]_{ss})$ is analytically intractable, because of the complexity of the phosphorylation cascade. But we know from simulations and experimental observations that it is an ultrasensitive function. From simulations, we find that a 1.3 fold change in pRaf leads to a 9-fold change in dpERK (from 10% to 90% of max, Figure S1B-C). We therefore approximate $[pRaf]_{ss}$ by a value R_s within this range, as indicated by the dashed line in Figure S2.1B. Substituting this into the equation above and rearranging, we find that $[dpERK]_{ss}$ becomes a linear function of input:

$$[dpERK]_{ss} \approx \frac{\alpha}{\beta} \cdot \left(\frac{Raf_{tot}}{R_s} - 1 - \frac{\gamma}{\alpha} \right) \cdot u - \frac{\delta}{\beta} \tag{2.54}$$

Lastly, we write the value of two terms in Equation S3.14 below, numerically calculated using the parameter values of the model:

$$\frac{\alpha}{\beta} \cdot \frac{Raf_{tot}}{R_s} = 140$$

$$\frac{\alpha}{\beta} \cdot (1 + \frac{\gamma}{\alpha}) = 13$$

We can neglect the second term, yielding:

$$[dpERK]_{ss} \approx \frac{\alpha}{\beta} \cdot \frac{Raf_{tot}}{R_s} \cdot u - \frac{\delta}{\beta} \tag{2.55}$$

Derivation for treating pRaf as a constant

Next, we analyze exactly how the level of pRaf changes with the input u. From earlier, we have that

$$[dpERK]_{ss} = g([pRaf]_{ss})$$

$$[pRaf_{tot}]_{ss} = f(u, [dpERK]_{ss})$$

We can now derive a general expression for the relative change of $[pRaf]_{ss}$ with respect to a relative change in u. We use the notation $d\hat{x} = dlnx = dx/x$

$$\frac{d\hat{f}}{d\hat{u}} = \frac{\partial f}{\partial u} \cdot \frac{u}{f} \cdot \left(1 - \frac{\partial f}{\partial [dpERK]_{ss}} \cdot \frac{dg}{d[pRaf]_{ss}}\right)^{-1} \tag{2.56}$$

Next, we define the response coefficient K between $[dpERK]_{ss}$ and $[pRaf]_{ss}$:

$$K \triangleq \frac{dg}{d[pRaf]_{ss}} \cdot \frac{[pRaf]_{ss}}{[dpERK]_{ss}} \tag{2.57}$$

From Equation S3.3, we get the partial derivatives:

$$\frac{\partial f}{\partial u} = \frac{f}{u} \cdot \frac{\beta [dpERK]_{ss} + \delta}{\beta [dpERK]_{ss} + (\alpha + \gamma) u + \delta} \tag{2.58}$$

$$\frac{\partial f}{\partial [dpERK]_{ss}} = -f \cdot \frac{\beta}{\beta [dpERK]_{ss} + (\alpha + \gamma) \cdot u + \delta} \tag{2.59}$$

Using these two equations, we find that:

$$\frac{d\hat{f}}{d\hat{u}} = \frac{\beta [dpERK]_{ss} + \delta}{(1 + k) \cdot \beta [dpERK]_{ss} + \alpha u + \delta} \tag{2.60}$$

When $K \gg 1$ and $\beta \cdot [dpERK]_{ss} \approx (\alpha + \gamma) \cdot u + \delta$, we see that

$$\frac{d\hat{f}}{d\hat{u}} \approx K^{-1} \tag{2.61}$$

Therefore, $[pRaf]_{ss}$ is held constant in the region where the kinase cascade is ultrasensitive and feedback is strong. In this region, it is easy to show that $[dpERK]_{ss}$ becomes a linear function of input.

$$\frac{dg \hat{-} g_0}{d\hat{u}} \approx 1; \qquad g_0 = -\frac{\delta}{\beta} \tag{2.62}$$

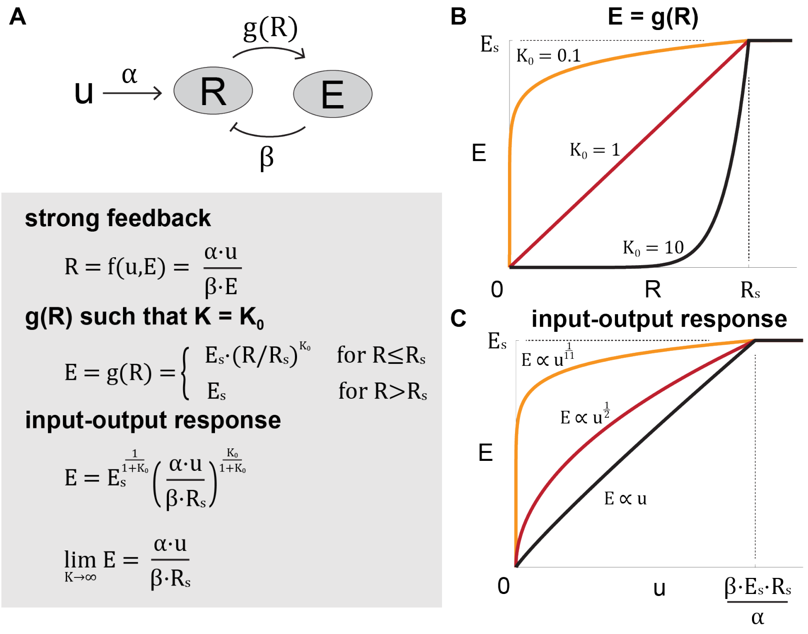
It is not guaranteed that the system is stable as K increases, but we see from simulations that our parameter regime provides a stable output.

Toy Model of the ERK pathway

Here we utilize a toy model to illustrate how ultrasensitivity and strong negative feedback combine to generate input-output linearity. In this model, induction of the output species E is a two-step process: An input u increases the amount of species R, which in turn influences E as $E = g(R)$. There is negative feedback from E to R, which in the limit of strong negative feedback is inversely proportional to E.

We specify the function $g(R)$ such that $K = K_0$, where K is the relative change of E with respect to R. As K_0 increases, therefore, the function $g(R)$ becomes more ultrasensitive. Solving for E, we see that in the limit of $K = K_0 \gg 1$, E becomes a linear function of u, and R is held constant at Rs.

While we do not have an explicit function for $g(R)$ for the full ERK model, we include derivations in section "Derivation for treating pRaf as a constant" that show that these results hold for any function $g(R)$ in the region where $K \gg 1$. We also show that these results hold outside the limit of strong negative feedback, as long as the feedback-inhibited pool of R is comparable to the remaining pool.



Tgfβ Model

We analyzed a mathematical model built by Schmierer et al. [40]. The model is illustrated in Figure 2.4A and consists of 10 ODEs and 14 parameters.

Solving the Tgfβ model at steady-state

We use the variable elimination technique described in section "Variable Elimination" to derive an analytical expression for the steady-state concentration of nuclear Smad complex. First, based on the measured parameter values, and as confirmed by simulations, the extent of Smad2-Smad2 binding is limited. We therefore neglect this reaction in subsequent analysis. We identify the following cut of the Tgfβ model:

$$\{S2c, pS2c, S24c, S2n, pS2n, S24n\}$$

which is subject to the conservation equation:

$$([S2c] + [pS2c] + [S24c]) + \frac{1}{a} \cdot ([S2n] + [pS2n] + [S24n]) = S2_{tot} \tag{2.63}$$

Thus, we can eliminate these variables from the steady-state polynomial solution, with dependence only on variables outside this cut:

$$\{S4c, S4n\}$$

Using this relationship, we derive an expression for the nuclear Smad complex (S24n) at steady-state,

$$[S24n]_{ss} = \frac{a \cdot \alpha \cdot u}{(\alpha + \gamma) \cdot u + \beta} S2_{tot} \tag{2.64}$$

where the parameter groups are:

$$\alpha = \frac{a \cdot (k_{on}[S4n]_{ss} + a \cdot k_{ex2})}{k_{off}} + \dots \tag{2.65}$$

$$\beta = PPase \cdot \frac{k_{dephos}}{k_{phos} \cdot R_{tot} \cdot k_{ex2}/(a \cdot k_{ex2} + k_{in2})} + \dots \tag{2.66}$$

$$\gamma = a \cdot (a \cdot k_{ex2} + PPase \cdot k_{dephos}) \left(\frac{1}{a \cdot k_{ex2}} + \frac{1}{CIF \cdot k_{in2}} \right) + \dots \tag{2.67}$$

Here the input function $u = u(Tgf\beta)$ is the fraction of receptors activated by Tgfβ ligands. The ellipses indicate additional small terms (i.e., <10% of the previous terms, as calculated using the model parameters, with the variables $[S4c]_{ss}$ and

$[S4n]_{ss}$ calculated for $u = 0$). All calculations for the paper use these truncated parameter groups. The complete parameter groups are written below:

$$\begin{aligned}\alpha =&(a \cdot ([S4n]_{ss} \cdot k_{off} + CIF \cdot [S4n]_{ss} \cdot k_{in2} + CIF \cdot PPase \cdot [S4c]_{ss} \cdot k_{dephos} \\ &+ CIF \cdot [S4c]_{ss} \cdot [S4n]_{ss} \cdot k_{on} + CIF \cdot [S4c]_{ss} \cdot a \cdot k_{ex2}))/(CIF \cdot [S4c]_{ss} \\ &\cdot k_{off})\end{aligned}$$

$$\begin{aligned}\beta =&(PPase \cdot k_{dephos} \cdot (k_{in2} + a \cdot k_{ex2}) \cdot (k_{off} + CIF \cdot k_{in2} + CIF \cdot [S4c]_{ss} \\ &\cdot k_{on}))/(CIF \cdot R_{tot} \cdot [S4c]_{ss} \cdot k_{ex2} \cdot k_{on} \cdot k_{phos})\end{aligned}$$

$$\begin{aligned}\gamma =&((PPase \cdot k_{dephos} + a \cdot k_{ex2}) \cdot (k_{in2} \cdot k_{off} + CIF \cdot k_{in2}^2 + a \cdot k_{ex2} \cdot k_{off} \\ &+ CIF \cdot [S4c]_{ss} \cdot k_{in2} \cdot k_{on} + CIF \cdot a \cdot k_{ex2} \cdot k_{in2} + [S4c]_{ss} \cdot a \cdot k_{ex2} \cdot k_{on})) \\ &/(CIF \cdot [S4c]_{ss} \cdot k_{ex2} \cdot k_{in2} \cdot k_{on})\end{aligned}$$

Physical significance of parameter groups

Next, we would like to develop an intuition for the physical significance of these parameter groups. As discussed above, $\alpha \cdot u$ relates to the amount of nuclear Smad complex, since $\alpha \cdot u/((\alpha+\gamma) \cdot u+\beta)$ is the fraction of Smad2 present as S24n. Thus, as $\alpha \cdot u$ increases relative to $\gamma \cdot u + \beta$, the amount of S24n also increases. By definition, the parameter groups β and $\gamma \cdot u$ capture the remaining input-independent and input-dependent polynomials, respectively. Nevertheless, we would like to understand the physical significance of the parameter groups. We can calculate the amount of unphosphorylated Smad2 as:

$$[S2c]_{ss} + \frac{1}{a}[S2n]_{ss} = \frac{\beta + \delta \cdot u}{\beta + (\alpha + \gamma) \cdot u} S2_{tot} \tag{2.68}$$

$$\delta = PPase \cdot k_{dephos} \cdot \frac{k_{off} + CIF \cdot k_{in2} + CIF \cdot [S4c]_{ss} \cdot k_{on}}{CIF \cdot [S4c]_{ss} \cdot k_{ex2} \cdot k_{on}} \tag{2.69}$$

δ captures the dependence of nuclear, unphosphorylated Smad on the input. With the measured parameters, $\beta \gg \delta \cdot u$, so we have

$$[S2c]_{ss} + \frac{1}{a}[S2n]_{ss} \approx \frac{\beta}{\beta + (\alpha + \gamma) \cdot u} S2_{tot} \tag{2.70}$$

This means that β relates to the amount of unphosphorylated Smad2 in the same sense that $\alpha \cdot u$ relates to nuclear Smad complex. We can also express the remaining Smad2 species as:

$$[pS2c]_{ss} + [S24c]_{ss} + \frac{1}{a}[pS2n]_{ss} = \frac{(\gamma - \delta) \cdot u}{((\alpha + \gamma) \cdot u + \beta)} \cdot S2_{tot} \tag{2.71}$$

However, as δ is of the same order of magnitude as γ, the parameter group γ only loosely relates to these remaining species of Smad2.

Derivation of linear behavior

Within the parameter values measured in cells, the behavior of Smad complex dramatically simplifies. Using the measured values, the parameter groups are

$$\alpha \cdot u = 3.1$$

$$\gamma \cdot u = 1.3$$

$$\beta = 46$$

where we have used a non-saturating input ($u = 0.2$). Therefore, within the physiological regime of parameters, $\beta \gg (\alpha + \gamma) \cdot u$. With this, the denominator in the $[S4n]_{ss}$ equation simplifies, and the concentration of Smad complex becomes a linear function of the input:

$$[S24n]_{ss} \approx \frac{\alpha \cdot S2_{tot}}{\beta} \cdot u \tag{2.72}$$

BIBLIOGRAPHY

[1] Shoeb Ahmed, Kyle G Grant, Laura E Edwards, Anisur Rahman, Murat Cirit, Michael B Goshe, and Jason M Haugh. Data-driven modeling reconciles kinetics of erk phosphorylation, localization, and activity states. *Molecular systems biology*, 10(1):718, 2014. ISSN 1744-4292.

[2] Bree B Aldridge, John M Burke, Douglas A Lauffenburger, and Peter K Sorger. Physicochemical modelling of cell signalling pathways. *Nature cell biology*, 8 (11):1195, 2006.

[3] Louise Ashall, Caroline A. Horton, David E. Nelson, Pawel Paszek, Claire V. Harper, Kate Sillitoe, Sheila Ryan, David G. Spiller, John F. Unitt, and David S. Broomhead. Pulsatile stimulation determines timing and specificity of nf-κb-dependent transcription. *Science*, 324(5924):242–246, 2009. ISSN 0036-8075.

[4] Roi Avraham and Yosef Yarden. Feedback regulation of egfr signalling: decision making by early and delayed loops. *Nature reviews Molecular cell biology*, 12(2):104–117, 2011. ISSN 1471-0072.

[5] Nikolaos Balaskas, Ana Ribeiro, Jasmina Panovska, Eric Dessaud, Noriaki Sasai, Karen M Page, James Briscoe, and Vanessa Ribes. Gene regulatory logic for reading the sonic hedgehog signaling gradient in the vertebrate neural tube. *Cell*, 148(1):273–284, 2012.

[6] Hans Clevers and Roel Nusse. Wnt/β-catenin signaling and disease. *Cell*, 149 (6):1192–1205, 2012. ISSN 0092-8674.

[7] Michael Cohen, Karen M Page, Ruben Perez-Carrasco, Chris P Barnes, and James Briscoe. A theoretical framework for the regulation of shh morphogen-controlled gene expression. *Development*, 141(20):3868–3878, 2014.

[8] Cellina Cohen-Saidon, Ariel A. Cohen, Alex Sigal, Yuvalal Liron, and Uri Alon. Dynamics and variability of erk2 response to egf in individual living cells. *Molecular cell*, 36(5):885–893, 2009. ISSN 1097-2765.

[9] Domitilla Del Vecchio, Alexander J Ninfa, and Eduardo D Sontag. Modular cell biology: retroactivity and insulation. *Molecular systems biology*, 4(1): 161, 2008.

[10] Michele K. Dougherty, Jürgen Müller, Daniel A. Ritt, Ming Zhou, Xiao Zhen Zhou, Terry D. Copeland, Thomas P. Conrads, Timothy D. Veenstra, Kun Ping Lu, and Deborah K. Morrison. Regulation of raf-1 by direct feedback phosphorylation. *Molecular cell*, 17(2):215–224, 2005. ISSN 1097-2765.

[11] Elisenda Feliu and Carsten Wiuf. Variable elimination in chemical reaction networks with mass-action kinetics. *SIAM Journal on Applied Mathematics*, 72(4):959–981, 2012. ISSN 0036-1399.

[12] James E. Ferrell and Ramesh R. Bhatt. Mechanistic studies of the dual phosphorylation of mitogen-activated protein kinase. *Journal of Biological Chemistry*, 272(30):19008–19016, 1997. ISSN 0021-9258.

[13] James E. Ferrell and Eric M. Machleder. The biochemical basis of an all-or-none cell fate switch in xenopus oocytes. *Science*, 280(5365):895–898, 1998. ISSN 0036-8075.

[14] James E Ferrell Jr. Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Current opinion in cell biology*, 14(2):140–148, 2002.

[15] Raphaela Fritsche-Guenther, Franziska Witzel, Anja Sieber, Ricarda Herr, Nadine Schmidt, Sandra Braun, Tilman Brummer, Christine Sers, and Nils Blüthgen. Strong negative feedback from erk to raf confers robustness to mapk signalling. *Molecular systems biology*, 7(1):489, 2011. ISSN 1744-4292.

[16] Sheng Gao, Claudio Alarcón, Gopal Sapkota, Sadia Rahman, Pan-Yu Chen, Nina Goerner, Maria J. Macias, Hediye Erdjument-Bromage, Paul Tempst, and Joan Massagué. Ubiquitin ligase nedd4l targets activated smad2/3 to limit tgf-β signaling. *Molecular cell*, 36(3):457–468, 2009. ISSN 1097-2765.

[17] Lea Goentoro and Marc W. Kirschner. Evidence that fold-change, and not absolute level, of β-catenin dictates wnt signaling. *Molecular cell*, 36(5): 872–884, 2009. ISSN 1097-2765.

[18] Mirko Hekman, Andreas Fischer, Lawrence P. Wennogle, Y. Karen Wang, Sharon L. Campbell, and Ulf R. Rapp. Novel c-raf phosphorylation sites: serine 296 and 301 participate in raf regulation. *FEBS letters*, 579(2):464–468, 2005. ISSN 1873-3468.

[19] Alexander Hoffmann, Andre Levchenko, Martin L. Scott, and David Baltimore. The iκb-nf-κb signaling module: temporal control and selective gene activation. *Science*, 298(5596):1241–1245, 2002. ISSN 0036-8075.

[20] Stefan P. Hoppler and Randall T. Moon. *Wnt signaling in development and disease: molecular mechanisms and biological functions*. John Wiley & Sons, 2014. ISBN 1118444159.

[21] Chi-Ying Huang and James E. Ferrell. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Sciences*, 93 (19):10078–10083, 1996. ISSN 0027-8424.

[22] Gareth J. Inman, Francisco J. Nicolás, and Caroline S. Hill. Nucleocytoplasmic shuttling of smads 2, 3, and 4 permits sensing of tgf-β receptor activity. *Molecular cell*, 10(2):283–294, 2002. ISSN 1097-2765.

[23] Shai Kaplan, Anat Bren, Erez Dekel, and Uri Alon. The incoherent feed-forward loop can generate non-monotonic input functions for genes. *Molecular systems biology*, 4(1):203, 2008.

[24] D. Kimelman and W. Xu. β-catenin destruction complex: insights and questions from a structural perspective. *Oncogene*, 25(57):7482–7491, 2006. ISSN 0950-9232.

[25] Walter Kolch. Coordinating erk/mapk signalling through scaffolds and inhibitors. *Nature reviews Molecular cell biology*, 6(11):827–837, 2005. ISSN 1471-0072.

[26] Giorgio Lagna, Akiko Hata, Ali Hemmati-Brivanlou, and Joan Massagué. Partnership between dpc4 and smad proteins in tgf-β signalling pathways. 1996.

[27] David Lake, Sonia A. L. Corrêa, and Jürgen Müller. Negative feedback regulation of the erk1/2 mapk pathway. *Cellular and Molecular Life Sciences*, 73 (23):4397–4413, 2016. ISSN 1420-682X.

[28] Ethan Lee, Adrian Salic, Roland Krüger, Reinhart Heinrich, and Marc W. Kirschner. The roles of apc and axin derived from experimental and theoretical analysis of the wnt pathway. *PLoS Biol*, 1(1):e10, 2003. ISSN 1545-7885.

[29] Robin E. C. Lee, Sarah R. Walker, Kate Savery, David A. Frank, and Suzanne Gaudet. Fold change of nuclear nf-κb determines tnf-induced transcription in single cells. *Molecular cell*, 53(6):867–879, 2014. ISSN 1097-2765.

[30] Bryan T. MacDonald, Keiko Tamai, and Xi He. Wnt/β-catenin signaling: components, mechanisms, and diseases. *Developmental cell*, 17(1):9–26, 2009. ISSN 1534-5807.

[31] Jeffrey P MacKeigan, Leon O Murphy, Christopher A Dimitri, and John Blenis. Graded mitogen-activated protein kinase activity precedes switch-like c-fos induction in mammalian cells. *Molecular and cellular biology*, 25(11):4676–4682, 2005.

[32] Joan Massagué, Joan Seoane, and David Wotton. Smad transcription factors. *Genes & development*, 19(23):2783–2810, 2005. ISSN 0890-9369.

[33] Francisco J. Nicolás, Karolien De Bosscher, Bernhard Schmierer, and Caroline S. Hill. Analysis of smad nucleocytoplasmic shuttling in living cells. *Journal of cell science*, 117(18):4113–4125, 2004. ISSN 0021-9533.

[34] Roel Nusse and Hans Clevers. Wnt/β-catenin signaling, disease, and emerging therapeutic modalities. *Cell*, 169(6):985–999, 2017. ISSN 0092-8674.

[35] Noah Olsman and Lea Goentoro. Allosteric proteins as logarithmic sensors. *Proceedings of the National Academy of Sciences*, 113(30):E4423–E4430, 2016. ISSN 0027-8424.

[36] Joseph R. Pomerening, Eduardo D. Sontag, and James E. Ferrell. Building a cell cycle oscillator: hysteresis and bistability in the activation of cdc2. *Nature cell biology*, 5(4):346–351, 2003. ISSN 1465-7392.

[37] Kenyi Saito-Diaz, Tony W. Chen, Xiaoxi Wang, Curtis A. Thorne, Heather A. Wallace, Andrea Page-McCaw, and Ethan Lee. The way wnt works: components and mechanism. *Growth factors*, 31(1):1–31, 2013. ISSN 0897-7194.

[38] Herbert M Sauro and Brian Ingalls. Mapk cascades as feedback amplifiers. *arXiv preprint arXiv:0710.5195*, 2007.

[39] Bernhard Schmierer and Caroline S. Hill. Kinetic analysis of smad nucleocytoplasmic shuttling reveals a mechanism for transforming growth factor β-dependent nuclear accumulation of smads. *Molecular and cellular biology*, 25(22):9845–9858, 2005. ISSN 0270-7306.

[40] Bernhard Schmierer, Alexander L. Tournier, Paul A. Bates, and Caroline S. Hill. Mathematical modeling identifies smad nucleocytoplasmic shuttling as a dynamic signal-interpreting system. *Proceedings of the National Academy of Sciences*, 105(18):6608–6613, 2008. ISSN 0027-8424.

[41] Birgit Schoeberl, Claudia Eichler-Jonsson, Ernst Dieter Gilles, and Gertraud Müller. Computational modeling of the dynamics of the map kinase cascade activated by surface and internalized egf receptors. *Nature biotechnology*, 20 (4):370–375, 2002. ISSN 1087-0156.

[42] Guy Shinar, Ron Milo, María Rodríguez Martínez, and Uri Alon. Input–output robustness in simple bacterial signaling systems. *Proceedings of the National Academy of Sciences*, 104(50):19931–19935, 2007.

[43] Jennifer L. Stamos and William I. Weis. The β-catenin destruction complex. *Cold Spring Harbor perspectives in biology*, 5(1):a007898, 2013. ISSN 1943-0264.

[44] O. E. Sturm, R. Orton, J. Grindlay, M. Birtwistle, V. Vyshemirsky, D. Gilbert, M. Calder, A. Pitt, B. Kholodenko, and W. Kolch. The mammalian mapk/erk pathway exhibits properties of a negative feedback amplifier. *Science signaling*, 3(153):ra90, 2010. ISSN 1937-9145.

[45] P. Vizán, D. S. Miller, I. Gori, D. Das, B. Schmierer, and C. S. Hill. Controlling long-term signaling: receptor dynamics determine attenuation and refractory

behavior of the tgf-β pathway. *Science signaling*, 6(305):ra106, 2013. ISSN 1937-9145.

[46] Joseph A Wayman and Jeffrey D Varner. Biological systems modeling of metabolic and signaling networks. *Current Opinion in Chemical Engineering*, 2(4):365–372, 2013.

[47] Angelique Whitehurst, Melanie H. Cobb, and Michael A. White. Stimulus-coupled spatial restriction of extracellular signal-regulated kinase 1/2 activity contributes to the specificity of signal-response pathways. *Molecular and cellular biology*, 24(23):10145–10150, 2004. ISSN 0270-7306.

[48] Lan Xu and Joan Massagué. Nucleocytoplasmic shuttling of signal transducers. *Nature Reviews Molecular Cell Biology*, 5(3):209–219, 2004. ISSN 1471-0072.

[49] Tau-Mu Yi, Yun Huang, Melvin I Simon, and John Doyle. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proceedings of the National Academy of Sciences*, 97(9):4649–4653, 2000.

[50] Seunghee Yoon and Rony Seger. The extracellular signal-regulated kinase: multiple substrates regulate diverse cellular functions. *Growth factors*, 24(1): 21–44, 2006. ISSN 0897-7194.

Chapter 3

LINEARITY IN CELL SIGNALING PATHWAYS: EXPERIMENTS

3.1 Introduction

In the previous chapter, we analyzed models of three signaling pathways and identified a convergent strategy of linear signal transmission. This work is built on the long track record of success in using math modeling and analysis to uncover behaviors of biological networks. In particular, our modeling provided qualitative insights into the input-output response of signaling pathways. In the Wnt, ERK, and Tgfβ pathways, our analysis indicates that the steady-state output is linear with respect to input. Furthermore, modeling also indicates for the ERK and Tgfβ pathways that peak output is linear with respect to transient inputs, e.g. receptor activation followed by attenuation.

In this chapter, we performed quantitative measurements of input-output responses in cell signaling. In addition to validating our model findings, experiments provide further insight to input-output response. For instance, while modeling revealed linearity across the entire dynamic range of three signaling pathways, it did not provide quantitative predictions of how large the dynamic range is, or the precise manner in which saturation occurs for high ligand dose. Experiments also provide insight into dynamic input-output responses, which may be more complex than our steady-state modeling. Lastly, perturbations can be employed to experimentally modulate input-output response, providing insights to each pathway.

To this end, we treated cultured human cells with varying concentrations of ligand and measured pathway response using quantitative Western blot. We used two human cell lines, RKO colon cells and H1299 lung cells, for studying the Wnt pathway and ERK pathway, respectively. Both cell lines have been used as model systems for the respective signaling pathways [1, 3]. We focused our efforts on the Wnt and ERK pathways, since we are limited by available antibodies in the Tgfβ pathway.

Western blots have many benefits as a quantitative tool. In particular, fluorescence-based protein measurements provide sensitive, highly linear measurements of protein level. Standard curves of protein load versus band intensity are used to validate each antibody, confirming that fluorescence increases linearly with protein level and that the signal is not saturated from excess protein. In addition, loading controls are employed to correct for differences in protein loading between lanes. Loading controls are typically constitutively expressed proteins that can be used as a proxy for total protein. The loading control corrects for variations in plating, cell growth due to treatments, and variations in loading volume. Lastly, many different gels, membranes, and lysis protocols can be used to optimize Western blots for size and solubility of protein, making the technique highly versatile. As a result, quantitative measurements can be conducted even on large, membrane associated proteins, such as signaling pathway receptors [6].

We confirmed that Western blotting is a quantitative tool to measure input-output response in the Wnt and ERK pathways. First, we performed standard curves to verify the linearity of each antibody used (Figure S3.1). Next, we verified that our Western blot measurements were reproducible across lanes within each gel, and across gels (Figure S3.2). Lastly, we confirmed that input-output response was not distorted by normalizing to loading controls (un-normalized measurements shown in Figure S3.3) or by averaging input-output responses across experiments (individual experiments shown in Figure S3.4).

In this chapter, we demonstrate linearity in the Wnt and ERK pathway. Interestingly, linearity in the ERK pathway extends to ligand dose, while in the Wnt pathway a nonlinear activation of Wnt co-receptor LRP is followed by linear accumulation of β-catenin. We then used chemical and genetic perturbations to modulate the linear input-output relationship of each pathway.

3.2 Results

Linear signal transmission in the Wnt pathway

To analyze the canonical Wnt pathway, we performed quantitative Western blot measurements in RKO cells, a model system for Wnt signaling. To track the input, we measured the level of phosphorylated Wnt co-receptor LRP5/6 (on Ser1490), which increases within minutes of ligand-receptor complex formation [13]. To track the output, we measured the level of β-catenin. We confirmed that the levels of phosphorylated LRP5/6 and β-catenin increase upon Wnt simulation and reach steady-state within 6 hours (Figure S3.5). Accordingly, all subsequent measurements were done at 6 hours after Wnt stimulation.

To measure the input-output relationship in the Wnt pathway, we treated RKO cells with varying doses of purified Wnt3A and measured how β-catenin (output) correlates with phosphorylated LRP (input). As shown in Figure 3.1A, the level of β-catenin increases linearly with the level of phosphorylated LRP. The linearity persists until saturation of the input, defined as 90% of maximal phosphorylated LRP response (blue circles, Figure 3.1A; Figure S3.6). Notably, at high doses of Wnt3A, β-catenin continues to show incremental activation, despite saturation in phosphorylation of LRP (grey circles, Figure 3.1A). This can be explained within some findings that, while Frizzled/LRP complex is the primary receptor input in β-catenin activation, β-catenin can be activated independently of LRP (e.g., ref. [11]).

Consistent with the mathematical analysis, we observed in RKO cells that the Wnt pathway behaves as a linear transmitter throughout the dynamic range of the input. As a control that is expected from the Michaelis-Menten kinetics that describe ligand binding in the model, we confirmed that the linearity does not extend upstream to Wnt dose: both phospho-LRP5/6 and β-catenin show nonlinear response to Wnt dose (Figure S3.6). Therefore, in the Wnt pathway, a nonlinear ligand-receptor processing step is followed by linear signal transmission through the core intracellular pathway.

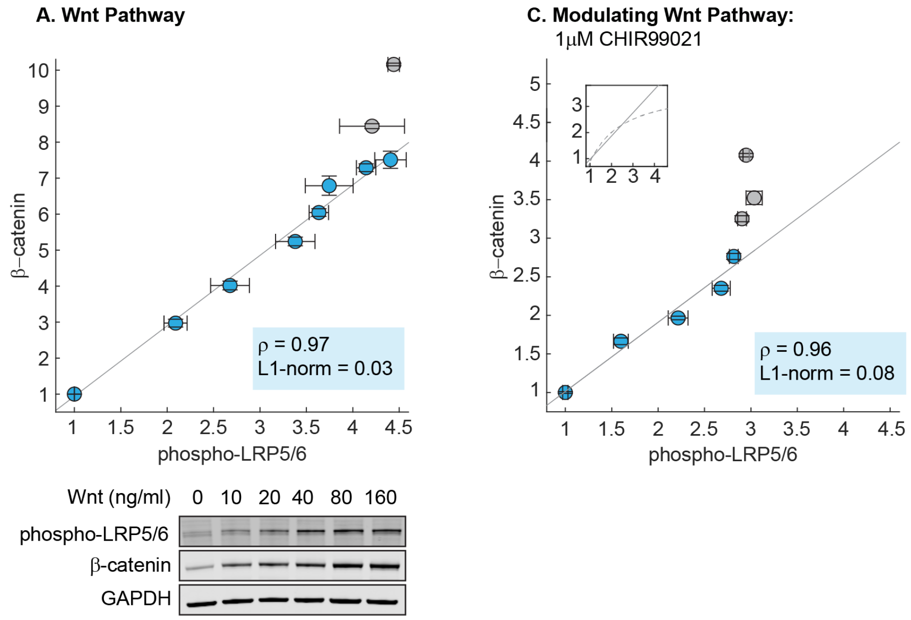


Figure 3.1: **(A)** Measurements of the input-output relationship in the Wnt pathway. In these experiments, RKO cells were stimulated with 0-1280 ng/mL purified Wnt3A ligand, harvested at 6 hours after ligand stimulation, and lysed for Western blot analyses. Shown on top is a representative Western blot. The data plotted come from 7 independent experiments (total N = 66). Each circle indicates the mean intensities of the phospho-LRP5/6 (x-axis) and β-catenin (y-axis) bands for all Western blot biological replicates, and error bars indicate the standard error of the mean. For each gel, we normalize the unstimulated sample (i.e., 0 ng/mL of Wnt3A) to one, and scale the magnitude of the dose response to the average of all gels (described in Methods). The grey line is a least squares regression line, and ρ is the Pearson's coefficient, where $\rho = 1$ is a perfect positive linear correlation. **(B)** As in **(A)**, except that cells were treated with 1μM CHIR99021 (detailed in Methods). The data plotted here come from 5 independent experiments (total N=59). The grey line is a least squares regression , and ρ is the Pearson's coefficient, where $\rho = 1$ is a perfect positive linear correlation. Shown in the subplot are the same least squares regression line (solid line), overlaid with the model prediction (dashed line).

Finally, the analytical expressions we derived in this study not only reveal linear signal transmission, but also the mechanisms by which it arises. In the model of the Wnt pathway, linear transmission occurs due to the futile cycle of β-catenin, in the parameter regime where β-catenin is continually synthesized and rapidly degraded (i.e. $\alpha/u \gg 1 + \gamma$). This regime is not infinite: for instance, a ten-fold decrease in α (e.g. by inhibiting the destruction complex) will break the futile cycle (grey line, Figure 2.2C).

To test if the futile cycle is indeed required for linear signal transmission, we inhibited the destruction complex using CHIR99021, an inhibitor of GSK3β kinase. As before, we measured the input-output relationship, β-catenin vs. phospho-LRP5/6 level, up to 90% of maximal phospho-LRP5/6 input (blue circles, Figure 3.1B). As expected, we found that inhibiting the destruction complex (decreasing α in the model) reduced the range of linearity. The non-treated cells (blue circles, Figure 3.1A) exhibit a linear input-output relationship over a 4.4-fold range of LRP input, whereas the CHIR-treated cells show a linear input-output relationship over only a 2.8-fold range of LRP input (blue circles, Figure 3.1B).

Further, our measurements also reveal an unexpected feature of the Wnt pathway. In the model, inhibiting GSK3β causes β-catenin response to become nonlinear for larger inputs (dashed line, Figure 3.1B subplot). In CHIR-treated RKO cells, however, this nonlinearity cannot be reached, as the maximal amount of phosphorylated LRP (input) is reduced by 50% (grey circles, Figure 3.1B; Figure S3.6), consistent with the dual-function of GSK3β identified by Zeng et al. [15, 16] in phosphorylating β-catenin for degradation as well as phosphorylation LRP for activation. Incorporating this dual-role of GSK3β into the model, we found that this expanded model can indeed recapitulate the data (Figure S3.8) . Therefore, our data indicate two findings: first, that inhibiting GSK3β reduces the range of linear input-output response in the Wnt pathway, as predicted by our analytics, and second, that GSK3β co-regulation of β-catenin and LRP unexpectedly constrains the system within the linear regime.

Linearity in the MAPK/ERK pathway

Next, to measure the input-output relationship in the ERK pathway, we performed quantitative Western blots in H1299 cells, one of the model systems used in the field. Detecting the input level, EGF-activated Ras-GTP, requires a pull-down step that makes it less quantifiable. We therefore tested the dose of EGF ligand itself, since a previous study indicates there could be linearity in ligand-receptor processing [10]. To track the output, we measured the level of doubly-phosphorylated ERK1/2 (on Thr202/Tyr204), dpERK. We first characterized the kinetics of response: dpERK peaks 5 minutes after EGF stimulation (Figure S3.7) and saturates at 4ng/ml EGF (grey circles, Figure 3.2A). Accordingly, all subsequent measurements were performed at 5 minutes after EGF stimulation, and linearity was assessed over the input range of 0-4 ng/mL EGF (blue circles, Figure 3.2A).

We observed linearity in the input-output relationship of the ERK pathway, with the level of dpERK increasing linearly with EGF dose (Figure 3.2A). The linearity holds throughout the dynamic range of the system, over at least 12-fold activation of dpERK. As the ERK pathway is sometimes observed to show bimodal response that would be masked by bulk measurements, we confirmed that the H1299 cells indeed show to graded dpERK response in single-cell level (Figure S3.9), in agreement with a previous single-cell, live imaging study [1]. Therefore, as in the Wnt pathway, signals are transmitted linearly in the ERK pathway throughout the dynamic range of the cell. Moreover, the linearity in the ERK pathway is more extensive than in the Wnt pathway, as linearity extends all the way upstream, such that the level of dpERK directly reflects the dose of extracellular EGF ligand.

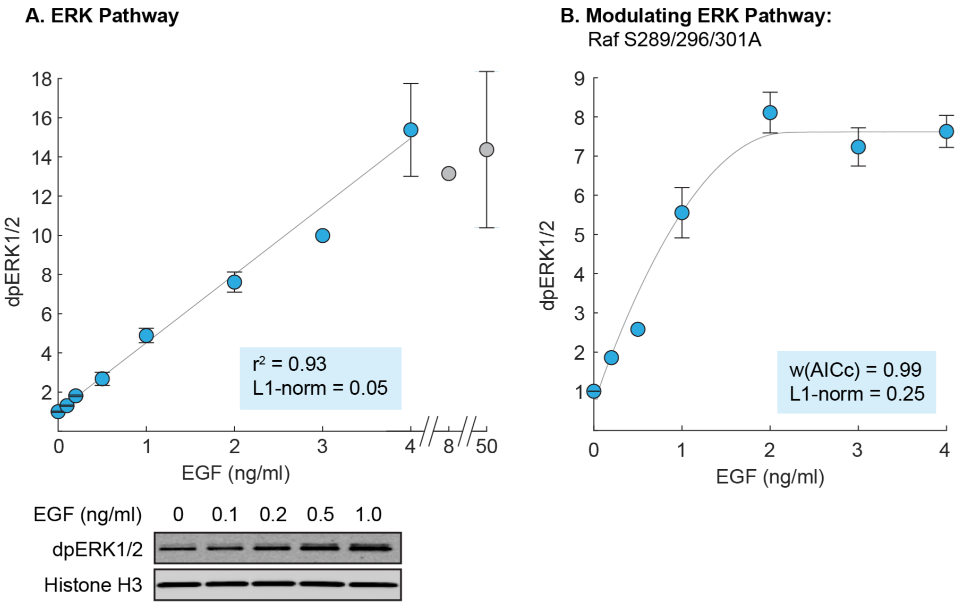


Figure 3.2: **(A)** Measurements of the input-output relationship in the ERK pathway. In these experiments, H1299 cells were stimulated with 0-50 ng/mL purified EGF ligand, harvested at 5 minutes after ligand stimulation, and lysed for Western blot analyses. Shown on top is a representative Western blot. The data plotted here come from 5 independent experiments (total N = 30). Each circle indicates the mean intensities of dpERK1/2 bands across Western blot biological replicates, and the error bars indicate standard error of the mean. Single replicates are plotted without error bars. All data is plotted relative to unstimulated sample. The grey line is a least squares regression line, and r2 is the coefficient of correlation where r2 = 1 is a perfect linear correlation. **(B)** As in **(A)**, but measurements were performed in H1299 cells expressing mutant Raf S289/296/301A. The data plotted here come from 3 independent experiments (total N = 15). The grey line is a fit using the ERK model. We first fitted the gain of the model to the data (i.e., the y-range), and afterward, varied the strength of dpERK feedback (k_{25}) to find the best fit. We used the weighted Akaike Information Criterion, w(AICc), to verify that the nonlinear fit from the ERK model outperforms a linear least squares fit (see Methods). 0< w(AICc)<1, with higher w(AICc) indicates better performance by the non-linear fit.

Next, we examine the requirements for linearity in the ERK pathway. Equation 2.10 reveals that linearity in the ERK pathway depends upon the coupling of strong nonlinearities – ultrasensitivity and negative feedback. As in the Wnt pathway, this regime is not infinite, e.g., decreasing the strength of feedback β enables the system to exit the ultrasensitive regime, and therefore reduces linearity (grey line, Figure 2.3C).

To test this requirement, we examined the effects of weakening the negative feedback.

We created a stable H1299 cell line expressing Raf S289/296/301A, a Raf-1 mutant in which three serine residues that are phosphorylated by dpERK are mutated to alanine [2, 4]. Assessing the dynamic range of the input as before (0-4 ng/mL EGF), we now found that dpERK responds nonlinearly to EGF dose (blue circles, Figure 3.2B), consistent with model predictions (grey line, Figure 3.2B). As a control, we found that overexpressing WT Raf-1 to a similar level does not perturb linearity (experiments, Figure S3.10; modeling, Figure S2.1). Lastly, mutating all 5 direct ERK feedback sites on Raf-1 to alanine had a similar effect to Raf S289/296/301A (Figure S3.11). Our results support the model requirement that strong negative feedback is critical to linear signal transmission in the ERK pathway.

3.3 Discussion

In this chapter, we demonstrated experimental input-output linearity for the Wnt and ERK pathways. We showed how linearity can arise on two vastly different timescales and dynamics. In the ERK pathway, the peak response of dpERK occurs after only five minutes and quickly decays. In the Wnt pathway, β-catenin accumulates to steady state after an hour and remains for high for up to 24 hours [3]. In both cases, however, we demonstrated that the output is linear with respect to input.

We also demonstrate how signal processing may be accomplished at the receptor level. In the ERK pathway, ligand level is transmitted linearly through the receptor such that dpERK is linear with respect to ligand concentration. In the Wnt pathway, however, the ligand-receptor step is nonlinear: β-catenin is linear with active LRP receptor, but nonlinear with respect to Wnt ligand concentration. It is unclear why the receptor exhibits saturating response to ligand, though it could be to prevent hyper-activity of the Wnt pathway.

Interestingly, unlike synthetic circuits whose linearity is often designed to extend across multiple orders of magnitude [8, 9], the linearity we observed in these two natural pathways extends only one order of magnitude, which is also the dynamic range of the pathways. However, we know that natural pathways can convey inputs varying across multiple orders of magnitude, e.g., vision. Thus, an advantage of linearity in natural pathways may be that, in conjunction with fold-change detection at the receptor-level (60), the system as a whole can continually adapt to a given input, hence maintaining sensitivity to future signals.

Why evolve complexity in signaling pathways only to produce seemingly simple behavior? We offer two thoughts. First, complexity of each pathway might afford tunability, in the sense that parameters can be tuned to produce different behaviors in different contexts. For instance, the ERK pathway produces digital, all-or-none response in some contexts [5], and analog response in others [7, 14]. Second – to take an example from engineering – in order to utilize physical processes that are not naturally linear, engineers must implement complex design features to approximate linearity. Similarly, many biochemical processes are inherently nonlinear, meaning that linearity does not arise from a reduction in complexity. Indeed, in each pathway we analyzed here, linearity emerges from complex interactions: a futile cycle in the Wnt pathway, ultrasensitivity coupled to feedback in the ERK pathway, and continual nucleocytoplasmic shuttling in the Tgfβ pathway. Therefore, analogous to engineered systems, complexity in the biochemical pathways we analyzed here might have evolved in part to produce linearity.

3.4 Materials and Methods

Expression Constructs. pBABEpuro-CRAF that contains the wt human Raf-1 clone was a gift from Matthew Meyerson (Addgene plasmid # 51124). Mutant Raf (S289/296/301A) and (S29/289/296/301/642A) were generated using the Q5 site-directed mutagenesis kit (New England Biolabs, E0554S). The mutant and wt Raf-1 were then placed downstream of a CMV promoter.

Cell Lines and Cell Culture. RKO cells (ATCC, CRL-2577) and H1299 cells (ATCC, CRL-5803) were authenticated by STR profiling and supplied by ATCC. RKO cells were cultured at 37°C and 5% (vol/vol) CO2 in DMEM (ThermoFisher Scientific; 11995) supplemented with 10% (vol/vol) FBS (Invitrogen; A13622DJ), 100 U/mL penicillin, 100μg/mL streptomycin, 0.25μg/mL amphotericin, and 2 mML-glutamine (Invitrogen). H1299 cells were cultured at 37C and 5% (vol/vol) CO2 in RPMI (ThermoFisher Scientific; 11875) supplemented with 10% (vol/vol) FBS (Invitrogen; A13622DJ), 100 U/mL penicillin, 100μg/mL streptomycin, 0.25μg/mL amphotericin, and 2 mML-glutamine (Invitrogen). Both cell lines tested negative for mycoplasma contamination.

Transfection of Raf-1 constructs. H1299 cells were transfected with the mutant and wt Raf-1 constructs using Lipofectamine 3000 (ThermoFisher Scientific, L3000). Stable expression was selected using puromycin at a concentration of 1.5 μg/mL for two weeks.

Reagents and Antibodies. The following antibodies were purchased from Cell Signaling Technologies: anti-Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204)

(E10) Mouse mAb #9106, anti-histone H3 (D1H2) XP® Rabbit mAb #4499, anti-c-Raf Antibody #9422, anti-phospho-LRP6 (Ser1490) Antibody #2568, anti-GAPDH (D4C6R) Mouse mAb #97166. Anti-β-catenin mouse mAb was purchased from BD Transduction Laboratories (#610153) and anti-GAPDH rabbit antibody was purchased from Abcam (ab9485). The following fluorescent secondary antibodies were purchased from Fisher Scientific: IRDye 800CW Goat anti-Mouse IgG (926-32210) and IRDye 680LT Goat anti-Rabbit IgG (926-68021). Recombinant human Wnt3A was purchased from Fisher Scientific (5036WN), and recombinant human EGF was purchased from Sigma (E9644). CHIR99021 was purchased from Sigma (SML1046). Halt™ Protease and Phosphatase Inhibitor Cocktail (100X) was purchased from Fisher Scientific (78440).

CHIR99021 Treatment. RKO cells were pre-treated with 1μM CHIR99021 for 24 hours before adding replacement media containing 1μM CHIR99021 and Wnt3A for six hours.

Cell Lysis. RKO cells at 70% confluency were scraped in PBS, pelleted, snap-frozen, and then thawed in NP-40 lysis buffer containing Halt inhibitor cocktail. Samples were spun down, and the supernatants were transferred to Laemmli sample buffer and boiled. The samples were then run onto a Bolt™ 4-12% Bis-Tris Plus Gel (Thermofisher, NW04120BOX). H1299 cells at 70% confluence were scraped in NP-40 lysis buffer containing Halt inhibitor cocktail and further lysed in Laemmli sample buffer. Samples were spun down, and the supernatants were boiled. The samples were then run onto a Novex™ 4-20% Tris-Glycine Mini Gel (ThermoFisher, XP04200BOX).

Quantitative Western blots. Proteins were transferred onto nitrocellulose membranes, blocked for one hour at RT with blocking buffer (Odyssey® Blocking Buffer (TBS) (927-50000) or 5% milk powder in TBS) and stained overnight at 4°C with primary antibody diluted in blocking buffer. The membranes were then stained with fluorescent IR secondary antibodies diluted in blocking buffer for one hour at RT. The fluorescent signal was then imaged using the LiCOR Odyssey Imager and quantified using Odyssey Application software version 3.0. The background-subtracted intensity of the protein bands were normalized to the loading control, GAPDH and/or Histone H3 (for RKO) or Histone H3 (for H1299). These values were then normalized to the reference lanes within each gel, to allow comparison across gels. For β-catenin and phospho-LRP, variation in the fold-activation from experiment to experiment could artificially stretch the data along the x- and y-axis and introduce artifacts into the relationship between phospho-LRP5/6 and β-catenin. Therefore, for Wnt3A dose responses, the data from each gel was scaled such that the mean of 80ng/mL and 160ng/mL samples was equal to the mean across all gels. Finally, for each antibody used in the study, we did careful characterization of the linear range and verified that our measurement conditions were within the linear range of the antibody (Figure S3.1).

Technical variability of Western blot quantitation. To confirm the effects reported, we verified that quantitation of the same sample loaded in multiple lanes in a gel gives CV < 10%, and quantitation of the same sample across multiple independent gels gives CV < 10% (Figure S3.2). As further control, we verified that normalization with loading control did not produce artificial distortion of the input-output relationship: linearity was observed without normalization in cases where loading was already uniform (Figure S3.3).

L-1 and L2-norm analysis. L1-norm analysis was performed as described in ref. [9]. Briefly, the data is fitted with a cubic Hermite polynomial and rescaled along the x and y axis to [0, 1]. The L1-norm is computed as the area between the polynomial fit and the diagonal. Linearity is defined in this context as L1-norm<0.1. L2-norm analysis for Wnt pathway data was performed using a Pearson's coefficient, and L2-norm analysis for ERK pathway data was performed using the coefficient of correlation, r^2.

Akaike Information Criterion. To score the validity of nonlinear model fits for Figure 3.2D, we used the bias-corrected Akaike Information Criterion as described in ref. [12], which assesses goodness-of-fit and model parsimony. The weighted Akaike $w(AIC)$ provides a comparison of all considered models, which in our case is the nonlinear ERK pathway model fit and a linear fit, with the higher score indicating a more valid model.

3.5 Supporting Figures

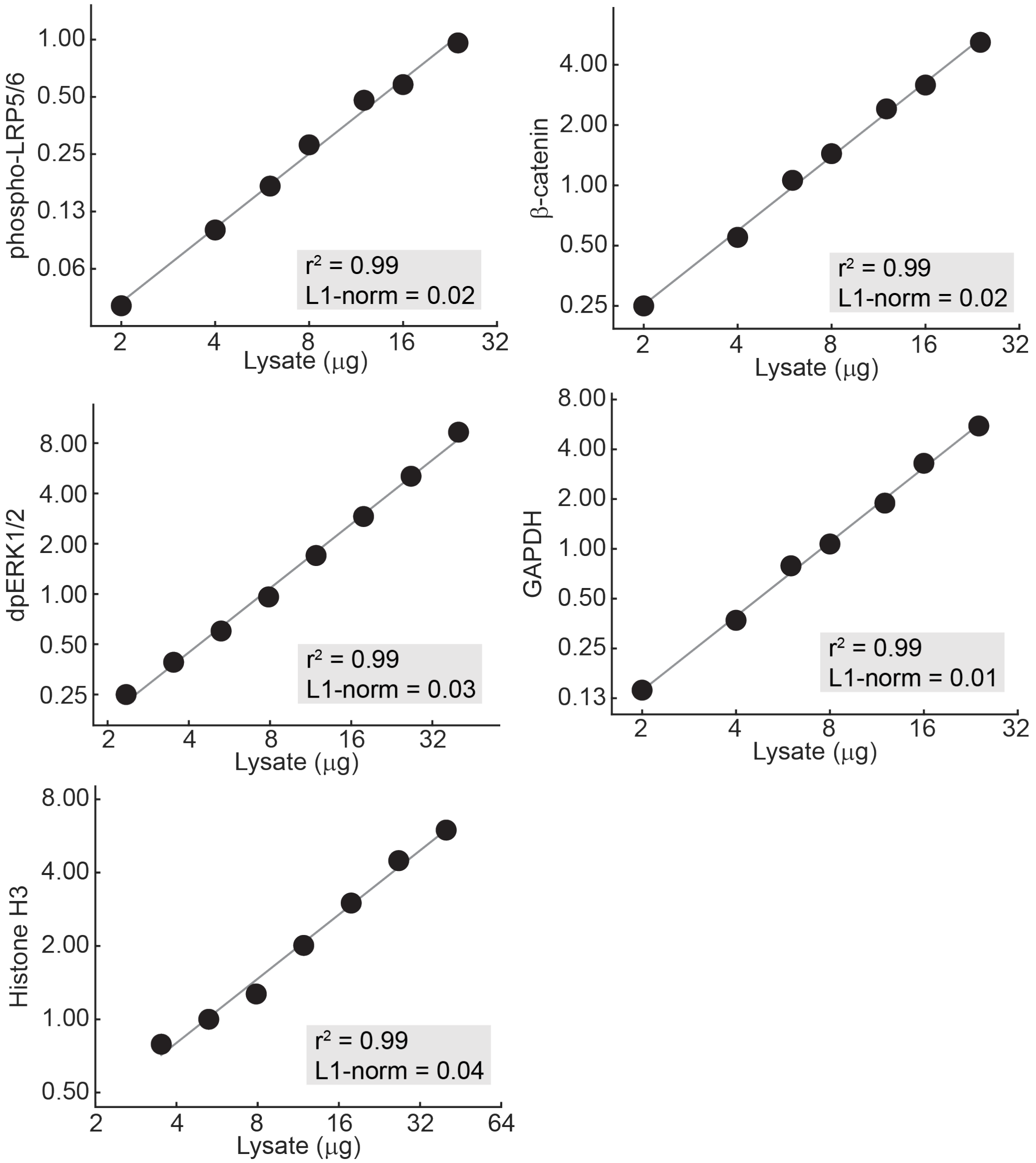


Figure S3.1: Standard curves are linear for all Western blot antibodies. Band intensity was measured against protein load for each antibody, and linearity was assessed by the correlation coefficient r^2 and the L1-norm. All measurements used for this study were within the linear range of antibody.

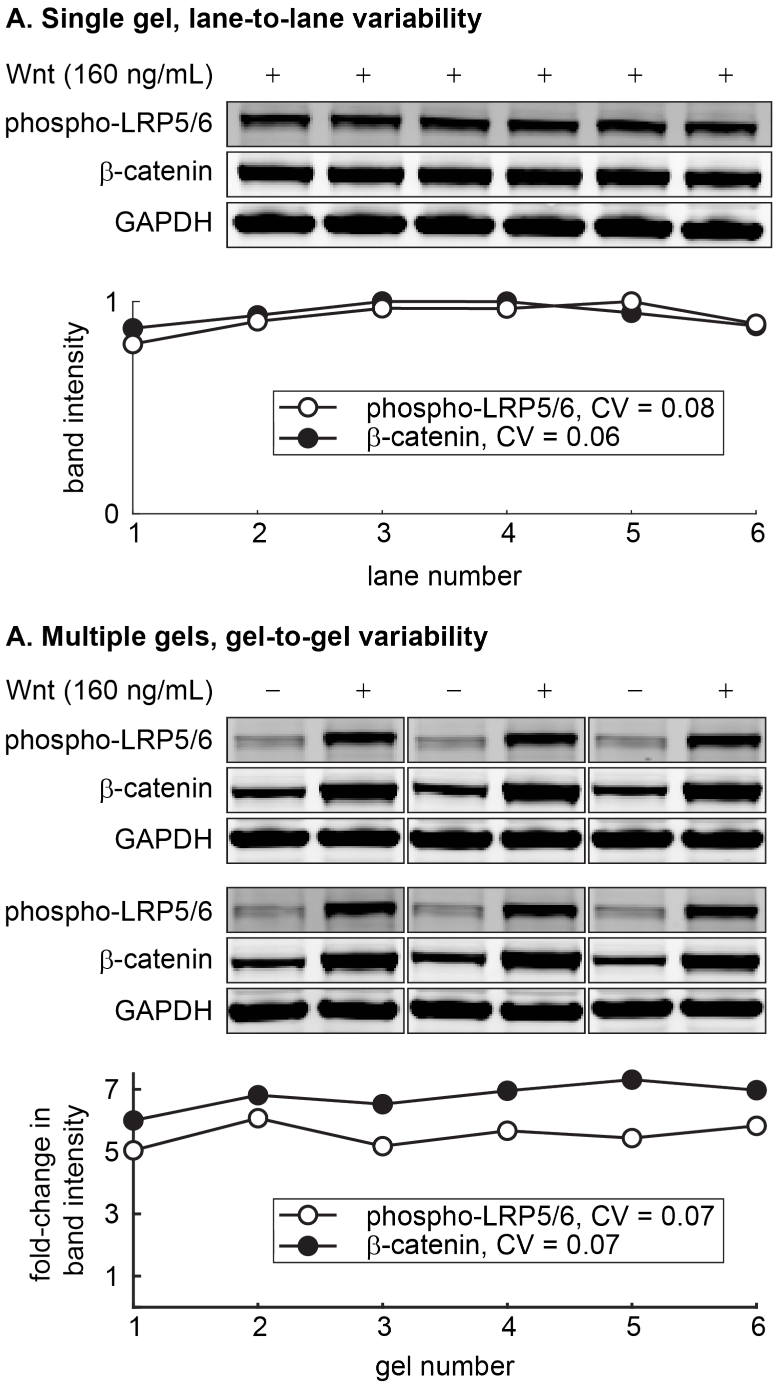


Figure S3.2: **(A)** The level of β-catenin and phosphorylated LRP, measured across different lanes. **(B)** Ligand-stimulated change in β-catenin and phophorylated LRP level, measured in 6 independent Western blots. CV is coefficient of variation, defined as standard deviation/mean.

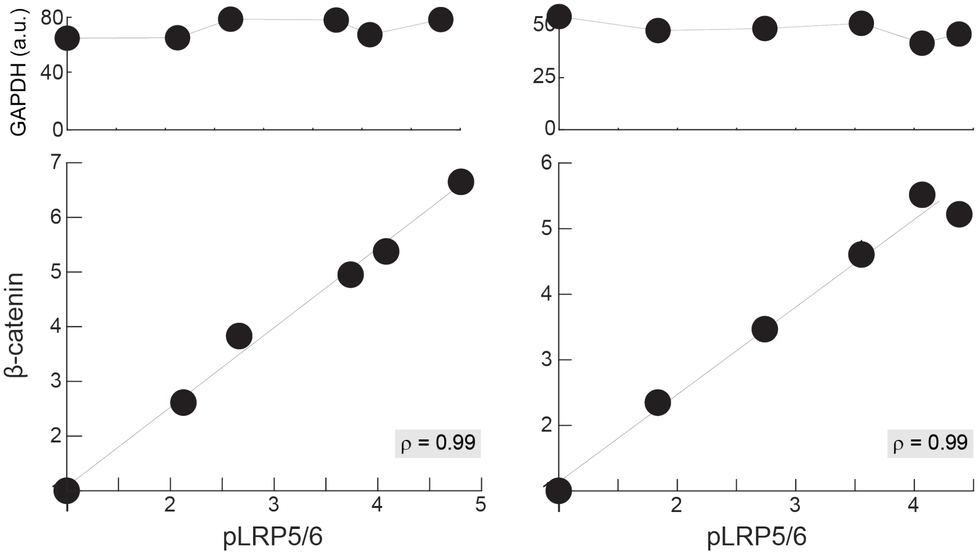


Figure S3.3: In these two independent experiments, RKO cells were stimulated with a range of Wnt3A dose (0-160 ng/mL), the cells lysed after 6 hours and analyzed for Western blot against β-catenin and phosphorylated LRP5/6 (pLRP5/6). Top row: In each experiment, GAPDH intensity varies with <10% CV across samples. Bottom row: Raw β-catenin and LRP intensity data without normalization with GAPDH loading control. The measurements are plotted relative to unstimulated cells. Grey lines are least squares regression lines, and ρ is the Pearson correlation coefficient.

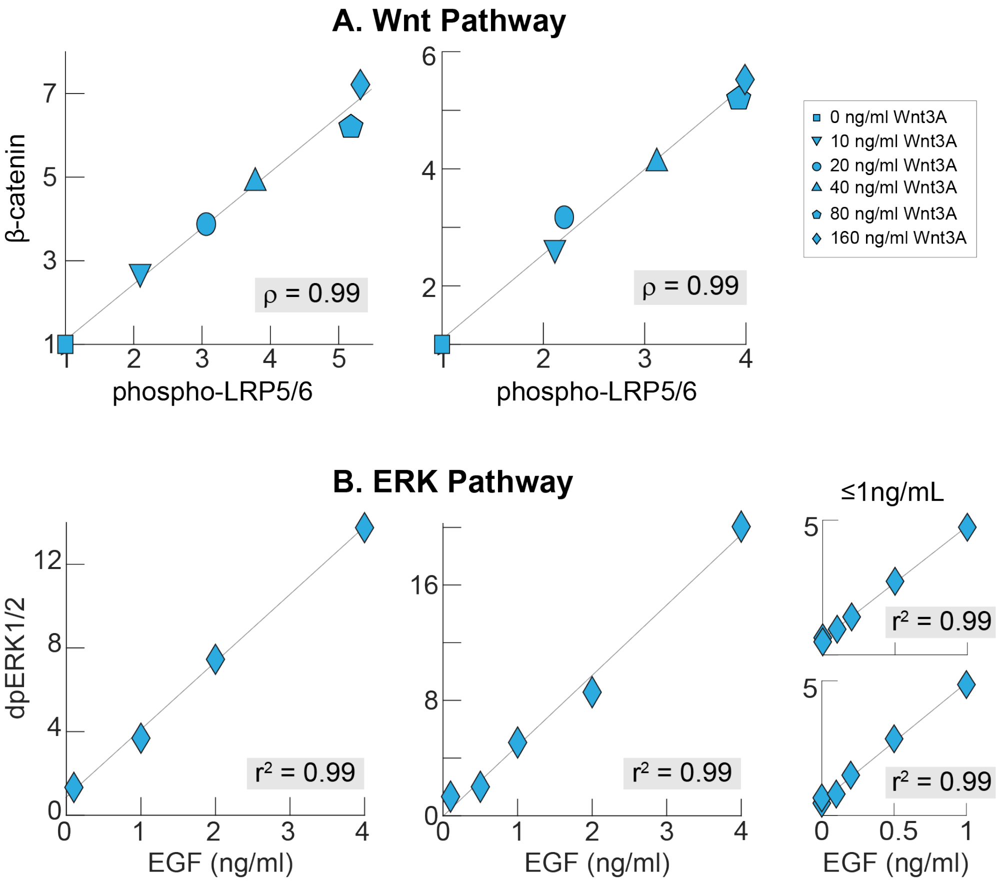


Figure S3.4: Linearity was observed across independent experiments. **(A)** In these 2 independent experiments, RKO cells were stimulated with a range of Wnt3A doses, lysed after 6 hours, and analyzed for Western blot against β-catenin and phospho-LRP5/6. **(B)** In these 4 independent experiments, H1299 cells were stimulated with a range of EGF doses, lysed after 5 minutes, and analyzed for Western blot against doubly-phosphorylated ERK. All measurements are plotted relative to unstimulated cells. Grey lines are least squares regression, ρ is Pearson correlation coefficient, and r^2 is correlation coefficient.

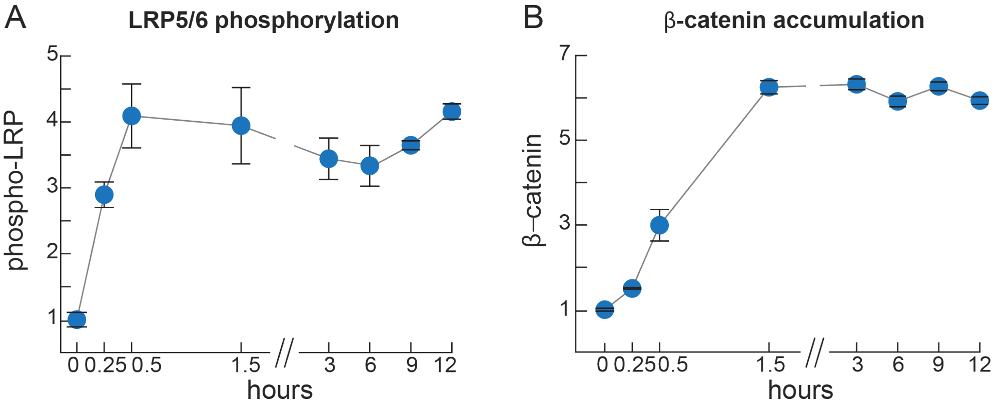


Figure S3.5: LRP5/6 phosphorylation and β-catenin accumulation are already at steady state at 6 hours after Wnt stimulation. RKO cells were treated with 160 ng/mL Wnt3A for the specified times, and then assayed for phospho-LRP5/6 and β-catenin level by Western blot. Error bars are standard error of the mean from 2-4 biological replicates. Data are plotted relative to the sample at time zero, and normalized to the average maximal activation across experiments. The grey lines connect the mean of each time point.

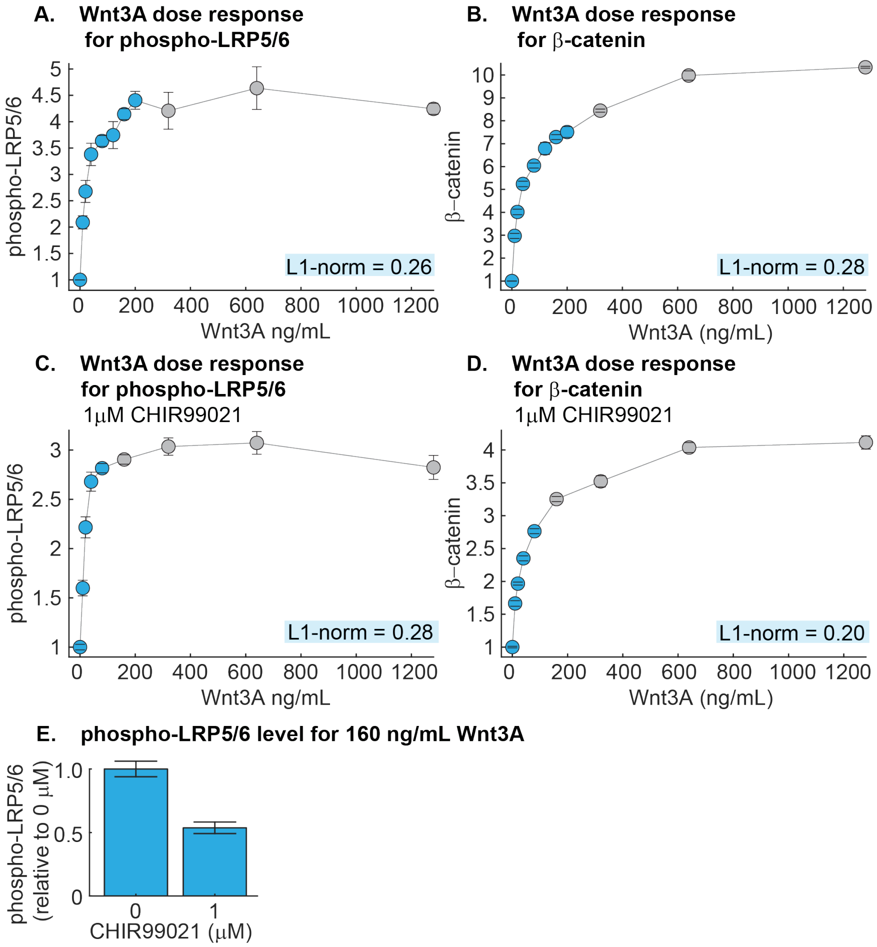


Figure S3.6: The dynamic range of Wnt signaling in RKO cells. RKO cells were treated with the specified dose of Wnt3A for six hours, and then assayed for phospho-LRP5/6 and β-catenin by quantitative Western blot. Data are plotted relative to unstimulated samples. **(A-B)** In wt cells, phospho-LRP5/6 (A) shows >90% of maximal response at 200 ng/mL Wnt3A, while β-catenin (B) shows 70% of maximal response at 200 ng/mL and subsequently incremental response until 640 ng/mL Wnt3A. (C-D) In cells pre-treated with 1μM CHIR99021, phospho-LRP5/6 (C) shows >90% of maximal response at 80ng/mL Wnt3A, while β-catenin (D) shows 70% of maximal response at 80 ng/mL and continues incremental activation at higher doses. (E). Cells were treated with 160 ng/mL Wnt3A and assayed for phospho-LRP5/6. Cells pre-treated with 1μM CHIR99021 (N = 3) exhibited 50% the level of phospho-LRP5/6 as untreated cells (N = 3). The grey lines simply connect the means of data.

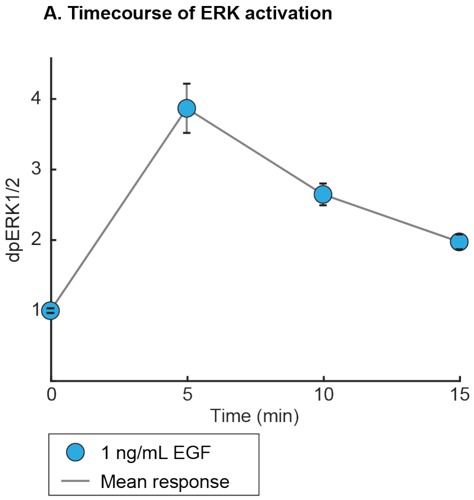


Figure S3.7: ERK activation peaks at 5 minutes after EGF stimulation. H1299 cells were treated with 1 ng/mL EGF for the specified times, and then assayed for dpERK1/2 by Western blot. Data is plotted relative to the samples at time zero, with at least three biological replicates per time point.

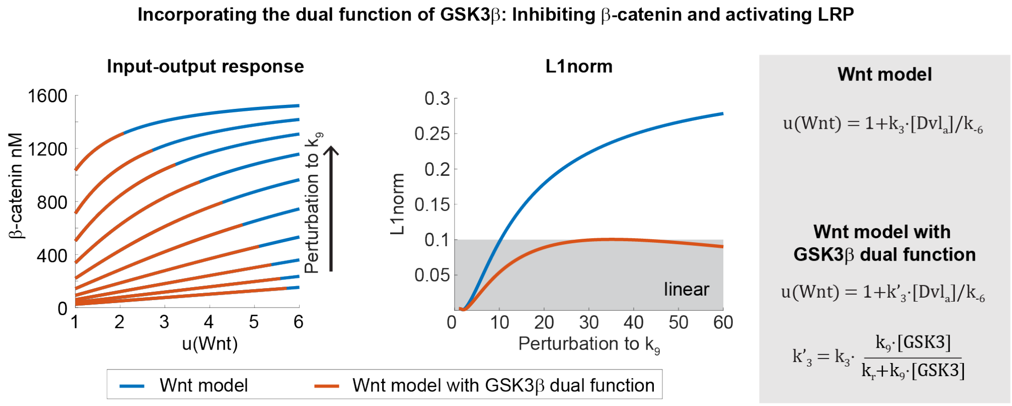


Figure S3.8: Incorporating into the Wnt model the dual function of GSK3β in phosphorylating β-catenin and LRP5/6. We include the role of GSK3β in phosphorylating LRP5/6 into the input function $u(Wnt)$, such that u(Wnt) is a function of GSK3β and the phosphorylation rate k¬9 (for simplicity, the same rate as GSK3β phosphorylation of β-catenin), and the reverse rate k¬r. In both models, β-catenin increases in response to GSK3β inhibition (e.g., by CHIR99021). However, only the model with the dual function of GSK3β shows a decrease in input range that we observed experimentally.

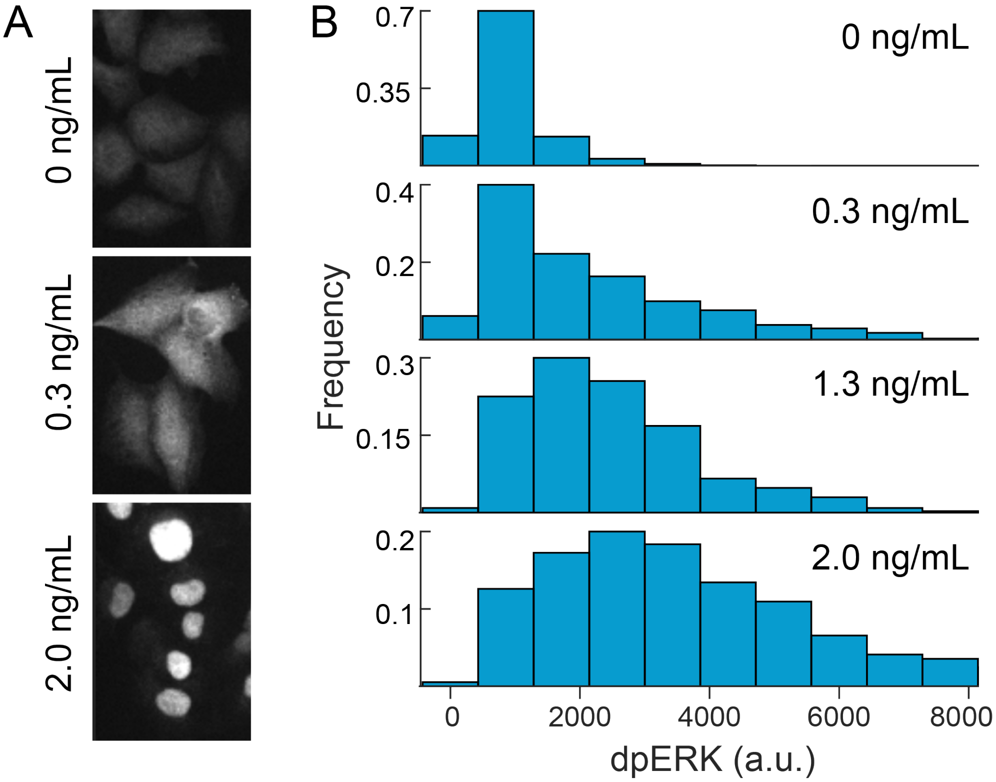


Figure S3.9: In these experiments, H1299 cells were treated with varying doses of EGF for 5 minutes and then fixed and analyzed for immunofluorescence against doubly phosphorylated ERK (dpERK). **(A)** Representative images of cells treated with the indicated doses of EGF. **(B)** The intensity of nuclear level of dpERK staining across individual cells. Cell nuclei were delineated using DAPI staining (for EGF doses 0, 0.3, 1.3, and 2.0 ng/mL, N = 453, 381, 373, and 413 cells, respectively).

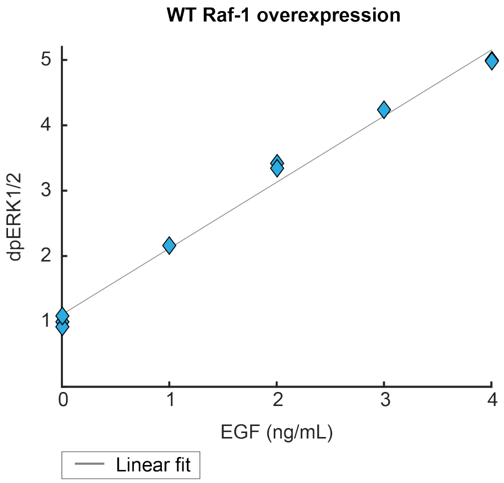


Figure S3.10: WT Raf-1 overexpression does not affect linear dose-response. H1299 cells over-expressing Raf-1 were treated with the indicated dose of EGF for 5 minutes, and then assayed for dpERK1/2 by Western blot. The grey line is a fit from a linear model with r2 = 0.99. Data is plotted relative to unstimulated samples, with total N = 9.

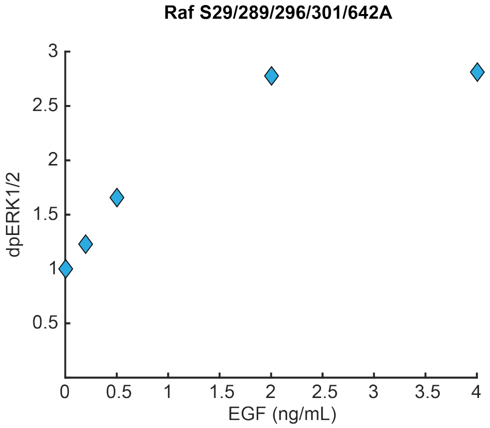


Figure S3.11: Expression of Raf S29/289/296/301/642A induces non-linear dose-response. H1299 cells expressing the Raf mutant Raf S29/289/296/301/642 were treated with the indicated dose of EGF for five minutes, and then assayed for dpERK1/2 by Western blot. Data is plotted relative to unstimulated samples, with total N = 5.

BIBLIOGRAPHY

[1] Cellina Cohen-Saidon, Ariel A. Cohen, Alex Sigal, Yuvalal Liron, and Uri Alon. Dynamics and variability of erk2 response to egf in individual living cells. *Molecular cell*, 36(5):885–893, 2009. ISSN 1097-2765.

[2] Michele K. Dougherty, Jürgen Müller, Daniel A. Ritt, Ming Zhou, Xiao Zhen Zhou, Terry D. Copeland, Thomas P. Conrads, Timothy D. Veenstra, Kun Ping Lu, and Deborah K. Morrison. Regulation of raf-1 by direct feedback phosphorylation. *Molecular cell*, 17(2):215–224, 2005. ISSN 1097-2765.

[3] Lea Goentoro and Marc W. Kirschner. Evidence that fold-change, and not absolute level, of β-catenin dictates wnt signaling. *Molecular cell*, 36(5): 872–884, 2009. ISSN 1097-2765.

[4] Mirko Hekman, Andreas Fischer, Lawrence P. Wennogle, Y. Karen Wang, Sharon L. Campbell, and Ulf R. Rapp. Novel c-raf phosphorylation sites: serine 296 and 301 participate in raf regulation. *FEBS letters*, 579(2):464–468, 2005. ISSN 1873-3468.

[5] Chi-Ying Huang and James E. Ferrell. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Sciences*, 93 (19):10078–10083, 1996. ISSN 0027-8424.

[6] Kevin A Janes. An analysis of critical factors for quantitative immunoblotting. *Sci. Signal.*, 8(371):rs2–rs2, 2015.

[7] Jeffrey P MacKeigan, Leon O Murphy, Christopher A Dimitri, and John Blenis. Graded mitogen-activated protein kinase activity precedes switch-like c-fos induction in mammalian cells. *Molecular and cellular biology*, 25(11):4676–4682, 2005.

[8] Dmitry Nevozhay, Rhys M Adams, Kevin F Murphy, Krešimir Josić, and Gábor Balázsi. Negative autoregulation linearizes the dose–response and suppresses the heterogeneity of gene expression. *Proceedings of the National Academy of Sciences*, 106(13):5123–5128, 2009.

[9] Dmitry Nevozhay, Tomasz Zal, and Gábor Balázsi. Transferring a synthetic gene circuit from yeast to mammalian cells. *Nature communications*, 4:1451, 2013. ISSN 2041-1723.

[10] Diego A. Oyarzún, Jo L. Bramhall, Fernando López-Caamal, Frances M. Richards, Duncan I. Jodrell, and Ben-Fillippo Krippendorff. The egfr demonstrates linear signal transmission. *Integrative Biology*, 6(8):736–742, 2014.

[11] Michael Rotherham and Alicia J El Haj. Remote activation of the wnt/β-catenin signalling pathway using functionalised magnetic particles. *PloS one*, 10(3): e0121761, 2015. ISSN 1932-6203.

[12] Andrej-Nikolai Spiess and Natalie Neumeyer. An evaluation of r 2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a monte carlo approach. *BMC pharmacology*, 10(1):6, 2010. ISSN 1471-2210.

[13] Keiko Tamai, Xin Zeng, Chunming Liu, Xinjun Zhang, Yuko Harada, Zhijie Chang, and Xi He. A mechanism for wnt coreceptor activation. *Molecular cell*, 13(1):149–156, 2004. ISSN 1097-2765.

[14] Angelique Whitehurst, Melanie H. Cobb, and Michael A. White. Stimulus-coupled spatial restriction of extracellular signal-regulated kinase 1/2 activity contributes to the specificity of signal-response pathways. *Molecular and cellular biology*, 24(23):10145–10150, 2004. ISSN 0270-7306.

[15] Xin Zeng, Keiko Tamai, Brad Doble, Shitao Li, He Huang, Raymond Habas, Heidi Okamura, Jim Woodgett, and Xi He. A dual-kinase mechanism for wnt co-receptor phosphorylation and activation. *Nature*, 438(7069):873–877, 2005. ISSN 0028-0836.

[16] Xin Zeng, He Huang, Keiko Tamai, Xinjun Zhang, Yuko Harada, Chika Yokota, Karla Almeida, Jianbo Wang, Brad Doble, and Jim Woodgett. Initiation of wnt signaling: control of wnt coreceptor lrp6 phosphorylation/activation via frizzled, dishevelled and axin functions. *Development*, 135(2):367–375, 2008. ISSN 0950-1991.

Chapter 4

LINEARITY UNDERLIES FOLD-CHANGE RESPONSE IN THE WNT PATHWAY

4.1 Introduction

What are potential advantages to linear signal transmission? Linearity is a feature of many engineering systems, where it serves several practical purposes. In particular, linear signal transmission enables the superposition of multiple signals, where the output of two simultaneous inputs is equal to the sum of the outputs for each input separately. Superposition enables multiple, dynamic signals to be faithfully transmitted and processed independently [21]. Thus, for instance, linearity enables people to listen to a phone call and interpret speech amongst background noise and allows a car radio to tune into one station out of multiple broadcasting on separate carrier frequencies. Notably, linearity is also a desired goal in synthetic biology, where it is often implemented using negative feedback [7, 18]. Analogous to engineered circuits, linearity in biological signaling pathways may facilitate multiplexing inputs into a single pathway (Figure 4.1A).

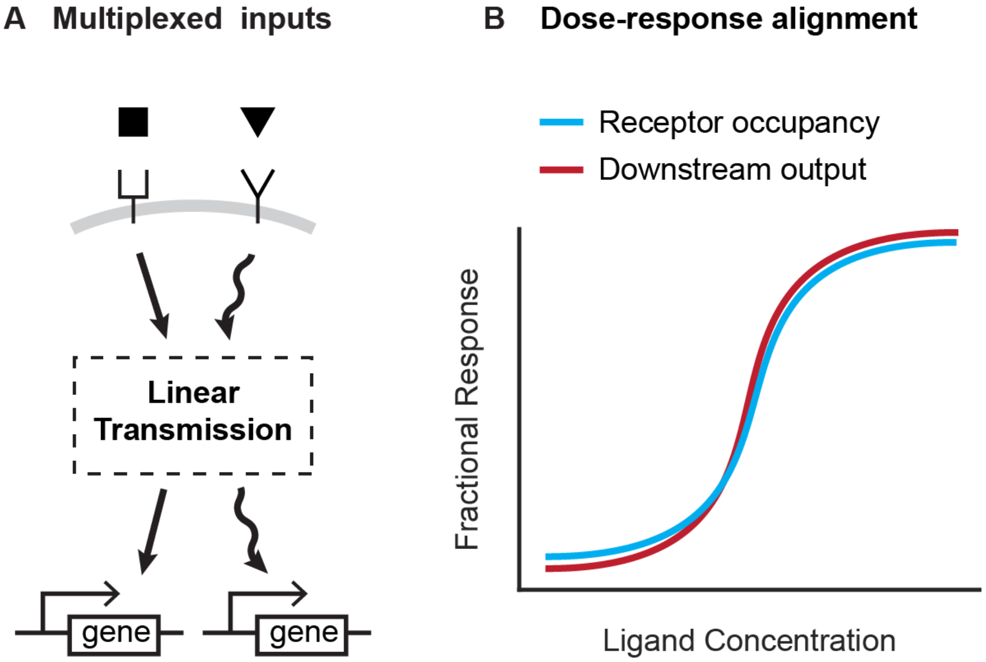


Figure 4.1: Benefits of linearity. **(A)** Linearity enables multiplexing of inputs to a signaling pathway. Multiplexed signals can be independently decoded downstream, and therefore regulate distinct transcriptional events. **(B)** Illustration for how linearity between the receptor occupancy and downstream outputs gives rise to dose-response alignment [2].

A second benefit is that a linear transmitter naturally gives rise to dose-response alignment [2], where one or more downstream responses of a pathway closely follows the fraction of occupied receptor (Figure 4.1B). Dose response alignment appears in many biological systems and is thought to improve the fidelity of information transfer through signaling pathways [2, 3, 19, 24].

Here, we investigate a third potential advantage of linearity that may be useful in biological contexts. We find that linearity facilitates fold change detection, where cells sense fold changes in signal, rather than absolute level, to buffer cellular noise [6, 8, 10, 17, 22]. Indeed, for the signaling pathways studied here, it has been shown experimentally that the robust outcome of ligand stimulation is the fold-change in the level of transcriptional regulator [6, 8, 10, 17]. We then demonstrate how fold-change detection may be implemented through a novel gene regulatory element.

4.2 Results

Linearity underlies fold-change response in the Wnt pathway

Increasingly, studies now indicate that many signaling pathways transmit signals relative to their basal state [6, 8, 10, 17]. A common outcome of the transduction process is an increase in the concentration of a transcriptional regulator that then activates or represses target genes. Interestingly, rather than regulate the absolute concentration (Figure 4.2A), many signaling pathways regulate fold-changes in concentration of the transcriptional regulator (Figure 4.2B) . Such a fold-change response, where the fold-change in transcriptional regulator is robust to variations in biochemical parameters, has been observed in the canonical Wnt [10], ERK [6], NF-κB [17], calcium [22], and Tgfβ pathways [8], as well as predicted in cytokine signaling in T-cells [23]. However, it is not understood how the signaling pathways produce fold-change response in first place. As illustrated in Figure 4.2B, producing fold-change response is not trivial, as the signaling pathways must produce response that is directly proportional to basal state, and do so in a robust way.

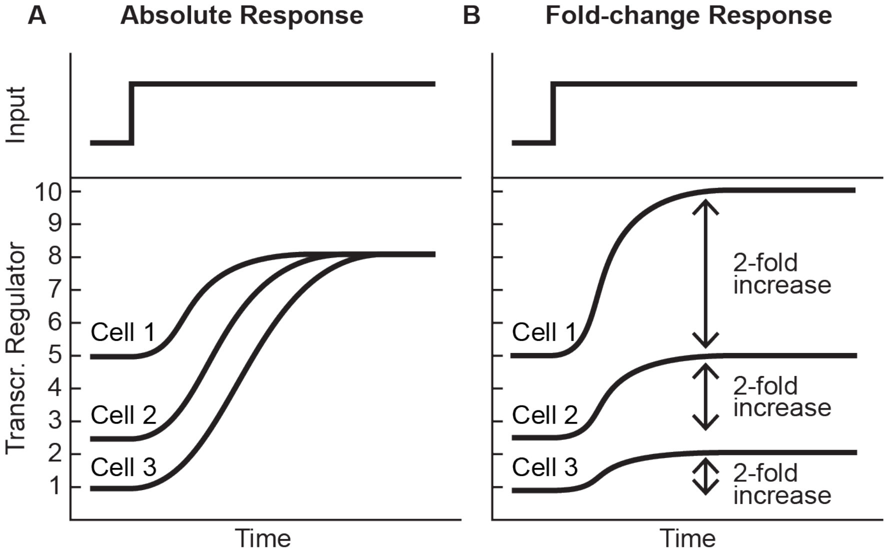


Figure 4.2: **(B-C)** In these illustrations, each line is a response from a single cell or average bulk response given biochemical perturbations. **(B)** With absolute response, concentration of the transcriptional regulator is the outcome of signal transmission that is robust to biochemical variations. **(C)** With fold-change response, fold-change in concentration of the transcriptional regulator is robust to biochemical variations.

Fold-change response in the Wnt pathway was characterized in the seminal work of Goentoro and Kirschner [10]. In this study, simulations of the Wnt pathway predicted that the fold-change in β-catenin is robust to most perturbations to the pathways. In contrast, the absolute level of β-catenin was predicted to be sensitive to variations in many pathway parameters. These predictions were then verified experimentally in *Xenopus* embryos and human RKO cells by perturbing components of the β-catenin destruction complex. Inhibition of GSK3 through lithium addition or overexpression of GSK3 binding protein (GBP) increased β-catenin levels, while Axin1 overexpression lowered β-catenin levels. In contrast, the fold-change in β-catenin was demonstrated to be insensitive to each perturbation [10].

To understand how fold-change response arises in the Wnt pathway, we examined the analytical solution to β-catenin concentration at steady-state, derived in ref. [10] and reproduced in Equation 2.1:

$$[\beta cat]_{ss} = K_{17} \cdot \frac{1-\gamma+\alpha/u}{2}\left(\sqrt{1+\frac{4\gamma}{(1-\gamma+\alpha/u)^2}}-1\right) \tag{2.1}$$

Here the input function $u(Wnt)$ describes the effect of Wnt stimulation on inhibiting the destruction complex and is nonzero in the absence of Wnt (refs. [10, 13, 16]). Basal input can arise from consitutive ligand expression, leaky kinase activity, or in the case of the Wnt model, basal dissociation of GSK3 from the β-catenin destruction complex (parameter k_{-6}), as observed in *Xenopus* egg extracts [16].

Equation 2.1 demonstrates that, in general, fold-change in β-catenin, $FC = [\beta cat]_{ss}(u_1) / [\beta cat]_{ss}(u_0)$, is sensitive to changes in all parameters in the system. Next, we examine fold-change behavior within the parameter regime measured in cells. As shown in Equation 2.4, with measured parameters β-catenin becomes a linear function of input:

$$[\beta cat]_{ss} \approx K_{17}\frac{\gamma}{\alpha}u \tag{2.4}$$

As a consequence of this linearity, the fold-change in β-catenin,

$$FC = \frac{[\beta cat]_{ss}(u_1)}{[\beta cat]_{ss}(u_0)} \approx \frac{K_{17}\frac{\gamma}{\alpha}u_1}{K_{17}\frac{\gamma}{\alpha}u_0} = \frac{u_1}{u_0} \tag{4.3}$$

becomes insensitive to variations in parameters – consistent with experimental observations and simulations of the model [10].

Thus, we find that linear signal transduction that we identified through modeling and experiments in Chapters 2-3 is sufficient to explain the complex relationship between the Wnt pathway and perturbation (Figure 4.3A-B). In particular, Equation 2.4 (reproduced above) demonstrates how the level of β-catenin is directly sensitive to parameter values, and Equation 4.3 demonstrates how fold-change in β-catenin is robust to perturbations to the Wnt pathway.

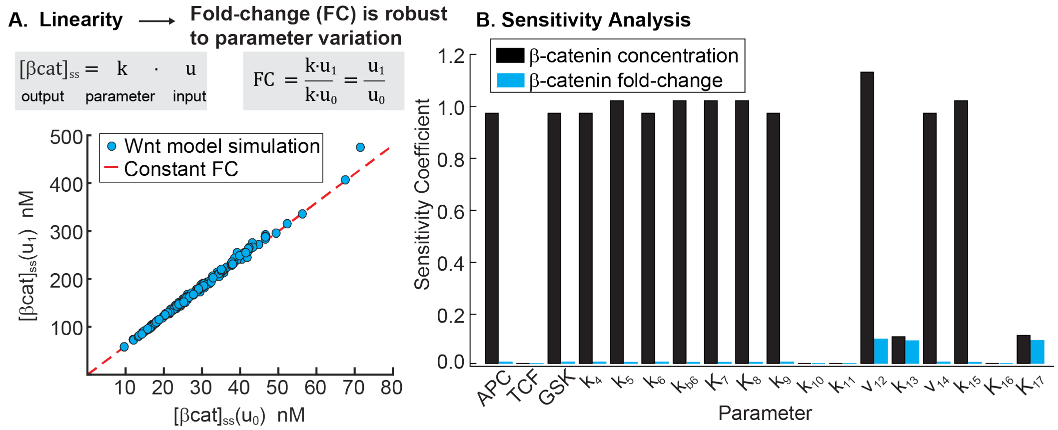


Figure 4.3: **(A)** Linearity produces fold-change response that is robust to parameter variations. To illustrate this, we added lognormal noise (0.1 CV) to all parameters of the Wnt model, and simulated the level of β-catenin before and after Wnt stimulation (blue circles). **(B)** In contrast to β-catenin fold-change, the concentration of β-catenin after Wnt signaling is directly sensitive to parameter variations. Sensitivity coefficients for β-catenin fold-change are plotted in blue, and sensitivity coefficients for β-catenin concentration are plotted in black. The sensitivity coefficient for output y and parameter k is defined as $s_k^y = \frac{\Delta y/y}{\Delta p/p}$

Fold-changes in β-catenin can be detected by novel gene regulatory element

In the previous section, we found that an additional benefit of linear signal transmission is to produce fold-change response. By computing fold-change response, gene regulatory networks can faithfully respond to ligand-receptor activation despite perturbations to the signaling pathway. Therefore, fold-change response as a mechanism for robust gene expression may act as a selective pressure for linearity in cell signaling pathways. In this section, we explore whether fold-change response is in fact utilized by gene regulatory networks downstream of the Wnt pathway.

Signal in the Wnt pathway is transduced by β-catenin, which in complex with Tcf/Lef regulates transcription primarily through a single cis element in the DNA, the Wnt responsive element (WRE), at times potentiated by a nearby Helper site [5]. Recently, an additional regulatory element was identified that regulates Wnt target genes such as *siamois* and *Xnr3* [15]. The 11-bp cis element is hundreds of basepairs upstream of the WRE and is acted upon by β-catenin and Tcf. Unlike the WRE, the 11-bp exerts a repressive effect on expression. This suggests a new model of gene regulation in the Wnt pathway, where β-catenin not only activates target genes through the WRE but also, in some contexts, negatively regulates expression through a distinct 11-bp element. The 11-bp elements co-localized with the WRE in 45%–71% of peaks from β-catenin ChIP sequencing in human cells, suggesting a widespread role for the mechanism [15].

A network in which input exerts both an activating and repressive action on a downstream output is known as an incoherent feed-forward loop (IFFL) [1]. The discovery of the novel 11-bp element suggests that β-catenin acts both as an activator and repressor of gene expression, forming an IFFL. Here, we incorporate the IFFL motif into an established model of Wnt signaling [16] in order to examine gene expression downstream of Wnt signaling.

The Wnt model was built on parameters measured in *Xenopus* extract, and reproduces experimental data from *Xenopus* and mammalian systems [16]. Here, we examine how β-catenin signal is transmitted through the combined Wnt pathway and β-catenin IFFL circuit (Figure 4.4A).

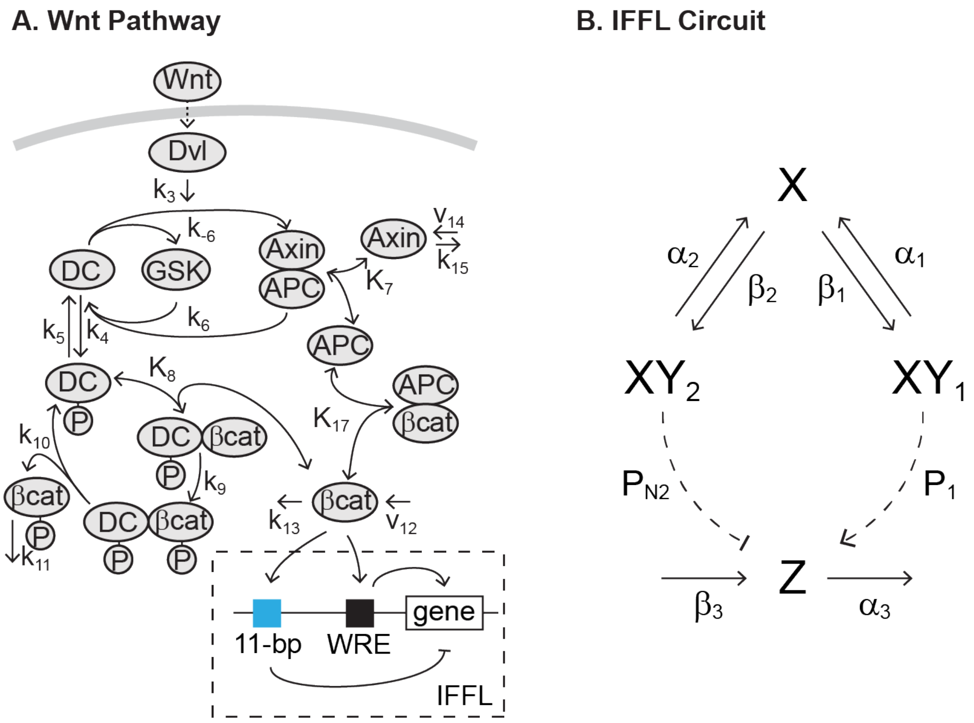


Figure 4.4: Wnt pathway with IFFL circuit **(A)** Diagram of the Wnt pathway, modified to include the gene regulatory IFFL circuit. **(B)** Diagram of the IFFL circuit. X is β-catenin, and forms protein complexes XY_1 and XY_2. XY_1 activates gene output Z by binding to the WRE, while XY_2 represses output by binding to the 11-bp element.

We modeled the IFFL circuit using a simple, phenomenological description previously derived in refs. [4, 11] (Figure 4.4B). To start, we modeled the dynamics of the β-catenin/Tcf-1 complex. This complex binds to the Wnt-responsive element (WRE):

$$\frac{dXY_1}{dt} = \beta_1 \cdot X - \alpha_1 \cdot XY_1 \tag{4.4}$$

XY_1 is the β-catenin/Tcf-1 complex, X is β-catenin, and Y_1 is Tcf-1. α_1 is the dissociation rate constant (k_{off}) of XY_1, and β_1 is the product of the binding rate constant (k_{on}) of X to Y_1 and concentration of Y_1. In the original Wnt model [16], Tcf-1 is present at much lower concentration than β-catenin, with Tcf-1/β-catenin complex increasing by only 1.8-fold upon Wnt signaling. Here, we model Tcf-1 in excess to capture the large dynamic range of Wnt-responsive gene expression [10].

Next, we implemented cis-regulation from the 11-bp element through a second complex, XY_2. While it is known that β-catenin binds to the 11-bp element, it is not currently known whether it is in complex with Tcf-1 or another protein, such as the repressive Tcf-3. We model XY_2 as a function of β-catenin, X, and an unknown binding partner Y_2:

$$\frac{dXY_2}{dt} = \beta_2 \cdot X - \alpha_2 \cdot XY_2 \tag{4.5}$$

α_2 and β_2 are defined in the same way as α_1 and β_2, respectively, applied to the binding partner Y_2.

XY_1 binds to the WRE and activates gene Z, whereas XY_2 binds to 11-bp element and represses gene Z. The dynamics of gene transcription are

$$\frac{dZ}{dt} = \beta_3 \cdot P_a - \alpha_3 \cdot Z \tag{4.6}$$

where β_3 is the rate of transcript production per unit time, α_3 is the degradation rate constant of the transcript Z, and P_a is the probability that transcription is active. In many promoters the WRE and 11-bp element are >500 base-pairs apart, so we modeled WRE occupancy by XY_1 and 11-bp occupancy by XY_2 as independent events. With independent binding of the activator and repressor, the probability function P_a is, as derived in refs. [S2-3],

$$P_a = (P_1) \cdot (P_{N2}) \tag{4.7}$$

$$P_1 = \frac{XY_1/K_1}{1 + XY_1/K_1} \tag{4.8}$$

$$P_{N2} = \frac{1}{1 + XY_2/K_2} \tag{4.9}$$

where P_1 is the probability that XY_1 is bound to WRE and P_{N2} is the probability that XY_2 is not bound to 11-bp element. K_1 is the dissociation constant of XY_1 to WRE and K_2 is the dissociation constant of XY_2 to 11-bp element. The full dynamic equation for Z is:

$$\frac{dZ}{dt} = \beta_3 \cdot \frac{XY_1/K_1}{1 + XY_1/K_1 + XY_2/K_2 + (XY_1 \cdot XY_2)/(K_1 \cdot K_2)} - a_3 \cdot Z \tag{4.10}$$

This trancriptional IFFL belongs to a class of recurring strategy in biological systems, where a biological molecule is used in a paradoxical manner [12]. Paradoxical circuits, and the incoherent feed-forward circuit specifically, moreover, are versatile circuits. By tuning the relative strengths and timescales of the activation and repression arm, an incoherent feed-forward circuit can generate a sustained net activation, a net repression, a temporal pulse, response acceleration, band-pass filtering, and fold change detection [12].

When repression through the 11-bp is strong and activation through the WRE is weak, the transcriptional IFFL can perform fold-change detection of β-catenin. Weak activation entails loose binding of XY_1 to the WRE, or $K_1 \gg XY_1$. Weak activation can also be understood as a non-saturating WRE response, such that an increase in XY_1 can stimulate further gene expression through WRE. Strong repression entails tight binding of XY_2 to the 11-bp, or $K_2 \ll XY_2$. This has been observed for many promoter binding sites [9], though more details are needed to determine this for the XY_2 complex.

With strong repression and weak activation, Equation 4.6 becomes:

$$\frac{dZ}{dt} = \beta_3 \cdot \frac{XY_1}{K_1} \cdot \frac{K_2}{XY_2} - \alpha_3 \cdot Z \tag{4.11}$$

Next, we employ dimensional analysis to demonstrate the fold-change detection capabilities of this IFFL, based on analysis by Goentoro et al. [11]:

$$FC = X/X_0 \tag{4.12}$$

$$xy_1 = \frac{\alpha_1 \cdot XY_1}{\beta_1 \cdot X_0} \tag{4.13}$$

$$xy_2 = \frac{\alpha_2 \cdot XY_2}{\beta_2 \cdot X_0} \tag{4.14}$$

$$\tau = \alpha_2 \cdot t \tag{4.15}$$

We apply these dimensionless variables to Eqs 4.4, 4.6 & 4.11:

$$\frac{dxy_1}{d\tau} = \frac{\alpha_1}{\alpha_2} \cdot (FC - xy_1) \tag{4.16}$$

$$\frac{dxy_2}{d\tau} = FC - xy_2 \tag{4.17}$$

$$\frac{dz}{d\tau} = \frac{\alpha_3}{\alpha_2} \cdot \left(\frac{xy_1}{K_1} \cdot \frac{K_2}{xy_2} - z \right) \tag{4.18}$$

From these equations, we see that the IFFL responds only to the fold-change $FC = X/X_0$ [11]. When X is a dynamic function $X = X(t)$, the IFFL can measure various fold-change features, including the rate and integrated area of FC. When the IFFL exhibits slow dynamics relative to β-catenin accumulation, however, FC can be approximated as a step function from steady-state X_0 to a new steady-state X. This slow dynamic can come from Y_2; for instance, if it is a transcriptional target of β-catenin, the timescale is set by the transcription and translation time of Y_2.

In the case of slow IFFL dynamics, gene expression is robust to variations in parameters of the Wnt pathway. As we demonstrated in the previous section, the fold-change in β-catenin does not depend on pathway parameters:

$$FC = \frac{u}{u_0} \tag{4.19}$$

Therefore, linear signal transmission combined with fold-change detection yields robust gene expression. Interestingly, existing data confirms that Wnt-responsive gene expression is robust to perturbations, suggesting a role for linearity and fold-change detection in the Wnt pathway.

Existing data supports either fold-change detection or amplitude filtering in the Wnt pathway

Wnt signaling in early *Xenopus* blastulas activates dorsal regulators, including *siamois* and *Xnr3* [15]. Treating embryos for 5–10 min with 300 mM lithium stabilizes β-catenin [7] and dorsalizes the embryos (Figures 4.5A-C). Embryos treated with moderate doses of lithium (150 and 200 mM) also showed increased β-catenin level (Figure 4.5D, see red arrows). However, the embryos largely retained a wild-type level of *siamois* and *Xnr3* expression (Figure 4.5B, see red arrows) and developed into wild-type tailbuds (Figure 4.5A), suggesting that gene expression is robust to perturbations to the Wnt pathway. A similar lack of embryo phenotypes despite increased β-catenin level was observed with other perturbations to the Wnt pathway, including injection of *axin1* and *GBP* mRNA [10].

The TopFlash reporter, by contrast, does not exhibit buffering to moderate lithium perturbation. The TopFlash reporter is a luciferase-based reporter plasmid that is driven by multiple Wnt-responsive elements (WREs) and lacks the 11-bp element. The TopFlash reporter showed increased activity at moderate doses of lithium (Figure 4.5C, red arrows). Moreover, deletion of the 11-bp element from the *siamois* promoter caused a similar loss of buffering [15]. This suggests a critical role for the 11-bp element in imparting robustness to endogenous gene expression.

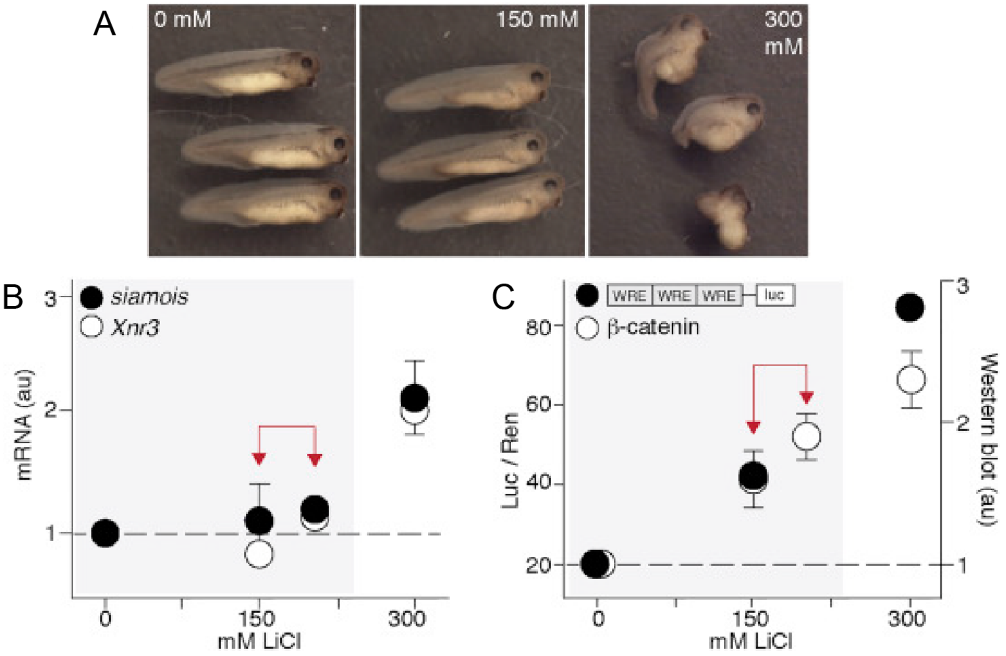


Figure 4.5: Endogenous Genes Show Regulation Not Captured by the WRE **(A)** Xenopus embryos were treated with LiCl for 5 min at the 32-cell stage, harvested at stage 10 for qRT-PCR assay, and scored 3–4 days later (shown here). **(B)** Expression of target genes, *siamois* (black circle) and *Xnr3* (white circle). Control embryos are untreated sibling embryos. Red arrows highlight how gene expression remains wild-type despite perturbations. **(C)** Black circle, luciferase/renilla signal from the TopFlash reporter injected at the four-cell stage; white circle, β-catenin level in the embryo measured using western blot. Red arrows highlight how β-catenin level and TopFlash expression change with moderate perturbations. **(A)–(C)** are reproduced from [10] with permission. Data are represented as mean ± SEM from three to five biological replicates. Error bars not visible have negligible SEM.

As we derived in the previous section, linearity and fold-change detection can give rise to robust gene expression. Here, we test where the observed gene expression robustness can indeed be explained by linear transmission in the Wnt pathway coupled by fold-change detection through β-catenin IFFL. To do this, we use the lithium perturbation data from Figure 4.5B-C. Specifically, we would like to capture β-catenin response to lithium perturbation and the resulting gene expression with or without 11-bp cis-regulation.

Lithium perturbs the Wnt pathway through inhbition of GSK3 kinase activity (the rate constant k_9 in the model, Figure 4.4A). Indeed, we find that the Wnt model recapitulates experimental findings, with a decreased k_9 leading to increased β-catenin level.

Next, we identify β-catenin IFFL parameters that are consistent with gene expression data (referred to as "consistent circuits"). TopFlash expression tracks lithium perturbation, with luciferase expression increasing concomitantly with β-catenin. Despite the increase to β-catenin level, however, endogenous *siamois* and *Xnr3* expression is buffered against moderate lithium pertubation. Therefore, consistent β-catenin IFFL circuits must show buffering against changes to k_9 and show sensitivity to changes in k_9 when the repressive 11-bp is removed (to simulate TopFlash) (Figure 4.6A).

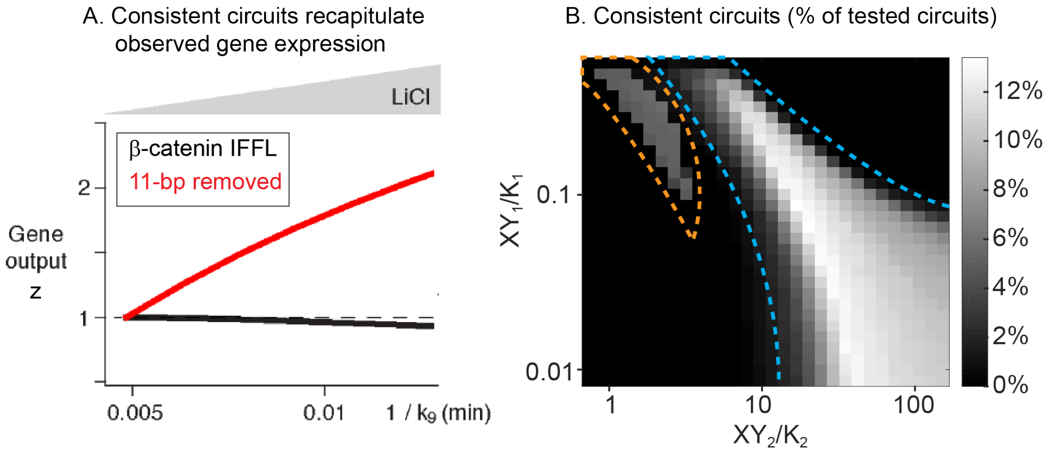


Figure 4.6: **(A)** We simulated lithium perturbation in the modified Wnt model by modifying the GSK3 phosphorylation rate k_9. An example of a consistent IFFL circuit that recapitulates lithium perturbation data is plotted. Consistent IFFL circuits show buffering against changes to k_9 and show increasing expression to decreases in k_9 when the repressive 11-bp is removed (to simulate TopFlash) **(B)** The IFFL was simulated for many values of parameters K_1, K_2, α_1, α_2, α_3. For each parameter set, the IFFL circuit was scored by its ability to recapitulate lithium perturbation data (see Methods). The intensity of each square in the K_1-K_2 plot indicates the percent of tested parameters $\{\alpha_1, \alpha_2, \alpha_3\}$ that were consistent with the data.

We tested a large range of IFFL circuits by systematically varying the IFFL parameters in Equations 4.4-4.6. The scaled Equations 4.16-4.18 indicate five relevant parameters to vary: α_1, α_2, α_3, K_1, and K_2. This parameter space was explored through millions of simulations, and the circuits consistent with experimental data are plotted in Figure 4.6B. The intensity of each square in K_1-K_2 space indicates the percent of combinations tested on parameter set $\{\alpha_1,\alpha_2,\alpha_3\}$ that yielded consistent circuits (Figure 4.6B).

From Figure 4.6B, we see that there are two regions of parameter-space that produce consistent circuits. As discussed in the previous section, fold-change detection occurs where activation is weak ($XY_1/K_1 \ll 1$) and repression is strong ($XY_2/K2 \gg 1$) - the lower right corner of Figure 4.6B. In this parameter regime, the β-catenin IFFL is invariant to scaling of β-catenin level, a sufficient condition for fold-change detection [20]. We identify the values of K_1 and K_2 for which the gene output Z is approximately scale-invariant: $Z(p \cdot X) \approx Z(X)$, for some scalar $p > 0$. Indeed, we find that this parameter regime (Figure 4.7) overlaps with the blue-outlined region in Figure 4.6B. We also find that the smaller, orange-outlined region in Figure 4.6B does not satisfy the requirements for fold-change detection.

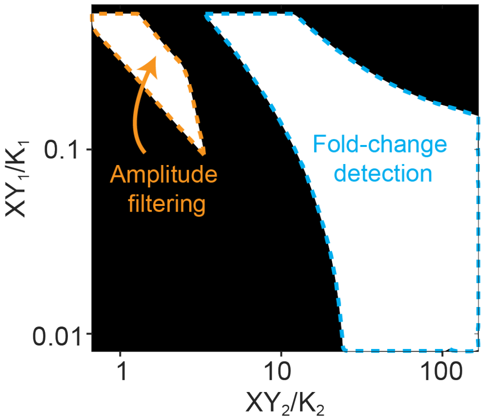


Figure 4.7: The β-catenin IFFL behaves either as a fold-change detector or amplitude filter. We identify the behavior within the two regions of consistent circuits from Figure X. First, we show the larger region (in blue) exhibits approximate scale invariance, with $Z(p \cdot X) \approx Z(X)$ for $p > 0$. Therefore, this region of circuits satisfies the conditions for fold-change detection from ref. [20]. The smaller region (in orange) does not meet these criteria. Instead, we find that in this region $X \approx \sqrt{K_1 \cdot K_2}$, indicating that the circuits in this region perform amplitude filtering as in ref. [14].

We find that the circuits in the smaller, orange-outlined region of Figure 4.6B behave as amplitude filters. An amplitude filter responds maximally to inputs within a particular range and produces no response for inputs outside this range. Specifically, the β-catenin IFFL performs amplitude filtering when $XY_1 \approx XY_2 \approx X$, with maximal response occuring when $X \approx \sqrt{K_1 \cdot K_2}$. Indeed, we find that all circuits in the smaller region of Figure 4.6B satisfy the requirements for amplitude filtering (Figure 4.7).

Discussion

Here, we show how linearity can facilitate reliable fold-change output in the face of parameter variations. For a linear input-output response, parameter variation affects only the gain of the response. In turn, fold-change output is computed as the ratio of two points on this linear input-output response and is invariant to changes in the gain. Importantly, there is evidence that all four pathways we examined in Chapter 2 produce reliable fold-change output. Therefore, fold-change response is a benefit of linearity and may act as a selective pressure on the evolution of input-output response of signaling pathways. The IFFL motif may be a pervasive mechanism for fold-change detection, as it has also been proposed in the NF-κB pathway [17].

In this chapter, we also find that the incoherent feed-forward loop formed by WREs and the 11-bp element can behave either as fold-change detector or as an amplitude filter. There are significantly more parameter combinations that support fold-change detection than amplitude filtering, suggesting that it may be a more plausible mechanism in the Wnt pathway. However, further experiments are needed to distinguish between the two. We propose two approaches to distinguishing between these behaviors (Figure S4.2): first, the fold-change detector leads to perfect adaptation of gene expression, while the amplitude filter does not. This can be tested by examining the basal *siamois*/*Xnr3* gene expression in *Xenopus* embryos (before β-catenin accumulation) in the presence or absence of perturbation. Second, fold-change detectors compute signal relative to basal, which means that the maximal gene output is inversely proportional to basal Wnt pathway activity. Counter-intuitively, an increase to the basal pathway activation decreases the maximal gene expression for fold-change detection (Figure X). In contrast, basal pathway activation would not affect the maximal response of the amplitude filter. An increase in basal Wnt pathway activity can be implemented, for instance, through embryo injection of low levels of *Wnt* mRNA.

What are the advantages of fold-change detection relative to the amplitude filtering, in the context of the Wnt pathway? Both behaviors can produce the observed gene expression buffering to GSK3 and Axin perturbation. That is, both fold-change detection and amplitude filtering can reject moderate disturbances to Wnt signal transmission. However, amplitude filtering accomplishes this in exchange for a loss of sensitivity to the Wnt signal: while β-catenin varies continuously with Wnt signal, the downstream gene output is binary (either on or off). Fold-change detection, on the other hand, can track changes in β-catenin while simultaneously rejecting prolonged disturbances (through adaptation). Thus, fold-change detection at the transcriptional level can overcome disturbances to Wnt signal transmission (e.g. chemical perturbation, variations in maternal mRNA inheritence, temperature), while responding sensitively to changes in external Wnt ligand.

4.3 Materials and Methods

To determine whether an IFFL circuit was consistent, we tested it against the following criteria:

1. Gene output is responsive to Wnt stimulation: IFFL gene output increased at least two-fold over basal.

2. For 1.5 and two-fold perturbations to GSK3 kinase activity, IFFL gene output changes by less than 5%.

3. For 1.5 and two-fold perturbations to GSK3 kinase activity, WRE-only gene output tracks β-catenin to within 50%. That is, if β-catenin increases by two-fold due to perturbation, gene output increases by at least 1.5-fold.

Simulations were performed in Matlab.

4.4 Supporting Figures

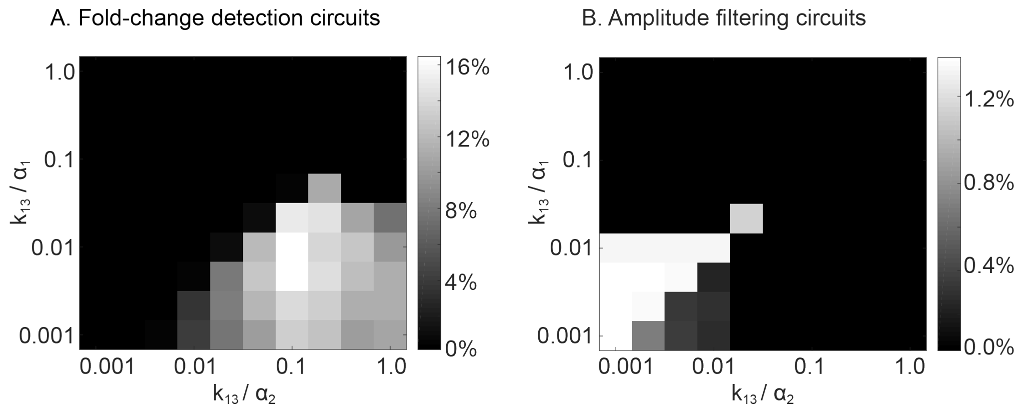


Figure S4.1: **(A)** Timescales α_1 and α_2 for activator XY_1 and repressor XY_2, respectively, that produce consistent, fold-change detection circuits. The timescales are normalized to destruction complex independent degradation, k_{13}. A value of one provides an upper bound for protein stability, while lower values indicate faster kinetics, such as protein-protein binding. The plot intensity indicates the percentage of tested parameters $\{K_1, K_2\}$ that produced consistent circuits. **(B)** Timescales α_1 and α_2 for activator XY_1 and repressor XY_2, respectively, that produce consistent, amplitude filtering circuits. The data is plotted as in **(A)**.

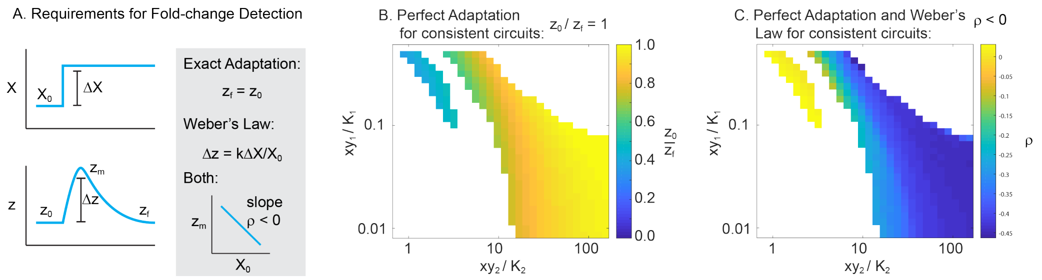


Figure S4.2: **(A)** Fold-change detection entails exact adaptation and Weber's Law. A consequence of both of these conditions is that an increase in X_0 (all else equal) yields a decrease in maximal gene output z_m. **(B)** The adaptation of all consistent circuits was calculated, with $y_0/y_f = 1$ indicating perfect adaptation. **(C)** The relationship between maximal gene output z_m and basal β-catenin X_0 was calculated, with a negative relationship (slope $\rho < 0$) required for fold-change detection.

BIBLIOGRAPHY

[1] Uri Alon. *An introduction to systems biology: design principles of biological circuits*. CRC press, 2006. ISBN 1584886420.

[2] Steven S. Andrews, William J. Peria, C. Yu Richard, Alejandro Colman-Lerner, and Roger Brent. Push-pull and feedback mechanisms can align signaling system outputs with inputs. *Cell systems*, 3(5):444–455. e2, 2016. ISSN 2405-4712.

[3] Verena Becker, Marcel Schilling, Julie Bachmann, Ute Baumann, Andreas Raue, Thomas Maiwald, Jens Timmer, and Ursula Klingmüller. Covering a broad dynamic range: information processing at the erythropoietin receptor. *Science*, 328(5984):1404–1408, 2010. ISSN 0036-8075.

[4] Lacramioara Bintu, Nicolas E Buchler, Hernan G Garcia, Ulrich Gerland, Terence Hwa, Jané Kondev, and Rob Phillips. Transcriptional regulation by the numbers: models. *Current opinion in genetics & development*, 15(2): 116–124, 2005.

[5] Ken M Cadigan and Marian L Waterman. Tcf/lefs and wnt signaling in the nucleus. *Cold Spring Harbor perspectives in biology*, page a007906, 2012.

[6] Cellina Cohen-Saidon, Ariel A. Cohen, Alex Sigal, Yuvalal Liron, and Uri Alon. Dynamics and variability of erk2 response to egf in individual living cells. *Molecular cell*, 36(5):885–893, 2009. ISSN 1097-2765.

[7] Domitilla Del Vecchio, Aaron J Dy, and Yili Qian. Control theory meets synthetic biology. *Journal of The Royal Society Interface*, 13(120):20160380, 2016. ISSN 1742-5689.

[8] Christopher L. Frick, Clare Yarka, Harry Nunns, and Lea Goentoro. Sensing relative signal in the tgf-β/smad pathway. *Proceedings of the National Academy of Sciences*, page 201611428, 2017. ISSN 0027-8424.

[9] Marcel Geertz and Sebastian J Maerkl. Experimental strategies for studying transcription factor–dna binding specificities. *Briefings in functional genomics*, 9(5-6):362–373, 2010.

[10] Lea Goentoro and Marc W. Kirschner. Evidence that fold-change, and not absolute level, of β-catenin dictates wnt signaling. *Molecular cell*, 36(5): 872–884, 2009. ISSN 1097-2765.

[11] Lea Goentoro, Oren Shoval, Marc W. Kirschner, and Uri Alon. The incoherent feedforward loop can provide fold-change detection in gene regulation. *Molecular cell*, 36(5):894–899, 2009. ISSN 1097-2765.

[12] Yuval Hart and Uri Alon. The utility of paradoxical components in biological circuits. *Molecular cell*, 49(2):213–221, 2013.

[13] Ana R. Hernández, Allon M. Klein, and Marc W. Kirschner. Kinetic responses of β-catenin specify the sites of wnt control. *Science*, 338(6112):1337–1340, 2012. ISSN 0036-8075.

[14] Shai Kaplan, Anat Bren, Erez Dekel, and Uri Alon. The incoherent feed-forward loop can generate non-monotonic input functions for genes. *Molecular systems biology*, 4(1):203, 2008.

[15] Kibeom Kim, Jaehyoung Cho, Thomas S Hilzinger, Harry Nunns, Andrew Liu, Bryan E Ryba, and Lea Goentoro. Two-element transcriptional regulation in the canonical wnt pathway. *Current Biology*, 27(15):2357–2364, 2017.

[16] Ethan Lee, Adrian Salic, Roland Krüger, Reinhart Heinrich, and Marc W. Kirschner. The roles of apc and axin derived from experimental and theoretical analysis of the wnt pathway. *PLoS Biol*, 1(1):e10, 2003. ISSN 1545-7885.

[17] Robin E. C. Lee, Sarah R. Walker, Kate Savery, David A. Frank, and Suzanne Gaudet. Fold change of nuclear nf-κb determines tnf-induced transcription in single cells. *Molecular cell*, 53(6):867–879, 2014. ISSN 1097-2765.

[18] Dmitry Nevozhay, Rhys M Adams, Kevin F Murphy, Krešimir Josić, and Gábor Balázsi. Negative autoregulation linearizes the dose–response and suppresses the heterogeneity of gene expression. *Proceedings of the National Academy of Sciences*, 106(13):5123–5128, 2009.

[19] Diego A. Oyarzún, Jo L. Bramhall, Fernando López-Caamal, Frances M. Richards, Duncan I. Jodrell, and Ben-Fillippo Krippendorff. The egfr demonstrates linear signal transmission. *Integrative Biology*, 6(8):736–742, 2014.

[20] Oren Shoval, Lea Goentoro, Yuval Hart, Avi Mayo, Eduardo Sontag, and Uri Alon. Fold-change detection and scalar symmetry of sensory input fields. *Proceedings of the National Academy of Sciences*, 107(36):15995–16000, 2010. ISSN 0027-8424.

[21] Steven Smith. *Digital signal processing: a practical guide for engineers and scientists*. Elsevier, 2013.

[22] Kevin Thurley, Stephen C. Tovey, Gregor Moenke, Victoria L. Prince, Abha Meena, Andrew P. Thomas, Alexander Skupin, Colin W. Taylor, and Martin Falcke. Reliable encoding of stimulus intensities within random sequences of intracellular ca2+ spikes. *Science signaling*, 7(331):ra59, 2014.

[23] Kevin Thurley, Daniel Gerecht, Elfriede Friedmann, and Thomas Höfer. Three-dimensional gradients of cytokine signaling between t cells. *PLOS Comput Biol*, 11(4):e1004206, 2015. ISSN 1553-7358.

[24] C. Richard Yu, C. Gustavo Pesce, Alejandro Colman-Lerner, Larry Lok, David Pincus, Eduard Serra, Mark Holl, Kirsten Benjamin, Andrew Gordon, and Roger Brent. Negative feedback that improves information transmission in yeast signalling. *Nature*, 456(7223):755–761, 2008. ISSN 0028-0836.

Chapter 5

CONCLUSIONS

This work identifies linear signal transmission as an overarching principle of the Wnt, ERK, and Tgfβ signaling pathways. Through mathematical analysis and experiments, I demonstrated that these three pathways converge upon linear input-output response despite significant structural differences between the pathways. Linearity is a desired property in engineering where it facilitates fidelity and superposition in signal transmission. This suggests that linear signal transmission may have a similar utility in cell signaling, creating a selective pressure for linearity to arise within many distinct signaling pathways.

First, in Chapter 2, I used mathematical analysis of pathway models to demonstrate linear input-output response, as well as the mechanisms by which linearity arises. I found that linearity is produced by distinct mechanisms in each pathway. Linearity in the Wnt pathway requires a futile cycle of β-catenin, which is maintained by rapid activity of the β-catenin destruction complex. In the ERK pathway, linearity is generated by an ultrasensitive kinase cascade coupled to negative feedback. In the Tgfβ pathway, linearity arises from continual nucleocytoplasmic shuttling of Smad, which maintains a Tgfβ-responsive pool of Smad in the cytoplasm. This mathematical analysis provides insight into how linear response is achieved in different signaling pathways.

In Chapter 3, I experimentally validated these findings in the Wnt and ERK pathways using quantitative Western blots and modulated the input-output response for both pathways. I found that in the Wnt pathway, β-catenin is linear with respect to Wnt co-receptor LRP activity up until receptor saturation. Further, by inhibiting β-catenin kinase GSK3 using CHIR99021, I demonstrated that the range of linearity can be reduced experimentally. Next, I showed a linear relationship between doubly-phosphorylated ERK and the concentration of EGF ligand, up until saturation of ERK. Expression of a feedback-resistant Raf-1 protein, however, weakened feedback in the pathway and yielded a nonlinear input-output response over the same range of EGF dose. These findings provide quantitative validation of the mathematical analysis from Chapter 2.

Finally, in Chapter 4, I used modeling to demonstrate how linearity can produce fold-change response in the Wnt pathway, and how a novel cis-regulatory element could detect fold-changes in β-catenin. First, I showed that linear input-output response is sufficient to produce fold-change response in the Wnt pathway, such that fold-change in β-catenin is robust to variations in pathway parameters. Next, I demonstrated how a novel cis-regulatory element in the promoter of many Wnt-responsive genes forms an incoherent feed-forward loop (IFFL) regulating gene expression. This β-catenin IFFL can detect fold-changes in β-catenin, and, when combined with linear signal transmission, yields robust gene expression.

Our work suggests two benefits for linearity in cell signaling. First, linearity enables the superposition of multiple signals within a single pathway. Superposition is a useful feature of engineered communications systems, where it enables the shared use of common communication channels. Similarly, linearity may mitigate the communication bottleneck created by the limited number of developmental signaling pathways through signal multiplexing. Indeed, work on dynamic signal encoding in the ERK, Notch, and p53 pathways provides evidence for this strategy in cell signaling [1, 6, 7]. Thus, superposition may not only have a role in engineered systems, but also in biological signaling systems.

Our work also suggests that linearity in cell signaling pathways imparts robustness to signaling. I demonstrated in Chapter 4 that linearity underlies fold-change response, an emerging property of signaling pathways in which signal is transmitted relative to basal pathway activity [2–5, 8, 9]. Specifically, the fold-change computed from a linear input-output system is insensitive to parameters. Interestingly, this strategy suggests that the amplification gain does not need to be tightly regulated. Instead, signaling fidelity is guaranteed as long as transmission is linear. Thus, this thesis demonstrates that linearity is a widespread property of cell signaling, and may increase the fidelity and rate of information transfer through signaling pathways.

BIBLIOGRAPHY

[1] Eric Batchelor, Alexander Loewer, Caroline Mock, and Galit Lahav. Stimulus-dependent dynamics of p53 in single cells. *Molecular systems biology*, 7(1): 488, 2011.

[2] Cellina Cohen-Saidon, Ariel A. Cohen, Alex Sigal, Yuvalal Liron, and Uri Alon. Dynamics and variability of erk2 response to egf in individual living cells. *Molecular cell*, 36(5):885–893, 2009. ISSN 1097-2765.

[3] Christopher L. Frick, Clare Yarka, Harry Nunns, and Lea Goentoro. Sensing relative signal in the tgf-β/smad pathway. *Proceedings of the National Academy of Sciences*, page 201611428, 2017. ISSN 0027-8424.

[4] Lea Goentoro and Marc W. Kirschner. Evidence that fold-change, and not absolute level, of β-catenin dictates wnt signaling. *Molecular cell*, 36(5):872–884, 2009. ISSN 1097-2765.

[5] Robin E. C. Lee, Sarah R. Walker, Kate Savery, David A. Frank, and Suzanne Gaudet. Fold change of nuclear nf-κb determines tnf-induced transcription in single cells. *Molecular cell*, 53(6):867–879, 2014. ISSN 1097-2765.

[6] Nagarajan Nandagopal, Leah A Santat, Lauren LeBon, David Sprinzak, Marianne E Bronner, and Michael B Elowitz. Dynamic ligand discrimination in the notch signaling pathway. *Cell*, 172(4):869–880, 2018.

[7] Silvia DM Santos, Peter J Verveer, and Philippe IH Bastiaens. Growth factor-induced mapk network topology shapes erk response determining pc-12 cell fate. *Nature cell biology*, 9(3):324, 2007.

[8] Kevin Thurley, Stephen C. Tovey, Gregor Moenke, Victoria L. Prince, Abha Meena, Andrew P. Thomas, Alexander Skupin, Colin W. Taylor, and Martin Falcke. Reliable encoding of stimulus intensities within random sequences of intracellular ca2+ spikes. *Science signaling*, 7(331):ra59, 2014.

[9] Kevin Thurley, Daniel Gerecht, Elfriede Friedmann, and Thomas Höfer. Three-dimensional gradients of cytokine signaling between t cells. *PLOS Comput Biol*, 11(4):e1004206, 2015. ISSN 1553-7358.