RESEARCH ARTICLE

Information transmission from NFkB signaling dynamics to gene expression

Alok Maity[1], Roy Wollman[1,2]*

1 Institute for Quantitative and Computational Biosciences, University of California, Los Angeles, California, United States of America, **2** Departments of Integrative Biology and Physiology and Chemistry and Biochemistry, University of California UCLA, California, United States of America

* rwollman@ucla.edu

Abstract

The dynamic signal encoding paradigm suggests that information flows from the extracellular environment into specific signaling patterns (encoding) that are then read by downstream effectors to control cellular behavior. Previous work empirically quantified the information content of dynamic signaling patterns. However, whether this information can be faithfully transmitted to the gene expression level is unclear. Here we used NFkB signaling as a model to understand the accuracy of information transmission from signaling dynamics into gene expression. Using a detailed mathematical model, we simulated realistic NFkB signaling patterns with different degrees of variability. The NFkB patterns were used as an input to a simple gene expression model. Analysis of information transmission between ligand and NFkB and ligand and gene expression allows us to determine information loss in transmission between receptors to dynamic signaling patterns and between signaling dynamics to gene expression. Information loss could occur due to biochemical noise or due to a lack of specificity. We found that noise-free gene expression has very little information loss suggesting that gene expression can preserve specificity in NFkB patterns. As expected, the addition of noise to the gene expression model results in information loss. Interestingly, this effect can be mitigated by a specific choice of parameters that can substantially reduce information loss due to biochemical noise during gene expression. Overall our results show that the cellular capacity for information transmission from dynamic signaling patterns to gene expression can be high enough to preserve ligand specificity and thereby the accuracy of cellular response to environmental cues.

Author summary

The fidelity of signal transduction depends on the accurate encoding of ligand information in specific signaling patterns and the reliable transmission of these patterns by downstream gene expression machinery. We present an analysis of the accuracy of information transmission from signaling dynamics into gene expression in the case of the transcription factor NFkB. We show that noiseless gene expression can preserve ligand identity with

minimal information loss. The addition of noise to gene expression model results in information loss, an effect that can be largely mitigated by choice of parameter values.

Introduction

The ability of cells to respond to environmental changes is key to their function. A ligand binding to a receptor initiates a cascade of biochemical transformations of cellular kinases, phosphatase, and other enzymes that connect the receptor to downstream effectors, often to change gene expression patterns [1]. Historically, these cascades of events were divided into distinct pathways. The pathway paradigm was appealing because it was easy to understand how information about ligand identity and abundance is preserved. However, as more and more signaling interactions were discovered, it became clear that the linear signaling pathway paradigm is insufficient in describing the complexity of how information propagates between receptors and effectors. The complexity of signaling networks with a large degree of crosstalk and feedback shifted the paradigm from pathway to network-centric view [2]. However, unlike an isolated signaling cascade, the network view poses a challenge. If the mapping between ligands and signaling nodes is many to many, how does specificity, i.e. information about ligand identify and concentration, is preserved?

A plausible solution for the specificity challenge in signaling networks is based on the concept of dynamic signal encoding [3–8]. Information about extracellular events undergoes multiple transformations. Initially, information on ligand identity and abundance is transformed, i.e. encoded, into a specific signaling activity pattern. Subsequently, downstream effectors such as transcription factors transform the specific dynamic signaling patterns to a specific cellular response [9–12]. The dynamic signal encoding view is useful since it explains how information can be preserved despite the complex many to many relationships between receptors and signaling nodes [13]. However, given the current understanding of transcription factor activity, it is unclear whether their dynamics can fully map distinct ligands to different patterns of gene expression. In other words, is the information transmission capacity between signaling dynamics to gene expression sufficiently large to transmit information encoded by these dynamic signaling patterns? It is possible that not all the information that exists within dynamic signaling patterns will be accessible to downstream gene expression machinery [14–16]. Information transmission into gene expression could be limited by specific biochemical constraints of gene expression machinery, i.e. are a property of the gene expression model [17–22] or simply due to additional layer of biochemical noise during the process of gene expression itself [15,23,24].

Information theory can be used to assess the quality of any input/output relationships between two random variables and was used extensively to quantify and understand information transmission in neurons [25–27] and developmental biology [28–31]. As the fundamental function of signaling networks is to reliably transmit specific information about ligand concentration to downstream effectors, information theoretical tools have been utilized as a way to assess the performance of a network [19,20,32–40]. Existing biochemical variability that occurs at multiple timescales can have an adverse effect on the quality of information transmission [38,41–46]. Using information theoretical tools, one can probe the operational quality of a signaling network in a quantitative manner by measuring the information transmission capacity of a network through a series of input/output measurements. Mutual information analysis can quantify the degree of overlap between cellular responses to multiple distinct inputs and thereby it is a good descriptor of the accuracy of information transmission through

a signaling network. These tools have been applied to signaling networks showing that indeed information loss occurs due to biochemical noise during the encoding step [16,47–53].

Within the dynamic signal encoding paradigm, information transmission depends on the information that exists within the dynamic signal as well as the ability to transmit this information into specific gene expression patterns. If gene expression is noisy and inaccurate, a high-quality encoding is of little use to the cell. Therefore, the question of how accurate is gene expression in capturing information from dynamic signals is paramount to our understanding of the overall performance characteristics of signaling networks. Previous work addressed this question using experimental and using theoretical approaches. Experimental measurement of the accuracy of mapping between signaling dynamics to single-cell gene expression is technically challenging. Work by Hansen et al [16] analyzed the degree of gene expression accuracy with regards to oscillatory signal and amplitude. While useful, it is unknown if cells rely on amplitude and frequency as the key features. In a more physiological setting, Lane et al measured the correspondence in a single cell between NFkB dynamics and resulting gene expression [54]. They were able to demonstrate that indeed the overall patterns of signaling dynamics result in distinct expression patterns. However, a high level of measurement error and a small sample size preclude an analysis of the reliability of information transmission accuracy from signaling dynamics into gene expression processes. Theoretical work on this question was more extensive [19,55–58]. However past work based the analysis on specific dynamic features (e.g. frequency or amplitude) or used simplified input signals that are unrealistic. Therefore, there is a gap in our understanding of the degree of accuracy of information transmission from ligand to gene expression patterns for realistic signaling dynamics patterns.

Here we address this gap by analyzing the information transmission accuracy of a simple gene expression model. We utilize NFkB signaling, as it is a system where dynamic signal encoding has been analyzed in depth [47–49,59–62]. We first generate simulated encoding data for multiple ligands that activate NFkB and ask how accurate is the mapping of this dynamic into distinct gene expression patterns. We show that without noise, gene expression accuracy is close to its maximal possible value for a large range of parameter values. However, when noise in gene expression is considered performance decreases. Still, close to maximal information transmission can still be achieved by a specific choice of gene expression model parameters. Overall our finding demonstrates that information transmission from signaling dynamic into gene expression can be achieved to allow cells to accurately respond to environmental changes.

Results

We first constructed mathematical models that represent the dynamic signal encoding and information transmission into gene expression steps in NFkB signaling (Fig 1). The first model (Fig 1A) captures the encoding step and the transformation of ligand identify to the dynamics of NFkB. The second model (Fig 1B) focuses on gene expression based on NFkB dynamics. As the goal of the encoding model was to capture the specific aspects of NFkB dynamics we opted to use a highly detailed model with 95 reactions of 48 reactants and 127 parameters [63]. The model connects five different receptors: TNFR, TLR4, TLR3, TLR9, and TLR1/2 through the key downstream pathways of MyD88 and TRIF to the IKK modules that controls NFkB dynamics [64]. The model generates realistic simulated time series up to 8 hours that match experimental data for five different ligands [65]. Given the scope of the model and the fact that includes multiple inputs to the same core network we focused on the question of ligand specificity and tested one ligand concentration for each condition (i.e., TNF = 10ng/ml, LPS = 10ng/ml, PolyIC = 30µg/ml, CPG = 1µM, Pam3CSK = 300ng/ml). The
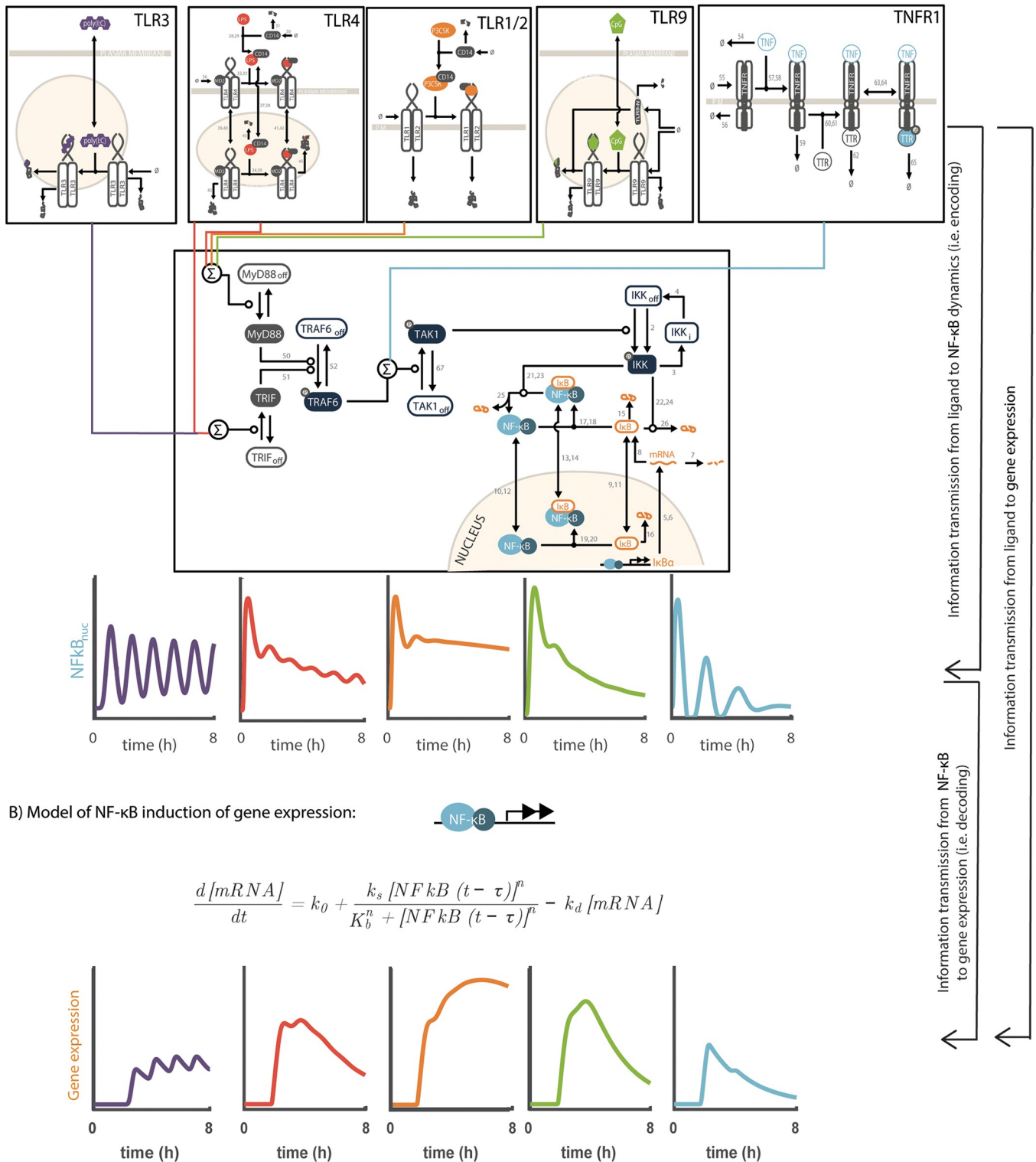
Fig 1. Models of dynamic signal encoding and gene expression in the NFkB network. A. The encoding model of NFkB signaling network. The model includes five distinct receptor modules (TLR3, TLR4, TLR1/2, TLR8, TNFR1) that feed at three distinct points (MyD88, TRIF, and TAK1) into the core IKK/NFkB module. The dynamic interaction in network results from the mapping between each receptor module and dynamic activation profiles of the transcription factor NFkB. **B**. The gene expression model contains a simplified a single ordinary differential equation that uses the nuclear concentration of

NFkB over time as an input and produces the gene expression pattern as an output. The model effectively maps the dynamic of NFkB into dynamic gene expression profile.

ligand concentrations were chosen to be close to the corresponding receptor Kd [63]. The construction of these two models is based on a functional separation of encoding ligand identity into NFkB dynamics and the transmission of this information into gene expression patterns. This functional separation is not the only one as the NFkB network by itself includes gene expression. However, this separation is useful in addressing the question of information transmission from NFkB dynamics into gene expression. Furthermore, this separation makes needed simplification that only NFkB dynamics matter while it is well known that these ligands not only regulate differential NFkB profiles but also activate additional transcription factors such as interferon regulatory factors (IRFs), kinases of the c-Jun N-terminal kinase (JNK) and MAPK/ERK families [66]. The key benefit of using a mathematical model to capture NFkB dynamic profile is that it allows us to control the degree of encoding noise, a manipulation that is technically impossible to perform experimentally. Furthermore, the analysis avoids the need to deal with experimental and technical measurement errors. The second model (Fig 1B) captures the gene expression step. Unlike the encoding model that aims at a highly realistic details description of the underlying system, the gene expression model is very simple with a single differential equation and six parameters. We chose a simple model to allow us to capture the essence of the gene expression step without unnecessarily increasing the number of parameters. Using the two models in tandem we can simulate how cells respond to extracellular stimuli through both NFkB signaling dynamics and resulting gene expression at multiple degrees of biological extrinsic noise, as was shown to be important for information capacity of signaling dynamics [44,67].

To analyze the effect of cellular heterogeneity on cellular information transmission we estimated the effect of biochemical variability on the accuracy of signal transduction using an information theoretic approach. The dominant source of variability in many signaling networks [41,47,68] and in NFkB specifically [69] is differences between cells in their underlying cell state (e.g. protein concentration, organelle structure, etc) between cells. To capture this effect, we repeated the simulation of the encoding model 1000 times per condition with parameter drawn from a log-normal distribution centered around the reference value with different degrees of the coefficient of variation. Our model makes the simplifying assumption that stochastic effects during dynamic signal encoding (i.e. intrinsic noise) are of smaller magnitude. We tested a range of variability magnitudes from 10–35% and at each condition estimated the mutual information between ligand identity and NFkB dynamics (Fig 2A, cyan line). As expected, the accuracy of the encoding step depended on the degree of biochemical variability. At low values (10%) there was very little information loss compared to the maximum possible value of $\log_2(5) \sim 2.3$. As expected, as the magnitude of variability increased the signaling accuracy decreased [44].

Next, we used the simulated trajectories as input to the gene expression model and asked how much additional information loss will occur during information transmission into gene expression. To separate the effect of biochemical variability during signaling and gene expression we first tested a gene expression model without biochemical variability. To make the analysis independent of the specific parameter values of encoding we sampled 1000 different parameter values of the 127-dimension parameter vector. We found that the noise-free gene expression model is able to capture most of the information that exists within NFkB dynamics (Fig 2A and 2B and Table B in the S1 Text). This indicates that the five tested ligands generate sufficiently distinct NFkB activation profiles such that a simple model of gene expression is
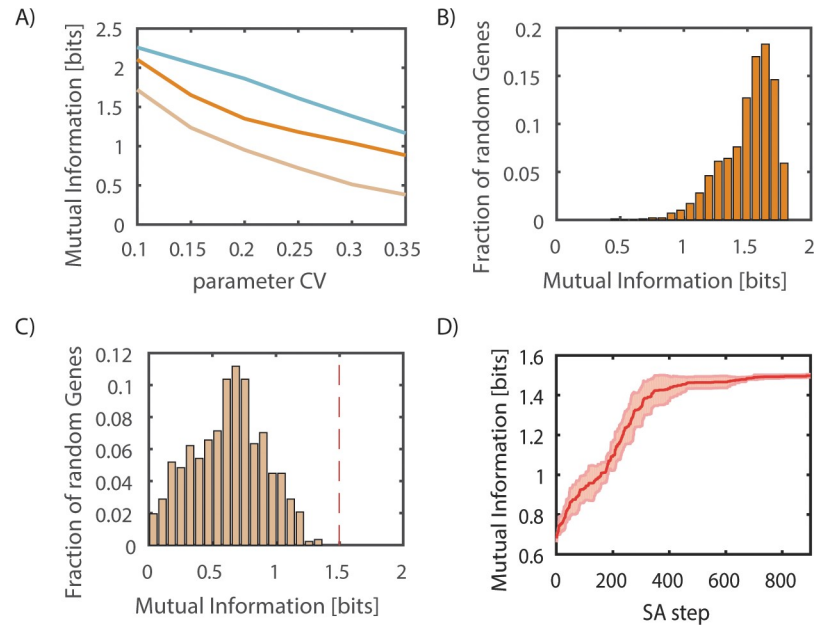
Fig 2. Information loss in dynamic signal encoding and gene expression. A. Mutual information as a function of parameter variability that exists in the encoded in NFkB dynamics (cyan), and noiseless (dark orange) and noisy (light orange) information transmission of NFkB dynamics into simple gene expression model. Values are the mean of 1000 samples. **B** and **C** distribution of mutual information values in noiseless (**B**) and noisy (**C**) information transmission of NFkB dynamics into a simple gene expression model. The distribution comes from the inference of mutual information after sampling different model parameters. In each case, 1000 parameter values are sampled from a log-normal distribution centered around reference values with a CV of 25%. The mutual information achieved after optimization of the parameter values of the gene expression model, the mean of panel B MI = 1.5, is shown as a red dashed line in panel C. **D.** Optimization of parameter values using heuristic simulated annealing show convergence of mutual information values. The red line shows the average and shading the standard deviation from 4 repeated optimization runs.

sufficient to generate distinct gene expression profiles that provide sufficient information capture the ligand identity to a similar degree that it exists in NFkB dynamics.

The analysis above takes into account noise within the encoding step but ignores the existence of noise within the gene expression step. As it is likely that biochemical variability plays a role in gene expression we tested the overall information transmission fidelity in the presence of noise in gene expression. The addition of biochemical variability into the gene expression step resulted in additional loss of information. This was true for large ranges of parameter values for the gene expression model (Fig 2C).

To maximize signaling fidelity, information loss should be minimized. We tested whether the information loss we observe during the gene expression step is an unavoidable property of the system, or whether it is possible to achieve higher fidelity information transmission with optimized parameters. We performed heuristic optimization of the gene expression model parameters with the objective function of maximizing the mutual information between ligands and gene expression dynamics (Fig 2D). In the figure, line and shading represent mean and standard deviation, respectively, generated from 4 repeated optimization runs of gene expression parameters (S1 Text). Interestingly, with the right parameters, there was little loss of information during gene expression step (Fig 2C red dashed line) despite the existence of biochemical noise in the gene expression model. Repeating the optimization multiple times identified similar parameter values. Overall these results indicate that careful choice of the population average gene expression parameters can mitigate the effect of biochemical
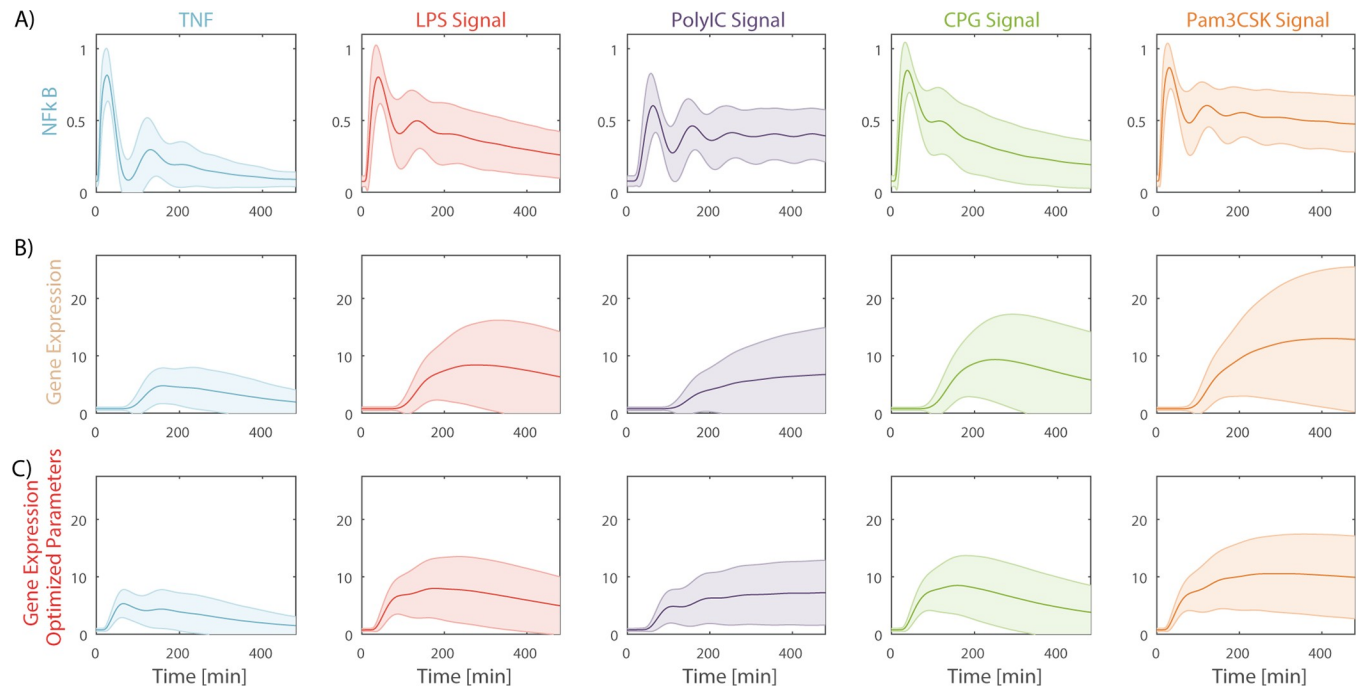
Fig 3. Simulation of NfkB dynamics and resulting gene expression. Average (center line) and standard deviation (shading) of NFkB dynamic (**A**), noisy gene expression model (**B**) and noisy gene expression model centered around parameter values that allowed maximal information transmission (**C**). Each condition was simulated 1000 times with parameters of both signaling and gene expression models sampled from a log-normal distribution with a CV of 25%.

variability around these reference values (Table A in the S1 Text). These results suggest that cells can potentially adopt an information transmission strategy to maximize the information extraction from the signal encoded by NFkB dynamics.

We next wanted to better understand the identified strategy that maximizes information transmission by answering two questions related to the identified solutions: 1. What changes in gene expression patterns cause it to increase its information about ligand identity and 2. What model parameters are changed and do these changes contribute to the increase in information transmission accuracy?

To assess what changed about the system after optimization we examined the dynamic trajectories of NFkB and resulting gene expression patterns (Fig 3). A strength of mutual information analysis is that it incorporates both differences in the overall separation between average responses to ligands and the variability within a response to a single ligand into the same framework. Analysis of the gene expression pattern before and after optimization indicates that the major change in gene response after optimization is a reduction in response variability per each ligand (Fig 4D) and not from an increased separation of the typical response to all ligands (Fig 4E). Interesting, mutual information was insensitive to up to 2-fold changes in dose of LPS or TNF suggesting that the ability to discriminate between ligands is not dose specific (Figure C and Figure D in S1 Text).

Next, we examined the contribution of each of the identified parameters to the increase in information transmission accuracy. Of the system's six parameters, only three changed their value in a consistent manner (Fig 4A): mRNA degradation rate (k_d), the cooperativity of the binding reaction (n), and the time lag between increased NFkB in the nucleus to changes in expression (τ). The bar graph values of each parameter, mean and standard deviation were calculated from the optimized parameter distribution that is generated from 4 repeated SA runs.
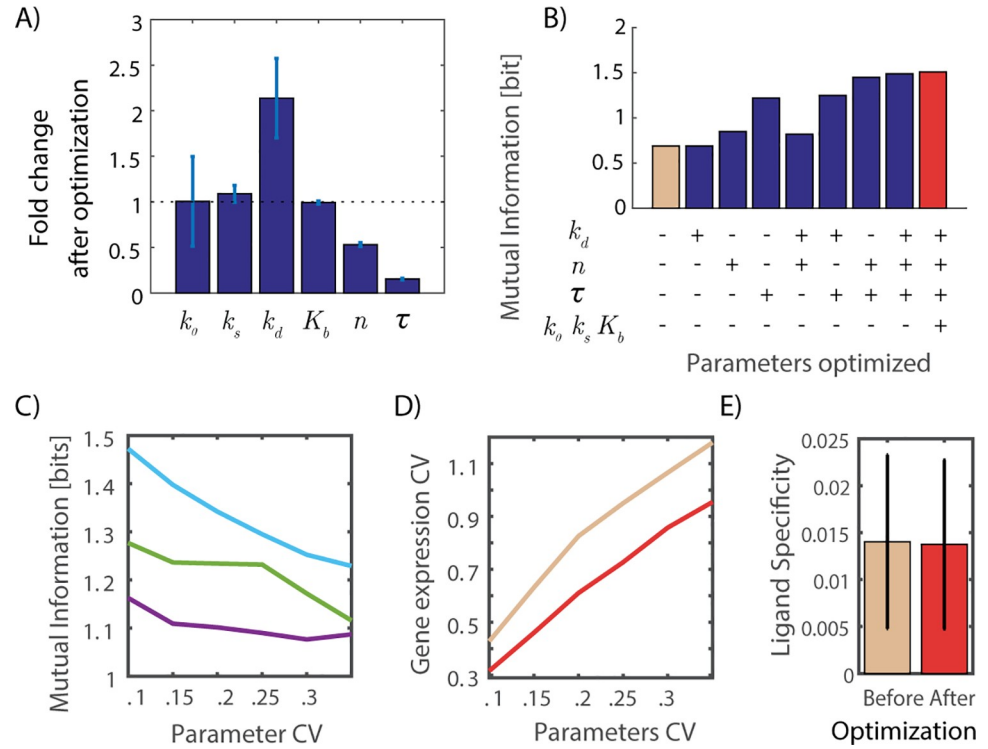
Fig 4. Effects of mutual information optimization on gene expression. A. Changes in the six parameters of the gene expression model as a result of optimization to increase mutual information. Error bars show standard deviation from 4 repeated optimization runs. **B.** Mutual information with all parameters used at nominal values (light orange), optimized values (red), or a mixture of nominal and optimized values (blue bars). **C.** Mutual information as a function of the amount of variability in a single model parameter shown for mRNA degradation rate k_d (purple), hill coefficient n (green) and time delay τ (cyan). All parameters considered at nominal values. **D.** Change in gene expression variability with an increase in signaling and gene expression models parameter variability for nominal (light orange) and parameter optimized (red) gene expression models. **E** Ligand specificity was calculated as the mean square error between the average responses to the five ligands shown before and after parameter optimization to maximize mutual information.

Figure E in the S1 Text also depicted the distribution of optimized gene expression parameters. However, the fact that a parameter changed through our optimization does not necessarily guarantee that this change is meaningful and is responsible for the increase in mutual information. To assess the specific contribution of these three parameters we score the information transmission accuracy for all possible combinations of these parameters for a specific value of variability (25% parameter CV). In each test, all other parameters have their original reference value before optimization. We saw that the parameter that contributed most to the reduced variability is the time delay parameter tau (Fig 4B). This was true for a wide range of degree of biochemical variability (Fig 4C). Given the dynamic nature of NFkB, a reduction in the value (and variance) of the time delay unifies the population response and can explain the reduction in gene expression variability we observed (Fig 4D). Interestingly, the lower value (14.22 minutes) matches the experimental estimate of this delay for genes that are part of the NFkB network such as A20 and IkB [62].

Discussion

Metazoan signaling architecture is unique with a high degree of crosstalk through the network [70]. The benefits of this architecture are that it is more plastic and allows multicellular

organisms to have specialized cells type that responds differently to the same environment without exponentially growing number of pathways [69]. However, this benefit comes at a cost. Cells need to accurately encode stimulus information in specific signaling patterns that can be transmitted to downstream effectors. Biochemical variability in either dynamic signal encoding or gene expression will prevent the accurate response to changes in the extracellular environment.

In this study, we analyzed the fidelity of information transmission from ligands, through signaling dynamics into gene expression patterns. We generated realistic signaling patterns using a complex mathematical model and asked how much information about ligand identity is lost in the transmission of signaling dynamics into gene expression patterns. It is important to emphasize that the information transmission of signaling dynamics into gene expression patterns does not necessarily imply perfect classification of dynamic patterns into completely distinct patterns. It is possible that the response to different dynamic patterns of transcription factor activity will result in overlapping expression patterns. Indeed, the overall information transmission between ligands and gene expression indicates that information was lost during transmission. Our results indicate that for realistic patterns of NFkB dynamic information transmission into gene expression is capable of preserving most of the information that is encoded in the signaling patterns of NFkB and that most of the information loss occurs during dynamic signal encoding.

We focused on information transmission into gene expression based on a simplified model with a single transcription factor and a single gene. This is a simplification as it was shown that the collective dynamics of multiple transcription factors are needed to encode sufficient information on the extracellular environment to generate signal specific downstream responses [51]. It is interesting to ask whether more complex gene expression models, i.e. multiple genes or more complex models of a single gene that were shown to be important for NFkB [60], can further increase information transmission fidelity. Our results show that even a single gene can recover most of the information that exists within NFkB dynamics. Therefore, while additional genes or more complex models of expression from a single gene that allow promoter to transition between states or include feedforward regulation will undoubtedly have higher capacity, information transmission will be bounded by the information loss in the encoding step itself. Where more complex gene expression models might be important is in allowing model parameters that are more restrictive. Restrictive parameters could be a result of evolutionary constraints and the need for different cell types to show distinct responses to the same ligands. Under such conditions, a more complex gene expression model might increase the overall information transmission fidelity. Similarly, more complex encoding, i.e. capturing ligand identify through the dynamics by multiple transcription factors, can only improve overall information transmission fidelity. Furthermore, our analysis was based on mapping the entire dynamics (i.e. full timeseries) of NFkB into the dynamics of gene expression. It will be interesting to analyze what aspects of the dynamic patterns are more meaningful using either a features based approach [56] or through the use of information transmission rate [57].

The numerical analysis we performed identified specific values that can be used to maximize transmission from signaling dynamics into gene expression patterns. It is interesting to ask whether 1) these parameters are physiological, and 2) what is the intuition behind the values generated by the optimization procedure. Given the simplicity of the model, it is unlikely that all genes in a cell have these parameters. It is likely that the simplistic model used here truly captures the complexity of physiological gene expression responses and therefore the physiological relevance of the specific parameter values identified is limited. However, that does not mean that the parameter values themselves do not have any meaning, intuition developed by interpreting these values could have significance to larger more complex models. Our

optimization showed that the key parameters changes were RNA half-life, the delay between signal and gene expression and cooperativity in gene expression. The effect of lowering RNA half-life is that more of the high frequency information in the dynamic patterns is preserved, pointing out that indeed fast changes in NFkB dynamics on the order of ~20 minutes are meaningful. The reduction in the delay is a bit more nuanced. In a deterministic system the delay value would simply cause a shift between signaling and expression and should not cause any changes to mutual information. However, with the addition of parameter variability a lower mean delay value, that will have lower variance, will help unify the mapping between NFkB and gene expression. Finally, the reduction in cooperativity could be explained as an effective increase in the dynamic range of the dynamics to expression mapping. Overall, while the specific values of the identified parameters should not be taken at face value, it is likely that the principles exposed by these changes will hold in more complex physiological setting.

Dynamical signaling patterns are observed in multiple signaling systems. Our analysis focused on a single case study of NFkB. It will be important to address questions of information transmission fidelity, both computationally and experimentally, in other signaling systems. As our understanding of the role of signaling patterns increases so is the desire to utilize this understanding to design better therapeutic approaches [71,72]. Understanding the mechanism of accurate information transmission from dynamic signals into gene expression, and the constraints that influence information transmission fidelity is an important step in the rational design and manipulation of these dynamical targets [71].

Methods

Quantification of information transduction

To estimate mutual information transduced by NFkB and downstream gene, we applied a binless strategy [47]. This approach uses an embedding of each simulated temporal response into a vector space and calculates Shannon's entropy based on the k-nearest neighbor Euclidean distance within these vector spaces. We started derivation of an algorithm to quantify information flow by considering a 5-dimensional column input signal vector L, where each vector element defines corresponding extracellular ligand concentration ($L = [l_1, l_2, l_3, l_4, l_5]$). After exposure to each l_i ligand, cells response (time trajectories of NFkB and downstream gene: $R_i = [r_{i1}, r_{i2}, \ldots \ldots, r_{in_i}]$) differentially due to cell-to-cell parameter variability. Each response trajectory was projected as a single point in continuous Euclidean space of dimension d, where d is the number of time points picked from an output trajectory. From log-normal distribution containing reference mean and CV values of each parameter, we sampled 1000 parameter sets for *in silico* experiment under each signaling condition. Thus, we generated $N = \sum_{i=1}^{5} n_i = 5000$ response trajectories in our response $R = [R_1, R_2, R_3, R_4, R_5]$ array that mitigates to construct a response distribution in d dimensional space having 5000 points. Following published formalism [66,73] accomplished with k-nearest neighbor estimator, we quantified Shannon differential entropies such as the sum of the conditional entropy of each signal to get overall conditional entropy and non-conditionall entropy $H(R)$. Both entropies quantification are unattainable without the knowledge of input signals' (L) probability (q_i). However the maximum information transduction defined as channel capacity is calculated by maximizing function in optimization method to optimize q_i's such that $\sum_{i=1}^{5} q_i = 1$ where $q_i \geq 0$. In this study, we considered equiprobable ligands probability ($q_i = \frac{1}{5}$) because our observable is mutual information not channel capacity. Using the classical formula ($MI(R;L) = H(R) - H(R|L)$),

information flow (*MI*) between input (*L*) and output (*R*) was estimated. In encoding step *MI* calculation, *R* array was generated from NFkB trajectories whereas downstream gene trajectories gave *R* array in gene expression step *MI* calculation

To convert a dynamic response into a *d* dimensional vector form, we have to opt for the appropriate time points. A response (*R*) array, evaluated at the time points, gives maximum information tranduction. We applied a simple strategy to determine vector dimension and the time points. We equally spliced the time frame of response dynamics (1 to 480 minutes) into *V* +2 number to get dimension *V*, after removing the first and last value. For *V* = 1, it choses response value at the center of the time frame and *V* = 2, it picks values at one third and two thirds of the time frame. In our analysis, we calculated *MI* for encoding (NFkB) and gene expression dynamics up to *V* = 8 (Figure A in S1 Text) considering 25% parameter CV. The results indicate that *MI* value converge. To avoid the curse of dimensionality in *MI* calculation, we considered response value at five [50, 100, 150, 300, 450] time points (in minutes) not only to generate vector space that capturing maximum dynamical feature of NFkB and target gene but also to attain the converged *MI* values. In *MI* estimation, we used k = 10 for k-nearest neighbor Euclidean distance calculation.

We checked possible bias in sample size associated with *MI* value that asymptotically converges for a large population of cell (sample size→∞). *MI* vs inverse of sample size measurement, performed at aforementioned time points, shows nearly horizontal line for encoding and gene expression response (Figure B in S1 Text). The results indicate the bias in mutual information estimation is nominal for our considering sample size.

Stochastic optimization

Optimization was implemented using simulated annealing (SA) to decipher the gene expression model parameter set that can transduce maximum information (Maximal transmission Parameter Set = *arg max MI*(*Gene*[*parameters*])). The optimization objective function follows an information-maximization approach between ligand and gene expression. During the simulation, a randomly chosen variable among 6 gene regulation variables was allowed to make a move for each sampling. The maximum step length taken in our simulation is of 5% with respect to the value of the particular variable (in log10 scale) at the prior sampling step. To be explicit, for any parameter (Param), the update was done by the SA rule, Param0 = Param + Param × $(-1)^n$ × δ × rand, where Param0 is the updated value of Param, n is a random integer [0 or 1], δ is the amplitude of allowed change (kept 5%), and rand is a uniform random number between 0 and 1. Using the information of the updated parameter, mutual information is calculated for the new set of vector spaces after each iteration and rejection or acceptance of updated parameter obeys SA rule. With the progress of iteration, the mutual information initially goes up and converges as the output becomes close to the desired value. The mutual information will be maximal if MI value at the ith step of the iteration is close to the i-1th step

$$\Delta MI = MI_i - MI_{i-1} \approx 0$$

In the optimization code, the precision criterion to terminate the iteration of SA running is $\Delta MI < 10^{-2}$. During the optimization, we didn't impose any boundary constraints on parameters value. Optimized parameter sets that are very close to each other, however each optimization run was started from different guess parameter sets. At this point, it is important to mention that we didn't investigate how parameters in the encoding step change information transmission from ligand to dynamic signaling patterns. The encoding model contains 95 kinetic parameters that are quite difficult to optimize.

Parameter sampling

To simulate with extrinsic noise, model parameters were sampled for the dynamic signal encoding and gene expression models (Fig 1A and 1B). For the encoding models, parameters were sampled from a log-normal distribution centered around the reference values with an identical coefficient of variability (CV) for all parameters. For the gene expression model, parameters were also sampled from a log-normal distribution centered around the guess or optimized values with an equal coefficient of variability (CV) for all parameters. To find good guess or starting parameters value for optimization, we used a multivariate 6-dimensional uniform distribution. Distribution scales are bounded within biologically relevant parameter range i.e., $k_0 = [27e^{-4}\ 0.7]$, $k_s = [6e^{-2}\ 1.72]$, $k_d = [5e^{-4}\ 23e^{-3}]$, $k_b = [4e^{-2}\ 2.46]$, $n = [2,6]$, $\tau = [24\ 120]$.

Here all parameters are unit less expect τ in minutes. Thus, we divided all NFkB response by the maximum response of it among the 5 ligands condition to make it dimensionless that regulates a non-dimensionalized ODE model of downstream gene. We sampled encoding parameters from a log-normal distribution centered around reference values with CV of (10–35)% and calculated *MI* for the encoding step (Fig 2A).

Ligand specificity calculation

To address ligand specificity in downstream gene expression, we evaluated population average response of gene expression for all time points in pre and post optimized parameters conditions for each ligand (mean of 1000 trajectories) and for all ligands (mean of 5000 trajectories). We scaled all the average responses by the total value (sum of all time points) of the average response of all ligands and quantified Euclidean distance between the response of each ligand and all ligands. Fig 4E depicts the mean and standard deviation of 5 Euclidean distances associated with 5 ligands.

Supporting information

S1 Text. Supporting text related to methods used in this paper. The supplementary text includes five figures and a table that support the validity of some of the technical aspects of the methods used.
(PDF)

Author Contributions

Conceptualization: Roy Wollman.

Formal analysis: Alok Maity.

Investigation: Alok Maity.

Methodology: Alok Maity.

Project administration: Roy Wollman.

Software: Alok Maity.

Supervision: Roy Wollman.

Validation: Roy Wollman.

Writing – original draft: Roy Wollman.

Writing – review & editing: Roy Wollman.

References

1. Lim W, Mayer B, Pawson T. Cell Signaling. 1 edition. Garland Science; 2014.

2. Bhalla US, Iyengar R. Emergent properties of networks of biological signaling pathways. Science. 1999; 283: 381–387. https://doi.org/10.1126/science.283.5400.381 PMID: 9888852

3. Behar M, Hoffmann A. Understanding the temporal codes of intra-cellular signals. Curr Opin Genet Dev. 2010; 20: 684–693. https://doi.org/10.1016/j.gde.2010.09.007 PMID: 20956081

4. Purvis JE, Lahav G. Encoding and decoding cellular information through signaling dynamics. Cell. 2013; 152: 945–956. https://doi.org/10.1016/j.cell.2013.02.005 PMID: 23452846

5. Behar M, Hao N, Dohlman HG, Elston TC. Dose-to-duration encoding and signaling beyond saturation in intracellular signaling networks. PLoS Comput Biol. 2008; 4: e1000197. https://doi.org/10.1371/journal.pcbi.1000197 PMID: 18846202

6. Sorre B, Warmflash A, Brivanlou AH, Siggia ED. Encoding of temporal signals by the TGF-β pathway and implications for embryonic patterning. Dev Cell. 2014; 30: 334–342. https://doi.org/10.1016/j.devcel.2014.05.022 PMID: 25065773

7. Thurley K, Tovey SC, Moenke G, Prince VL, Meena A, Thomas AP, et al. Reliable encoding of stimulus intensities within random sequences of intracellular Ca2+ spikes. Sci Signal. 2014; 7: ra59. https://doi.org/10.1126/scisignal.2005237 PMID: 24962706

8. Marshall CJ. Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. Cell. 1995; 80: 179–185. https://doi.org/10.1016/0092-8674(95)90401-8 PMID: 7834738

9. Dolmetsch RE, Xu K, Lewis RS. Calcium oscillations increase the efficiency and specificity of gene expression. mediated by reversible bonds between cell surface molecules Science. 1998; 392: 933–936.

10. Hoffmann A, Levchenko A, Scott ML, Baltimore D. The I B–NF- B Signaling Module: Temporal Control and Selective Gene Activation. Science. 2002; 298: 1241. https://doi.org/10.1126/science.1071914 PMID: 12424381

11. Hao N, O'Shea EK. Signal-dependent dynamics of transcription factor translocation controls gene expression. Nat Struct Mol Biol. 2012; 19: 31–39.

12. Batchelor E, Loewer A, Mock C, Lahav G. Stimulus-dependent dynamics of p53 in single cells. Mol Syst Biol. 2011; 7: 488. https://doi.org/10.1038/msb.2011.20 PMID: 21556066

13. Singh V, Nemenman I. Simple biochemical networks allow accurate sensing of multiple ligands with a single receptor. PLoS Comput Biol. 2017; 13: e1005490. https://doi.org/10.1371/journal.pcbi.1005490 PMID: 28410433

14. Tikhonov M, Little SC, Gregor T. Only accessible information is useful: insights from gradient-mediated patterning [Internet]. arXiv [q-bio.MN]. 2015. Available: http://arxiv.org/abs/1501.07342

15. Hansen AS, O'Shea EK. Promoter decoding of transcription factor dynamics involves a trade-off between noise and control of gene expression. Mol Syst Biol. John Wiley & Sons, Ltd; 2013; 9. https://doi.org/10.1038/msb.2013.56 PMID: 24189399

16. Hansen AS, O'Shea EK. Limits on information transduction through amplitude and frequency regulation of transcription factor activity. Elife. 2015; 4. https://doi.org/10.7554/eLife.06559 PMID: 25985085

17. Sanchez A, Garcia HG, Jones D, Phillips R, Kondev J. Effect of promoter architecture on the cell-to-cell variability in gene expression. PLoS Comput Biol. 2011; 7: e1001100. https://doi.org/10.1371/journal.pcbi.1001100 PMID: 21390269

18. Rieckh G, Tkačik G. Noise and information transmission in promoters with multiple internal States. Biophys J. 2014; 106: 1194–1204. https://doi.org/10.1016/j.bpj.2014.01.014 PMID: 24606943

19. Tkacik G, Callan CG Jr, Bialek W. Information capacity of genetic regulatory elements. Phys Rev E Stat Nonlin Soft Matter Phys. 2008; 78: 011910. https://doi.org/10.1103/PhysRevE.78.011910 PMID: 18763985

20. Tkačik G, Callan CG, Bialek W. Information flow and optimization in transcriptional regulation. Proc Natl Acad Sci U S A. National Academy of Sciences; 2008; 105: 12265–12270. https://doi.org/10.1073/pnas.0806077105 PMID: 18719112

21. Tkacik G, Walczak AM, Bialek W. Optimizing information flow in small genetic networks. Phys Rev E Stat Nonlin Soft Matter Phys. APS; 2009; 80: 031920. https://doi.org/10.1103/PhysRevE.80.031920 PMID: 19905159

22. Walczak AM, Tkacik G, Bialek W. Optimizing information flow in small genetic networks. II. Feed-forward interactions. Phys Rev E Stat Nonlin Soft Matter Phys. APS; 2010; 81: 041905. https://doi.org/10.1103/PhysRevE.81.041905 PMID: 20481751

23. Sanchez A, Golding I. Genetic determinants and cellular constraints in noisy gene expression. Science. 2013; 342: 1188–1193. https://doi.org/10.1126/science.1242975 PMID: 24311680

24. Carey LB, van Dijk D, Sloot PMA, Kaandorp JA, Segal E. Promoter sequence determines the relationship between expression level and noise. PLoS Biol. 2013; 11: e1001528. https://doi.org/10.1371/journal.pbio.1001528 PMID: 23565060

25. Granot-Atedgi E, Tkačik G, Segev R, Schneidman E. Stimulus-dependent maximum entropy models of neural population codes. PLoS Comput Biol. 2013; 9: e1002922. https://doi.org/10.1371/journal.pcbi.1002922 PMID: 23516339

26. Schneidman E, Bialek W, Berry MJ 2nd. Synergy, redundancy, and independence in population codes. J Neurosci. 2003; 23: 11539–11553. https://doi.org/10.1523/JNEUROSCI.23-37-11539.2003 PMID: 14684857

27. Tkačik G, Marre O, Amodei D, Schneidman E, Bialek W, Berry MJ 2nd. Searching for collective behavior in a large network of sensory neurons. PLoS Comput Biol. 2014; 10: e1003408. https://doi.org/10.1371/journal.pcbi.1003408 PMID: 24391485

28. Petkova MD, Tkačik G, Bialek W, Wieschaus EF, Gregor T. Optimal Decoding of Cellular Identities in a Genetic Network. Cell. 2019; 176: 844–855.e15. https://doi.org/10.1016/j.cell.2019.01.007 PMID: 30712870

29. Dubuis JO, Tkacik G, Wieschaus EF, Gregor T, Bialek W. Positional information, in bits. Proc Natl Acad Sci U S A. 2013; 110: 16301–16308. https://doi.org/10.1073/pnas.1315642110 PMID: 24089448

30. Gregor T, Tank DW, Wieschaus EF, Bialek W. Probing the limits to positional information. Cell. 2007; 130: 153–164. https://doi.org/10.1016/j.cell.2007.05.025 PMID: 17632062

31. Zagorski M, Tabata Y, Brandenberg N, Lutolf MP, Tkačik G, Bollenbach T, et al. Decoding of position in the developing neural tube from antiparallel morphogen gradients. Science. 2017; 356: 1379–1383. https://doi.org/10.1126/science.aam5887 PMID: 28663499

32. Uda S, Kuroda S. Analysis of cellular signal transduction from an information theoretic approach. Semin Cell Dev Biol. 2016; 51: 24–31. https://doi.org/10.1016/j.semcdb.2015.12.011 PMID: 26710655

33. Tkačik G, Bialek W. Information Processing in Living Systems. Annu Rev Condens Matter Phys. Annual Reviews; 2016; 7: 89–117.

34. Levchenko A, Nemenman I. Cellular noise and information transmission. Curr Opin Biotechnol. 2014; 28: 156–164. https://doi.org/10.1016/j.copbio.2014.05.002 PMID: 24922112

35. Bowsher CG, Swain PS. Environmental sensing, information transfer, and cellular decision-making. Curr Opin Biotechnol. 2014; 28: 149–155. https://doi.org/10.1016/j.copbio.2014.04.010 PMID: 24846821

36. Waltermann C, Klipp E. Information theory based approaches to cellular signaling. Biochim Biophys Acta. 2011; 1810: 924–932. https://doi.org/10.1016/j.bbagen.2011.07.009 PMID: 21798319

37. Nandagopal N, Santat LA, LeBon L, Sprinzak D, Bronner ME, Elowitz MB. Dynamic Ligand Discrimination in the Notch Signaling Pathway. Cell. 2018; 172: 869–880.e19. https://doi.org/10.1016/j.cell.2018.01.002 PMID: 29398116

38. Wilson MZ, Ravindran PT, Lim WA, Toettcher JE. Tracing Information Flow from Erk to Target Gene Induction Reveals Mechanisms of Dynamic and Combinatorial Control. Mol Cell. Elsevier; 2017; 67: 757–769.e5. https://doi.org/10.1016/j.molcel.2017.07.016 PMID: 28826673

39. Ziv E, Nemenman I, Wiggins CH. Optimal signal processing in small stochastic biochemical networks. PLoS One. 2007; 2: e1077. https://doi.org/10.1371/journal.pone.0001077 PMID: 17957259

40. Tostevin F, ten Wolde PR. Mutual information between input and output trajectories of biochemical networks. Phys Rev Lett. 2009; 102: 218101. https://doi.org/10.1103/PhysRevLett.102.218101 PMID: 19519137

41. Toettcher JE, Weiner OD, Lim WA. Using optogenetics to interrogate the dynamic control of signal transmission by the Ras/Erk module. Cell. 2013; 155: 1422–1434. https://doi.org/10.1016/j.cell.2013.11.004 PMID: 24315106

42. Sigal A, Milo R, Cohen A, Geva-Zatorsky N, Klein Y, Liron Y, et al. Variability and memory of protein levels in human cells. Nature. 2006; 444: 643–646. https://doi.org/10.1038/nature05316 PMID: 17122776

43. Wada T, Wataya M, Fujii M, Hironaka K-I, Eto M, Uda S, et al. Single-cell information analysis reveals small intra- and large intercellular variations increase cellular information capacity [Internet]. bioRxiv. 2019. p. 653832. https://doi.org/10.1101/653832

44. Tudelska K, Markiewicz J, Kochańczyk M, Czerkies M, Prus W, Korwek Z, et al. Information processing in the NF-κB pathway. Sci Rep. Nature Publishing Group; 2017; 7: 15926. https://doi.org/10.1038/s41598-017-16166-y PMID: 29162874

45. Chevalier M, Venturelli O, El-Samad H. The Impact of Different Sources of Fluctuations on Mutual Information in Biochemical Networks. PLoS Comput Biol. 2015; 11: e1004462. https://doi.org/10.1371/journal.pcbi.1004462 PMID: 26484538

46. Tsuchiya T, Fujii M, Matsuda N, Kunida K, Uda S, Kubota H, et al. System identification of signaling dependent gene expression with different time-scale data. PLoS Comput Biol. 2017; 13: e1005913. https://doi.org/10.1371/journal.pcbi.1005913 PMID: 29281625

47. Selimkhanov J, Taylor B, Yao J, Pilko A, Albeck J, Hoffmann A, et al. Systems biology. Accurate information transmission through dynamic biochemical signaling networks. Science. 2014; 346: 1370–1373. https://doi.org/10.1126/science.1254933 PMID: 25504722

48. Uda S, Saito TH, Kudo T, Kokaji T, Tsuchiya T, Kubota H, et al. Robustness and compensation of information transmission of signaling pathways. Science. 2013; 341: 558–561. https://doi.org/10.1126/science.1234511 PMID: 23908238

49. Cheong R, Rhee A, Wang CJ, Nemenman I, Levchenko A. Information transduction capacity of noisy biochemical signaling networks. Science. 2011; 334: 354–358. https://doi.org/10.1126/science.1204553 PMID: 21921160

50. Potter GD, Byrd TA, Mugler A, Sun B. Dynamic Sampling and Information Encoding in Biochemical Networks. Biophys J. 2017; 112: 795–804. https://doi.org/10.1016/j.bpj.2016.12.045 PMID: 28256238

51. Granados AA, Pietsch JMJ, Cepeda-Humerez SA, Farquhar IL, Tkačik G, Swain PS. Distributed and dynamic intracellular organization of extracellular information. Proc Natl Acad Sci U S A. 2018; 115: 6088–6093. https://doi.org/10.1073/pnas.1716659115 PMID: 29784812

52. Voliotis M, Perrett RM, McWilliams C, McArdle CA, Bowsher CG. Information transfer by leaky, heterogeneous, protein kinase signaling systems. Proc Natl Acad Sci U S A. 2014; 111: E326–33. https://doi.org/10.1073/pnas.1314446111 PMID: 24395805

53. Bowsher CG, Voliotis M, Swain PS. The fidelity of dynamic signaling by noisy biomolecular networks. PLoS Comput Biol. 2013; 9: e1002965. https://doi.org/10.1371/journal.pcbi.1002965 PMID: 23555208

54. Lane K, Van Valen D, DeFelice MM, Macklin DN, Kudo T, Jaimovich A, et al. Measuring Signaling and RNA-Seq in the Same Cell Links Gene Expression to Dynamic Patterns of NF-κB Activation. Cell Syst. 2017; 4: 458–469.e5. https://doi.org/10.1016/j.cels.2017.03.010 PMID: 28396000

55. Micali G, Aquino G, Richards DM, Endres RG. Accurate encoding and decoding by single cells: amplitude versus frequency modulation. PLoS Comput Biol. 2015; 11: e1004222. https://doi.org/10.1371/journal.pcbi.1004222 PMID: 26030820

56. Makadia HK, Schwaber JS, Vadigepalli R. Intracellular Information Processing through Encoding and Decoding of Dynamic Signaling Features. PLoS Comput Biol. journals.plos.org; 2015; 11: e1004563. https://doi.org/10.1371/journal.pcbi.1004563 PMID: 26491963

57. Grabowski F, Czyż P, Kochańczyk M, Lipniacki T. Limits to the rate of information transmission through MAPK pathway [Internet]. bioRxiv. 2018. p. 402750. https://doi.org/10.1101/402750

58. Hormoz S. Cross talk and interference enhance information capacity of a signaling pathway. Biophys J. 2013; 104: 1170–1180. https://doi.org/10.1016/j.bpj.2013.01.033 PMID: 23473500

59. Habibi I, Cheong R, Lipniacki T, Levchenko A, Emamian ES, Abdi A. Computation and measurement of cell decision making errors using single cell data. PLoS Comput Biol. 2017; 13: e1005436. https://doi.org/10.1371/journal.pcbi.1005436 PMID: 28379950

60. Lee REC, Walker SR, Savery K, Frank DA, Gaudet S. Fold change of nuclear NF-κB determines TNF-induced transcription in single cells. Mol Cell. 2014; 53: 867–879. https://doi.org/10.1016/j.molcel.2014.01.026 PMID: 24530305

61. Zhang Q, Gupta S, Schipper DL, Kowalczyk GJ, Mancini AE, Faeder JR, et al. NF-κB Dynamics Discriminate between TNF Doses in Single Cells. Cell Syst. 2017; 5: 638–645.e5. https://doi.org/10.1016/j.cels.2017.10.011 PMID: 29128333

62. Zambrano S, De Toma I, Piffer A, Bianchi ME, Agresti A. NF-κB oscillations translate into functionally related patterns of gene expression. Elife. 2016; 5: e09100. https://doi.org/10.7554/eLife.09100 PMID: 26765569

63. Taylor B, Adelaja A, Liu Y, Luecke S, Hoffmann A. Identification and physiological significance of temporal NFκB signaling codewords deployed by macrophages to classify immune threats. bioRxiv. 2020. p. 2020.05.23.112862. https://doi.org/10.1101/2020.05.23.112862

64. Cheng Z, Taylor B, Ourthiague DR, Hoffmann A. Distinct single-cell signaling characteristics are conferred by the MyD88 and TRIF pathways during TLR4 activation. Sci Signal. 2015; 8: ra69. https://doi.org/10.1126/scisignal.aaa5208 PMID: 26175492

65. Lee TK, Denny EM, Sanghvi JC, Gaston JE, Maynard ND, Hughey JJ, et al. A noisy paracrine signal determines the cellular NF-kappaB response to lipopolysaccharide. Sci Signal. 2009; 2: ra65. https://doi.org/10.1126/scisignal.2000599 PMID: 19843957

66. Hoffmann A. Immune Response Signaling: Combinatorial and Dynamic Control. Trends Immunol. 2016; 37: 570–572. https://doi.org/10.1016/j.it.2016.07.003 PMID: 27461000

67. Selimkhanov J, Taylor B, Yao J, Pilko A, Albeck J, Hoffmann A, et al. Systems biology. Accurate information transmission through dynamic biochemical signaling networks. Science. 2014; 346: 1370–1373. https://doi.org/10.1126/science.1254933 PMID: 25504722

68. Yao J, Pilko A, Wollman R. Distinct cellular states determine calcium signaling response. Mol Syst Biol. 2016; 12: 894. https://doi.org/10.15252/msb.20167137 PMID: 27979909

69. Hughey JJ, Gutschow MV, Bajar BT, Covert MW. Single-cell variation leads to population invariance in NF-κB signaling dynamics. Mol Biol Cell. 2015; 26: 583–590. https://doi.org/10.1091/mbc.E14-08-1267 PMID: 25473117

70. Rowland MA, Greenbaum JM, Deeds EJ. Crosstalk and the evolvability of intracellular communication. Nat Commun. 2017; 8: 16009. https://doi.org/10.1038/ncomms16009 PMID: 28691706

71. Behar M, Barken D, Werner SL, Hoffmann A. The dynamics of signaling as a pharmacological target. Cell. 2013; 155: 448–461. https://doi.org/10.1016/j.cell.2013.09.018 PMID: 24120141

72. Bugaj LJ, Sabnis AJ, Mitchell A, Garbarino JE, Toettcher JE, Bivona TG, et al. Cancer mutations and targeted drugs can disrupt dynamic signal encoding by the Ras-Erk pathway. Science. 2018; 361. https://doi.org/10.1126/science.aao3048 PMID: 30166458

73. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. Phys Rev E Stat Nonlin Soft Matter Phys. 2004; 69: 066138. https://doi.org/10.1103/PhysRevE.69.066138 PMID: 15244698