

# SCIENTIFIC REPORTS



OPEN

## Computational Methods for Estimating Molecular System from Membrane Potential Recordings in Nerve Growth Cone

Tatsuya Yamada<sup>1</sup>, Makoto Nishiyama<sup>2</sup>, Shigeyuki Oba<sup>3</sup>, Henri Claver Jimbo<sup>4</sup>, Kazushi Ikeda<sup>1</sup>, Shin Ishii<sup>3</sup>, Kyonsoo Hong<sup>2,5</sup> & Yuichi Sakumura<sup>4,6</sup>

Biological cells express intracellular biomolecular information to the extracellular environment as various physical responses. We show a novel computational approach to estimate intracellular biomolecular pathways from growth cone electrophysiological responses. Previously, it was shown that cGMP signaling regulates membrane potential (MP) shifts that control the growth cone turning direction during neuronal development. We present here an integrated deterministic mathematical model and Bayesian reversed-engineering framework that enables estimation of the molecular signaling pathway from electrical recordings and considers both the system uncertainty and cell-to-cell variability. Our computational method selects the most plausible molecular pathway from multiple candidates while satisfying model simplicity and considering all possible parameter ranges. The model quantitatively reproduces MP shifts depending on cGMP levels and MP variability potential in different experimental conditions. Lastly, our model predicts that chloride channel inhibition by cGMP-dependent protein kinase (PKG) is essential in the core system for regulation of the MP shifts.

Estimation or determination of unknown functions or targets by indirect measurements, e.g. the estimation of the DNA double helix from its X-ray diffraction pattern<sup>1</sup>, the functional connectivity of neurons from the activation pattern<sup>2</sup>, the identification of cell type from the gene expression pattern<sup>3</sup>, and cancer diagnosis from breath gas components<sup>4</sup> have been demonstrated. We present a computational derivation for the estimation of biomolecular interactions from an observed time series of electrophysiological activities recorded from nerve growth cones.

Axon guidance is essential for establishing a neuronal network during nervous system development<sup>5,6</sup>. The biomolecular signaling pathways that instruct the direction of a navigating growth cone have been intensively investigated<sup>7–10</sup>. Many studies<sup>11–13</sup>, including our's<sup>8,14</sup>, have shown that the second messenger, cGMP, is a downstream effector of the guidance cue, Sema3A. A growth cone normally exhibits a repulsive response to a Sema3A gradient<sup>11,15</sup>. However, this repulsion converts to attraction if the intracellular cGMP level is elevated<sup>8,16</sup>. Our studies revealed that the growth cone turning direction depends on the state of the growth cone membrane potential (MP); a hyperpolarized or depolarized state induces, respectively, either repulsion or attraction in response to many diffusible guidance molecules<sup>16</sup>. Furthermore, it has been shown that a low level of cGMP causes growth cone hyperpolarization, whereas a high level of cGMP causes depolarization<sup>16</sup>, demonstrating that a cGMP signal regulates the MP shifts, which determine the growth cone turning direction. The signaling cascade that converts the guidance cue-induced biomolecular system to electrical signals, however, remains largely unknown.

To computationally estimate the biomolecular network responsible for axon guidance from growth cone MP recordings, three major hurdles must be overcome: 1. the limited availability of the recording data due to the amount of labor required; 2. the large cell-to-cell variability<sup>17–19</sup>, which affects the observed MP; 3. multiple unknown factors that potentially cooperate to regulate the molecular network. The current computational study

<sup>1</sup>Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan. <sup>2</sup>Department of Biochemistry, New York University School of Medicine, New York, USA. <sup>3</sup>Graduate School of Informatics, Kyoto University, Kyoto, Japan. <sup>4</sup>Graduate School of Biological Sciences, Nara Institute of Science and Technology, Nara, Japan. <sup>5</sup>KASAH Technology, Inc, New York, USA. <sup>6</sup>School of Information Science and Technology, Aichi Prefectural University, Aichi, Japan. Correspondence and requests for materials should be addressed to K.H. (email: [kyonsoo.hong@kasahtech.com](mailto:kyonsoo.hong@kasahtech.com)) or Y.S. (email: [saku@bs.naist.jp](mailto:saku@bs.naist.jp))

considers the limited availability of data by utilizing each MP time series (MPTS) that contain over 10,000 data points within one recording sample, thereby providing sufficient data points to perform quantitative computational analysis and to fit a deterministic model to a cell-dependent characteristic. To extract a reliable estimation of a biomolecular network from the small sample size of MP data sets, we applied Bayesian a reverse-engineering framework<sup>20,21</sup> that has been an especially effective method for studying small number data sets and has been successfully applied in many neuroscience studies<sup>2,3,22–24</sup>. Briefly, by the Bayesian reverse-engineering framework of the system comprised of different physical quantities, we computed the posterior distributions of the parameters that are derived from a fitness of the deterministic biochemical reaction model developed using the experimental MP data sets and prior constraints. Second, the computational study considers the cell-to-cell variability by expressing it as probability distributions of the model parameters<sup>18,20,21,25</sup>. Lastly, the study considers the involvement of multiple unknown factors in the signaling pathway by developing a signaling cascade-based model that simplifies the multiple bio-molecular cascades, and introduced MPTSs data sets into the model.

Many studies have utilized the Bayesian framework to deduce potential unknown biochemical interactions from direct biochemical reaction measurements<sup>18,19,26–28</sup>. We present a novel approach that estimates the molecular signaling pathway from electrophysiological recordings. The current study addresses the system uncertainty and cell-to-cell variability by applying parameter categorization to separate the model parameters into two categories: cell-common core system and cell-dependent peripheral properties. To select the most plausible molecular pathway from multiple model candidates, we applied the Bayesian evidence<sup>29,30</sup> that satisfies model simplicity while considering all possible parameter ranges. Our combined Bayesian reverse-engineered framework with a mathematical model approach successfully reproduces the quantitative dependency of steady-state MP shifts based on cGMP levels, and reveals that chloride channel inhibition by PKG is essential for MP shifts. Thus, we provide a novel computational methodology to estimate the essential molecular signaling components in transducing the electrical responses elicited during growth cone turning.

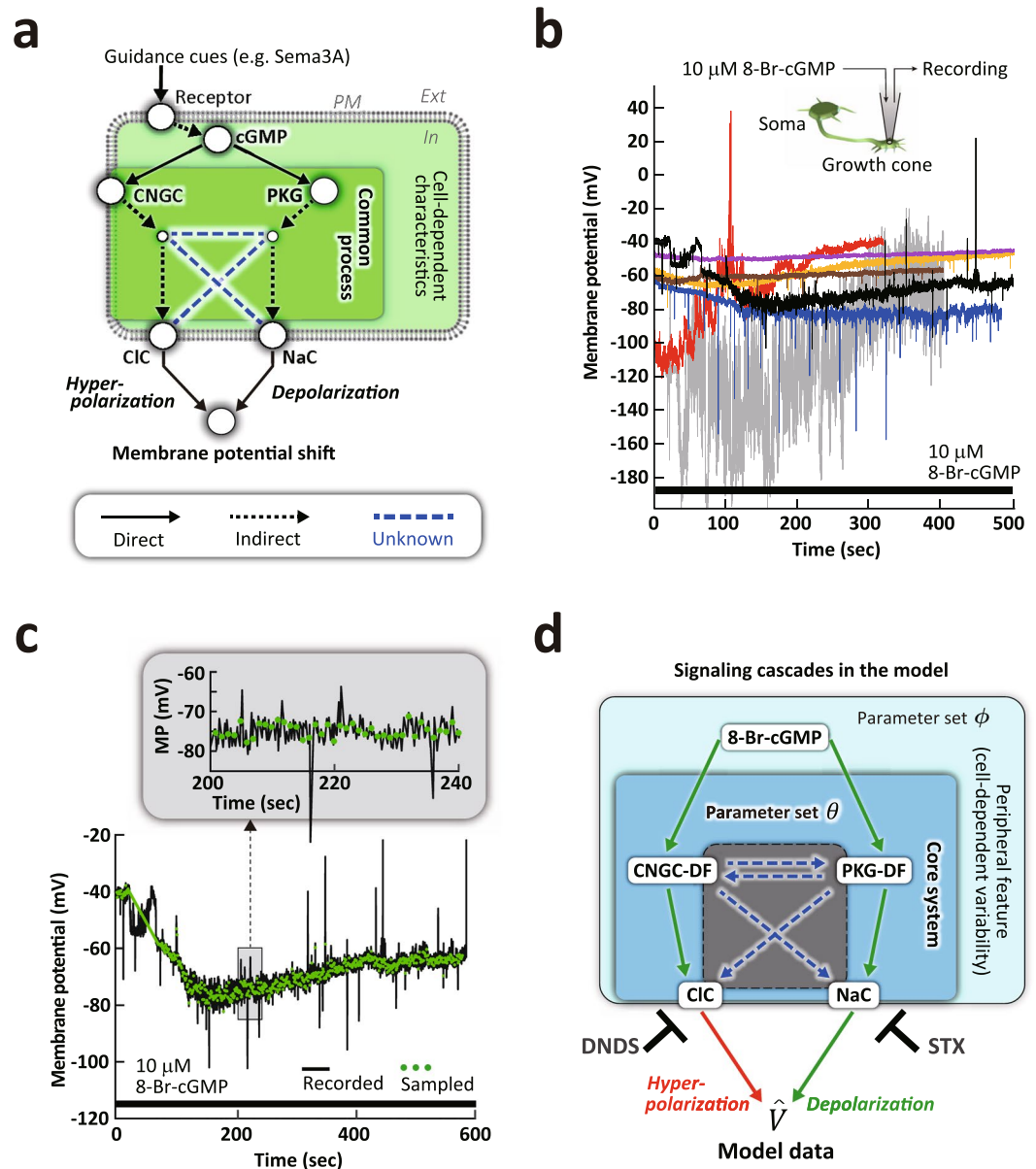
## Results

### A mathematical model of cGMP signaling inducing growth cone membrane potential shifts.

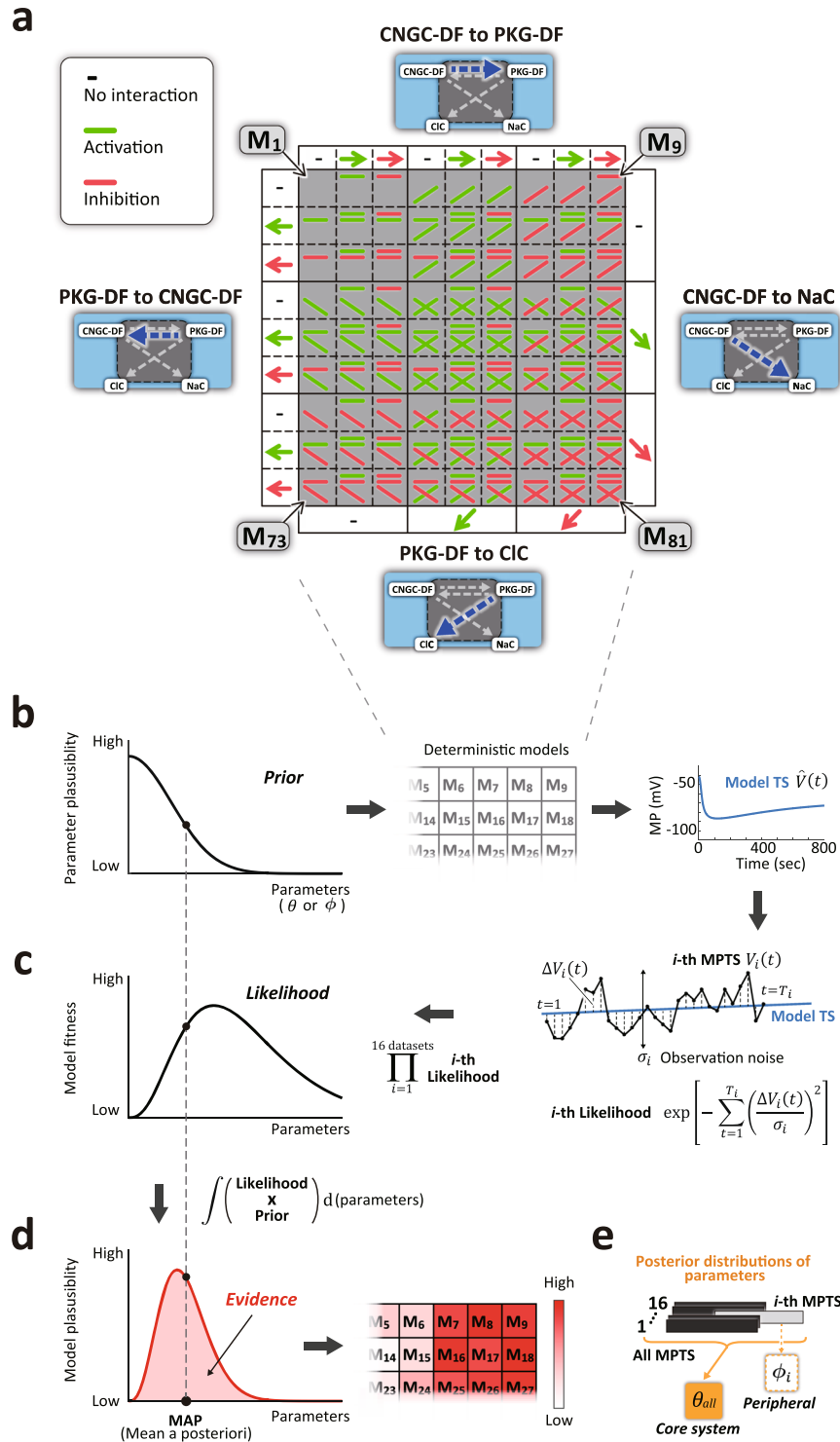
The induction of membrane potential (MP) shifts, i.e., hyperpolarization to depolarization and vice versa by cGMP signaling, involves multifaceted molecular, biochemical, and biophysical processes in the growth cone (Fig. 1a). Diffusible guidance molecules, such as *Sema3A*, increase the intracellular cGMP level, and depending on the level of increase, either cGMP-mediated hyperpolarization (at low cGMP level) or PKG-mediated depolarization (at high cGMP level) occurs, which subsequently causes, respectively, growth cone repulsion or attraction<sup>16</sup>. However, how hyperpolarizing and depolarizing channel activities are regulated in response to different levels of a cGMP signal that regulates the bidirectional MP shifts is largely unknown.

We used the abundant data points from MP recordings from *Xenopus* spinal neuron growth cones (Fig. 1b; see Methods). The time series of each MP (MPTS) recorded during 10 minutes in response to stimulation by 8-Br-cGMP (a membrane permeable cGMP analogue injected into the growth cone by a recording pipette) is comprised of more than 10,000 data points (see Fig. 1b). We eliminated the noise, such as spikes, by sampling the recorded MPTS at 1 sec intervals (green dots in Fig. 1c, Supplementary Fig. 1). The sampled MPTS provides a sufficient number of data points (several hundred) to allow us to develop a quantitative and deterministic mathematical model to dissect the cGMP signaling responsible for inducing growth cone MP shifts. Our model incorporates the mesoscopic molecular signal flows within a core system of the model with parameter set,  $\theta$ , that regulate chloride and sodium channels (ClC and NaC), respectively, hyperpolarizing and depolarizing channels that regulate MP shifts (Fig. 1d). To extract a partial subsystem within the core system, the model also considers the absence of each channel function conditions (Fig. 1d). Previous studies revealed potential interactions between CNGC-downstream factor (DF) and ClC<sup>31,32</sup> and between PKG-DF and NaC<sup>33,34</sup> (Methods). Therefore, our model considers four possible unknown molecular interactions within the core system (Fig. 1d, blue dashed arrows in the gray box), which were modeled by deterministic equations (see details in Methods). As shown in Fig. 1b, the growth cone MPTSs are relatively diverse despite the induction by an identical cGMP stimulation. This is likely due mainly to the difference in intrinsic properties of each growth cone. Therefore, the model also considers the peripheral parameter set,  $\phi$  (Fig. 1d), which is modeled by diffusion process and the well-established Hodgkin and Huxley quantitative mathematical model of MP regulation<sup>35</sup>, to characterize cell-to-cell variability, such as growth cone size (from 5  $\mu\text{m}$  up to 10  $\mu\text{m}$ )<sup>36,37</sup> and shape that affect the intracellular cGMP diffusion rate, as well as growth cone dynamic behavior, i.e., exocytosis and endocytosis that modulate ion channel densities<sup>38</sup>.

**System identification based on a Bayesian framework.** In the mathematical model, we considered three potential interactions between one molecule and another within the core system: activation, inhibition, and no interaction. Considering the four possible unknown molecular interactions within the core system (Fig. 1d), a total of 81 ( $3^4 = 81$ ) possible models arise, from the simplest model,  $M_1$  (Model 1) with no interaction to the most complex model,  $M_{81}$  with four possible interactions ( $M_1$  to  $M_{81}$ ; Fig. 2a). We examined the model fitness of each of these 81 deterministic candidate models. Briefly, we computed the model MPTS using the core system and peripheral parameters ( $\theta$  and  $\phi$ ) constrained by prior probability distributions (Fig. 2b). Due to our configuration of model candidates, complex models have more capability to fit the given dataset than the simple ones that are enclosed by the complex ones (e.g.  $M_{81}$  encloses  $M_1$ ). To extract a sufficient condition of signaling pathways, it is necessary to give room for complex models to approximate simple ones by reducing the values of parameters that are not common between the complex and the simple ones. We designed the priors as left-truncated normal distributions for most of the model parameters to satisfy this requirement. If using a log-normal distribution or a Gamma distribution whose shape parameter is greater than one, the probability density of parameter value being zero is zero. In contrast, a left-truncated normal distribution where the peak probability density becomes zero is useful because complex model can cut off unnecessary pathways by setting unnecessary parameters to zero. Thus



**Figure 1.** Conversion of a biomolecular signal to an electrophysiological signal in a nerve growth cone. **(a)** Signaling pathways from an extracellular guidance cue, e.g. Semaphorin 3A, to membrane potential (MP) shifts in a growth cone. Semaphorin 3A, upon binding to its receptor on the plasma membrane (PM), increases the intracellular cGMP level. The effector proteins, Cyclic nucleotide-gated ion channels (CNGC) and cGMP-dependent protein kinases (PKG) that directly bind to cGMP (solid arrows) trigger, respectively, the activation of chloride channels (CIC) and sodium channels (NaC), which induce either hyperpolarization or depolarization. The potential cross interactions between the downstream effectors, which are unknown, are indicated by dashed blue lines. **(b)** Growth cone recording stimulated by a stimulant, 10  $\mu\text{M}$  8-Br-cGMP, a membrane permeable cGMP analogue (Top). Samples of recorded growth cone MP time series (MPTS; control:  $n = 7$ ). **(c)** The MPTS contains over 10,000 data points (black). Outlier noises, such as spikes, are removed by sampling the data points at a 1 sec interval (green dots; see upper inset, enlargement of dashed square region; Supplementary Fig. 1). **(d)** The schematic diagram illustrates the mathematical model that functions as an alternative system and is designed in mesoscopic scale that expresses effective molecular signal flows. The control model contains the signaling pathway from the cGMP stimulation to MP shifts, via CNGC- and PKG-downstream factors (DFs), corresponding to the pathways in (a). The unknown signal flows between CNGC-DF, PKG-DF, CIC, and NaC in the core system are indicated within the dashed grey box. The parameters of the unknown flows and known factors, i.e., CNGC-DF to CIC and PKG-DF to NaC, are involved in the blue box and characterized by the core system parameter set,  $\theta$ . The 8-Br-cGMP stimulation and MP, which are outside of the core system, are characterized by cell-dependent peripheral parameter set,  $\phi$ . We modeled the diversified data as a core system with peripheral elements representing cell-to-cell variability. Two pharmacological experimental conditions were also modeled, which represent complete inhibition of CIC, blocked by DNDS ( $n = 5$ ) and NaC, blocked by STX ( $n = 4$ ) (see Supplementary Fig. 2). A total of 16 datasets (control:  $n = 7$ ; DNDS:  $n = 5$ ; STX:  $n = 4$ ) were analyzed.



**Figure 2.** Model candidates and Bayesian system identification procedure. **(a)** The matrix shows the possible interactions in the gray region in Fig. 1d; pathways from CNGC-DF to PKG-DF, CNGC-DF to NaC, PKG-DF to CIC, and PKG-DF to CNGC-DF in the clockwise direction. Each of four pathways has three potential signals, i.e. activation (green), inhibition (red), or no interaction, as modeled by the core system parameters ( $\theta$ ). Thus, there are  $3^4 = 81$  model candidates labeled as  $M_1$  to  $M_{81}$ . **(b)** Schematic illustration of prior distributions of model parameters that models the core system uncertainty and cell-dependent variability, expressing that extremely large values are biologically implausible and result in overfitting to the data (left; left-truncated Gaussians). Parameter values are applied to each of the models listed in **a** to generate model time series (TS) ( $\hat{V}(t)$ ; blue line in the right). **(c)** Schematic illustration of model fitness, which depends on the parameters (left). The model fitness is the product of all the model likelihood ( $10\mu\text{M}$  8-Br-cGMP; control/DNDS/STX;  $n = 16$ ), and the  $i$ -th likelihood with the parameter values in **b** was computed as the product of Gaussian probability density functions of the  $i$ -th MPTS (black;  $V(t)$ ) for the entire time points ( $t = 1, \dots, T_i$ ) with the mean  $\hat{V}(t)$

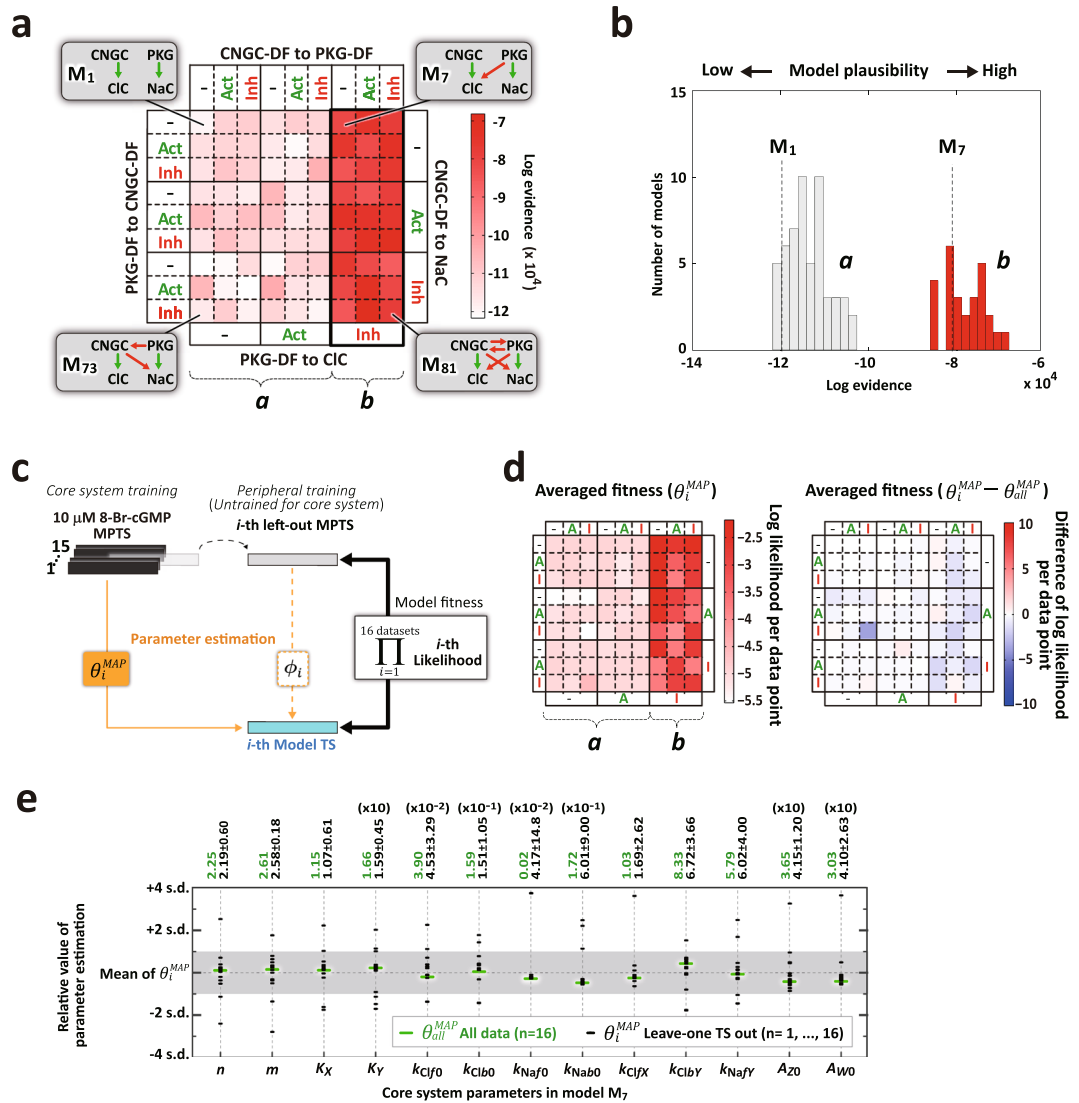
(blue) and the s.d.  $\sigma_i$  (right). **(d)** Illustrates the model plausibility (left; Bayesian evidence), which evaluates both model fitness and parameter plausibility. The model plausibility was computed by integrating the product of the model fitness and the parameter plausibility over the entire ranges of all the parameters (red area in the left). The logarithmic evidences for all the 81 models were displayed in the matrix corresponding to that in **a** with red-white color (right; higher evidence is colored by red). **(e)** The core system parameters were estimated from all the datasets ( $\theta_{all}$ ;  $n = 16$ ) and cell-dependent peripheral parameters of the  $i$ -th cell ( $\phi_i$ ;  $n = 1$ ) were from only the  $i$ -th cell's MPTS, as posterior probability distributions. When we use representative specific parameter values instead of distributions during validation steps, we selected mean a posteriori (MAP) values (black circle in **d**) and denote the parameter set with superscript "MAP" like  $\theta_{all}^{MAP}$ .

we introduced the left-truncated Gaussians for the priors of the most model parameters (Supplementary Table 1). We determined the prior s.d. values based on the Monte Carlo simulation of cGMP diffusion (Supplementary Fig. 4), the number of binding sites of cGMP to CNGC and PKG, and the range and time scale of MP shift of experimental measurements. We then compared the error between the single experimental MPTS and model MPTS as the likelihood function. Due to the presence of cell-dependent noise in the experimental MPTS (Fig. 1b,c; Supplementary Fig. 1), we standardized the error by each noise level (Fig. 2c, right). We estimated the noise level ( $\sigma_i$ ) from the errors between the sampled MPTS (green dots in Fig. 1c) and the smoothed time average of the same MPTS (red line in Supplementary Fig. 1; see Methods). This process was iterated for total given data points of all the MPTSs and multiplied all the likelihoods to obtain the total likelihood as model fitness for all the data sets. The model fitness, which depends on the parameter values (Fig. 2c, left), provides only the similarity between the experimental and the model MPTSs and ignores the parameter plausibility. Thus, the model fitness alone is insufficient to select a plausible model, as the parameters are likely to take unreasonable values when the model overfits the data set.

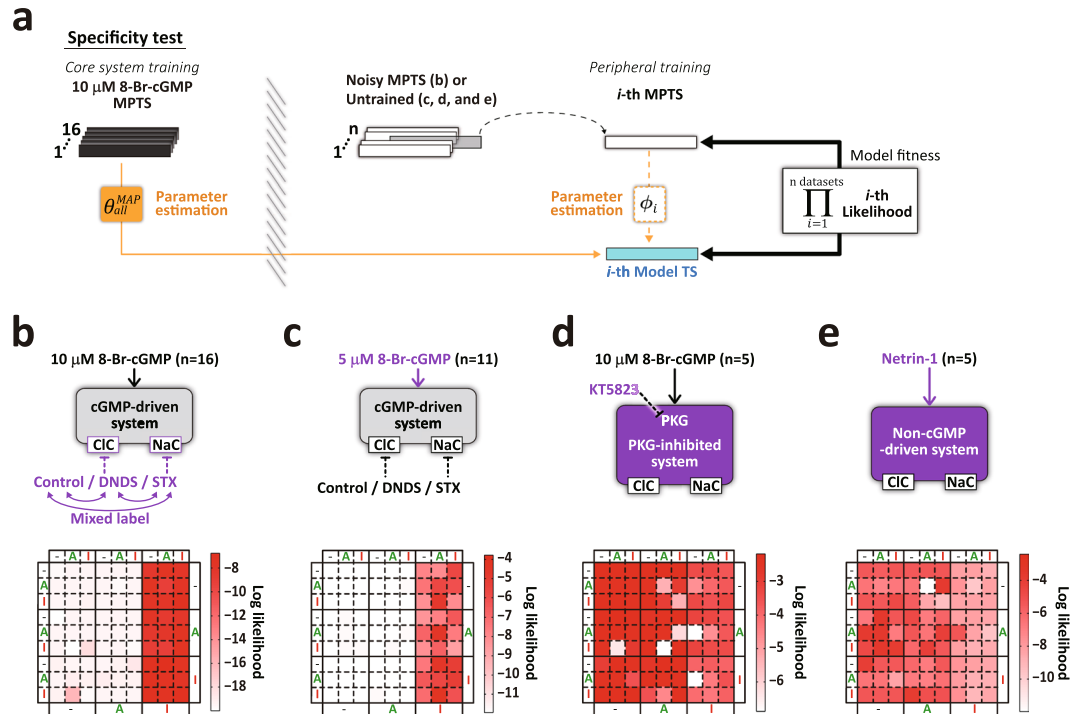
Therefore, we further computed the model plausibility by incorporating the parameter plausibility using a Bayesian framework that avoids the parameters from overfitting the data set. The model plausibility is then expressed as the product of the likelihood as model fitness and the prior as parameter plausibility. To take into account the model plausibility for the entire parameter range, we calculated the integral of the product, the Bayesian evidence<sup>29,30</sup> (see Methods; Fig. 2d, left) for each of the total 81 deterministic models (Fig. 2d, right), by a Monte Carlo simulation (Supplementary Methods). We took the logarithm of the evidence (log-evidence) as index of model plausibility; the larger the log-evidence, the more plausible the model. For parameter estimation, we computed the core system parameters that were estimated from all the experimental MPTS data sets (total  $n = 16$ : 10  $\mu$ M 8-Br-cGMP, control,  $n = 7$ ; with DNDS blocking CIC,  $n = 5$ ; with STX blocking NaC,  $n = 4$ ) expressed as  $\theta_{all}$  (Fig. 2e). The peripheral parameters (expressed as  $\phi_i$ ), which are cell-dependent, were only estimated from the  $i$ -th MPTS from the total MPTS data sets for computing the model fitness (Fig. 2e). As we aim to identify the molecular signaling network within the core system from the 10  $\mu$ M 8-Br-cGMP-induced MPTS data set, we utilized the core system parameters that were estimated from the total data sets for the validation, specificity, and predictability tests.

**Identification and validation of the core system.** As shown in Fig. 3a, the evidence computed from the experimental MPTS data sets ( $n = 16$ ; see Methods) for each deterministic models revealed two distinct groups: group *a* with smaller evidences (light pink region) and group *b* with larger evidences (red region) (Fig. 3a). The difference of the evidences between the models within each group is much smaller than between the two groups (Fig. 3b), suggesting that the molecular pathway commonly present in the group *b* models may play a significant role. Noticeably,  $M_7$ , which represents the simplest molecular interactions present within group *b*, reveals that PKG-mediated CIC inhibition is involved in cGMP signaling during MP shifts.

To evaluate how stable the core system parameters were estimated, we examined the generality of the estimated core system parameters by performing leave-one-out (LOO) cross validation (LOO-CV)<sup>39</sup>. We first separated the 16 data sets into two subsets: one with 15 data sets to train the core system parameters ( $\theta_i^{MAP}$ ; except for the  $i$ -th MPTS) and the other with only one data set (left-out MPTS) to train the peripheral parameters ( $\phi_i$ ; the  $i$ -th MPTS). The left-out MPTS was used for computing the likelihood as an untrained data, as it was not incorporated into the core system parameters. We iterated the training and computing procedure 16 times for the total samples ( $i = 1, \dots, 16$ ) and multiplied 16 likelihoods to obtain the model fitness (Fig. 3c; see Method). If the core system parameter estimation is stable, the matrix representation of model fitness to untrained data sets should be similar to that of the trained data sets and the core system parameters in  $M_7$  should not be significantly altered by LOO-CV. The average of the model fitness values (averaged fitness with  $\theta_i^{MAP}$ ), which are expressed in log likelihood per data point (green dots in Fig. 1c), appears small in group *a* and large in group *b* (Fig. 3d, left panel). This average fitness shows insignificant differences between LOO data sets ( $\theta_i^{MAP}$ ) and all data sets ( $\theta_{all}^{MAP}$ ) (Fig. 3d, right panel), suggesting that the untrained MPTS (left-out MPTS) can be reproduced by the original models with the parameters estimated from the 15 separated MPTSs. In addition, the maximum log likelihood evaluation for the training with LOO data set (Supplementary Fig. 8) showed that the models are clustered into same two groups as those by the evidence (Fig. 3a,b). Although these are only maximum likelihoods, considering Supplementary Fig. 7, these results suggest that model group *b* is stably superior to group *a* independent of data set combination within LOO-CV. Furthermore, for the core system parameter values, the computation that resulted from performing LOO-CV shows that each of the mean a posteriori (MAP) parameters in  $M_7$  ( $\theta_i^{MAP}$ ;  $i = 1, \dots, 16$ ) is close to the original MAP parameters estimated from the complete data set ( $\theta_{all}^{MAP}$ ) (Fig. 3e; within the LOO mean  $\pm$  s.d.). These results verify that our identification method, which depends on the stability of the core system



**Figure 3.** Most plausible model and the generality test by leave-one-out cross validation. **(a)** Computed model evidences (logarithm evidences; Fig. 2d) are expressed as a matrix representation following Fig. 2a (10  $\mu$ M 8-Br-cGMP; control/DNDS/STX;  $n = 16$ ). A model with large log evidence (toward greater red) indicates greater plausibility than that with smaller log evidence (toward white). Models in left six columns are labeled as model group *a*; those in right three columns are labeled as model group *b*. Models indicated by  $M_1$  and  $M_7$  (insets) contain, respectively, the fewest and common interactions in groups *a* and *b*, representing minimal models within the respective model groups. **(b)** Histograms of log evidence in the model groups *a* and *b*, as shown in **a**. Vertical dashed lines indicate the log evidence of the two representative models,  $M_1$  and  $M_7$ . **(c)** Schematic procedure of the generality test of the core system parameter estimations by applying the leave-one-out (LOO) data selection to the models. Black and gray horizontal bars represent experimental MPTS. When the *i*-th MPTS is left out from the 10  $\mu$ M 8-Br-cGMP datasets ( $n = 16$ ), it is used to estimate peripheral parameters ( $\phi_i$ ) and validate model fitness (product of all the likelihood in white box), whereas the remaining MPTS ( $n = 15$ ) are completely separated from the left-out MPTS to estimate the MAP core system parameters ( $\theta_i^{MAP}$ ). Repeating LOO for all the datasets, 16 system parameter sets are obtained ( $\theta_i^{MAP}$ ;  $i = 1, \dots, 16$ ). The model fitness to the left-out MPTS was taken the logarithm and averaged over all the data points. **(d)** Matrix representation of the averaged model fitness by LOO-derived parameter ( $\theta_i^{MAP}$ ) (left) and the difference between the fitness and that computed from all the dataset ( $\theta_i^{MAP} - \theta_{all}^{MAP}$ ) (right). **(e)** Distributions of the estimated MAP core system parameter values given by LOO in a ( $\theta_i^{MAP}$ ; black point) compared with those by all the datasets ( $\theta_{all}^{MAP}$ ; green circle;  $n = 16$ ). The parameters exhibited in the horizontal axis and their values are represented in the vertical axis with relative representation with respect to the mean (actual values are listed in upper axis with the corresponding color). Most of the estimations are within the mean  $\pm$  s.d. (gray region).



**Figure 4.** Specificity of identified core system to 8-Br-cGMP-induced MPTS. **(a)** Schematic procedure of model specificity test using MPTSs in different conditions using MAP core system parameters estimated from all the 10  $\mu\text{M}$  8-Br-cGMP datasets ( $\theta_{all}^{MAP}$ ;  $n = 16$ ; black horizontal bars in the left side of the shaded band). The model likelihood to the  $i$ -th MPTS from other testing datasets (white bars) was computed while computing the cell-dependent parameters,  $\phi_i$ , and by repeating this the model fitness to all the  $n$  datasets were obtained. Testing datasets are the datasets untrained by the core system (5  $\mu\text{M}$  8-Br-cGMP-induced MPTS, PKG-inhibited condition (KT5823), and netrin-1-induced MPTS), those with noise (10  $\mu\text{M}$  8-Br-cGMP-induced MPTS with randomly labeled channel inhibitor). **(b and c)** Matrix representations of model fitness by the datasets derived by the control core system; 10  $\mu\text{M}$  8-Br-cGMP-induced MPTS with mixed labels of channel conditions (control, DNDS, STX) (**b**;  $n = 16$ ) by randomly selecting channel blocker conditions in the model (see Supplementary Methods), and 5  $\mu\text{M}$  8-Br-cGMP-induced MPTS (**c**; total  $n = 11$ ; control:  $n = 2$ ; with DNDS blocking CIC:  $n = 5$ ; with STX blocking NaC:  $n = 4$ ). **(d and e)** Same as **b and c**, but the datasets derived by the different core systems; 10  $\mu\text{M}$  8-Br-cGMP-induced MPTS under the PKG-inhibited condition (KT5823) (**d**), and netrin-1-induced MPTS (**e**).

parameters, is not significantly affected by the data set combination of LOO-CV, indicating that the estimated core system parameters are reasonably stable.

**Specificity of trained membrane potential time series.** Next, we tested the specificity of the identified model to the given trained data sets – 10  $\mu\text{M}$  8-Br-cGMP-induced MPTS (Fig. 1b). To test the degree of specificity, we introduced untrained data sets of MPTSs induced under different experimental conditions. We then compared their model fitness with that derived from the original trained data sets (Fig. 3a) by computing log likelihoods per data point using the core system parameters estimated from the experimental 10  $\mu\text{M}$  8-Br-cGMP MPTSs ( $\theta_{all}^{MAP}$ ; Fig. 4a). First, we introduced the mixed labeled channel condition models (control/STX/DNDS to random labels) (Fig. 4b). As each MPTS is the sum of CIC and NaC components, the pattern of the model fitness matrix is not expected to be significantly altered by mixing the label of the data sets. As shown in Fig. 4b, the model fitness matrix clearly shows separation of group *a* and *b* as in Fig. 3a. When we introduced the MPTS induced by a different concentration of stimulant – 5  $\mu\text{M}$ , instead of 10  $\mu\text{M}$  8-Br-cGMP, the model fitness matrix shows a similar pattern of model groups, as in Fig. 3a (Fig. 4c). This might indicate that core system parameters have a small dependency on the stimulus intensity between 5  $\mu\text{M}$  and 10  $\mu\text{M}$  concentration of 8-Br-cGMP.

In contrast, when we introduced the MPTS induced by 10  $\mu\text{M}$  8-Br-cGMP in the presence of KT5823 ( $n = 5$ ), a PKG activity inhibitor (Supplementary Fig. 3a) that abolishes the PKG activation in the core system, the model fitness matrix shows the total absence of the distinct segregation between the model group *a* and *b* (Fig. 4d). This supports that PKG activation is required for the pattern of model groups observed in Fig. 3a, at least partially. Lastly, we introduced the MPTS induced by netrin-1, a diffusible guidance molecule (Supplementary Fig. 3b) that regulates MP shifts through a different molecular signaling pathway<sup>40</sup>. As shown in Fig. 4e, the pattern of the model fitness matrix shown in Fig. 3a is totally abolished (Fig. 4e; Supplementary Methods), supporting that our models are specifically trained for cGMP signaling. Taken together, these tests indicate that our system identification method has high specificity.

**Predictability of the untrained late phase of MPTS.** We have demonstrated that our integrated methodology for training the model parameters specifically derived from electrophysiological data sets monitored in response to cGMP stimulation in nerve growth cones has the capability to identify the core molecular system of cGMP signaling. We further determined whether the identified models in group *b* have the capability of predicting the untrained, late phase of MPTS, when the peripheral parameters were trained for the initial phase of MPTS. Although the estimation of the initial phase MPTS from the late phase is possible in principle, it is practically impossible because the late phase MPTS contains little information on the signaling pathways due to steady state of cGMP concentration. Specifically, we examined the accuracy of forecasting the late phase data points of the reference initial phase of MPTS, using a data assimilation test<sup>41</sup> in which the core system, as well as the peripheral parameters for the late phase MPTS, are not incorporated. First, we set the core system parameters to the MAP values that derived from the data sets of MPTS induced by 10  $\mu$ M 8-Br-cGMP, except for the *i*-th MPTS ( $\theta_i^{MAP}$ ;  $n = 15$ ). The peripheral parameters were trained for the initial phase of the *i*-th left-out MPTS ( $\phi_i$ ), which was iterated 16 times ( $i = 1, \dots, 16$ ;  $n = 16$ ). Subsequently, we computed the model fitness for the late phase of the MPTS (250 to 800 sec) and projected the predicted late phase of MPTS (indicated as the predicted MPTS in Fig. 5a). As shown in Fig. 5b, the representative models,  $M_1$  and  $M_7$ , predicted the trajectories of the late phase MPTSs induced by 10  $\mu$ M 8-Br-cGMP, demonstrating the significant superiority of  $M_7$  over  $M_1$  in its capability to predict the late phase of MPTS. Normalized root mean squared errors (RMSEs) were also computed with the time series used in Fig. 5c to examine the superiority of group *b* to group *a* by other evaluation criteria (Supplementary Fig. 9a). We found that RMSEs also shows that group *b* models have a high ability to predict the late phase MPTS, consistent with Fig. 5c. We further confirmed the essential requirement of the core molecular system for the predictability by replacing the 5  $\mu$ M 8-Br-cGMP data sets for the peripheral parameters training. As in Fig. 5d, the core system parameters were trained for all the MPTSs of 10  $\mu$ M 8-Br-cGMP data sets ( $\theta_{all}^{MAP}$ ;  $n = 16$ ) while the peripheral parameters were trained only for the initial phase of the *i*-th left-out MPTS ( $\phi_i$ ) induced by 5  $\mu$ M 8-Br-cGMP, and the model fitness was computed similarly as in Fig. 5a ( $i = 1, \dots, 11$ ; Fig. 5d). RMSEs with the time series used in Fig. 5f also showed the superiority of the group *b* to the group *a* (Supplementary Fig. 9b). These tests support that the models in the group *b* have the capacity to predict the untrained late phase of MPTS given the generalized core molecular system parameters derived from the 10  $\mu$ M 8-Br-cGMP MPTSs.

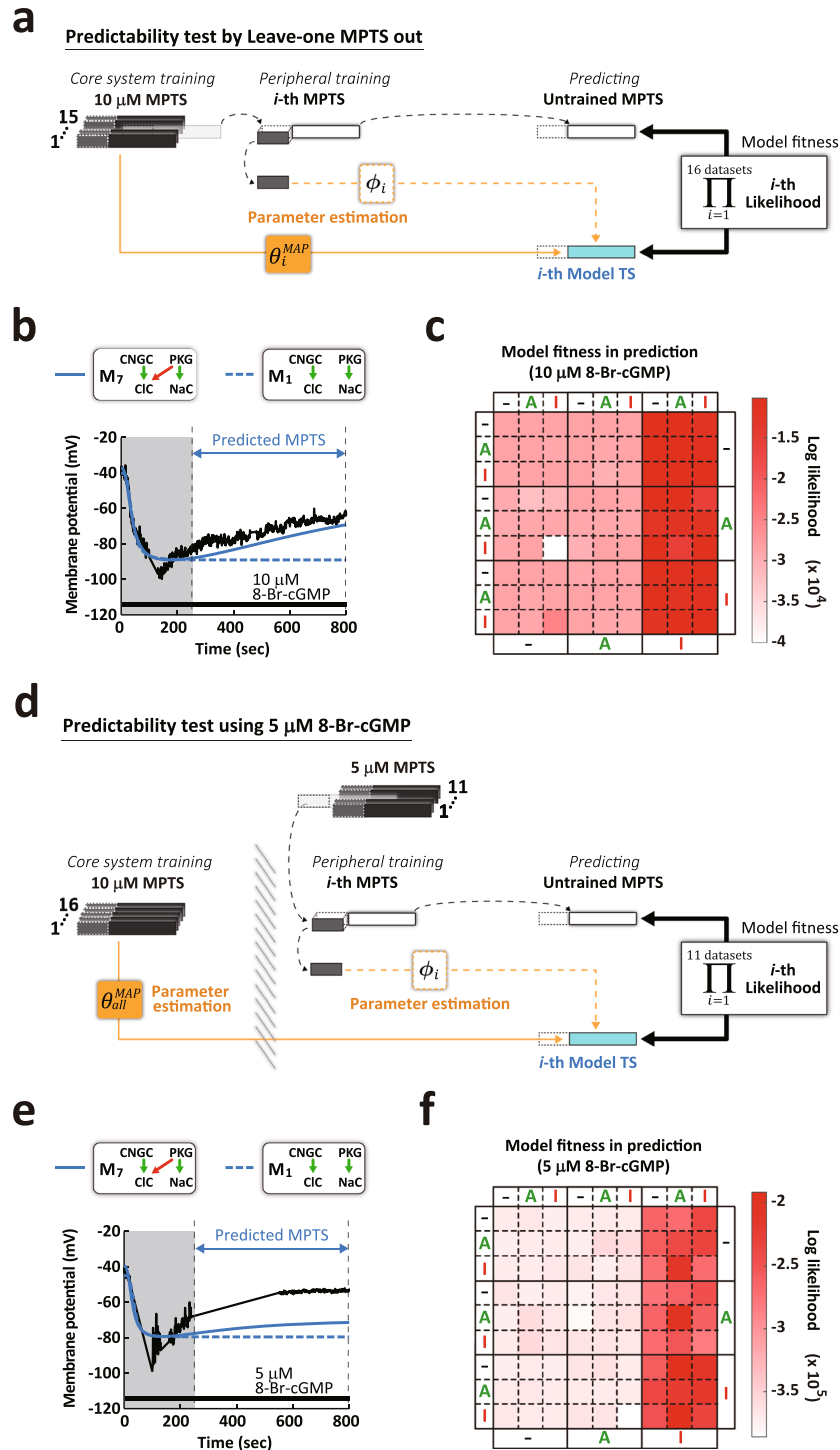
**Reproducibility of cGMP-dependent bidirectional MP shifts.** Upon the injection of cGMP stimulant through the recording pipette into the nerve growth cone, the resting membrane potential (from about  $-80$  to  $-70$  mV) becomes hyperpolarization<sup>16</sup>. As the intracellular cGMP increases, the hyperpolarization slowly converts to depolarization, causing a bidirectional MP shift<sup>16</sup>. We simulated these bidirectional MP shifts, which show temporal hyperpolarization in the initial phase of the stimulation that converts to depolarization and eventually reaches the steady state (after 10 min; Fig. 6a). At the steady state, MPs maintain almost constant values, which have bidirectional dependency on the level of 8-Br-cGMP stimulation<sup>16</sup>. We examined whether the core molecular system containing the PKG inhibition of CIC that was highlighted in  $M_7$  is also required during the induction of cGMP-dependent bidirectional MP shifts. We computed the  $M_7$  with  $\theta_{all}^{MAP}$  estimated from 10  $\mu$ M 8-Br-cGMP ( $n = 16$ ) at different concentrations of cGMP. Our model simulation shows: (i) a gradual hyperpolarization stimulated by a low 8-Br-cGMP (0.5  $\mu$ M) concentration, (ii) a sharp hyperpolarization that gradually recovered to the resting MP level stimulated by a moderate 8-Br-cGMP (6  $\mu$ M) concentration, and (iii) a sharp hyperpolarization that converted to depolarization stimulated by a high concentration of 8-Br-cGMP (20  $\mu$ M). We then compared the steady state MP of the model (blue circles in Fig. 6a) with the experimental data (Fig. 5e in ref.<sup>16</sup>). The simulated cGMP concentration-dependent MP shifts show a bidirectional MP shift (Fig. 6b, a blue dotted line). Although our model simulation showed bidirectional MP shifts, the magnitude of depolarization was much greater and the occurrence of depolarization appeared at much lower 8-Br-cGMP concentration than that which was observed experimentally. This difference may result from the difference between the modeled direct stimulation by 8-Br-cGMP through the recording pipette compared to the bath application of 8-Br-cGMP in the experiment<sup>16</sup>. When we introduced a simple Hill-like model of membrane permeation of 8-Br-cGMP into the model to mimic the bath-application of 8-Br-cGMP (Supplementary Fig. 6; Supplementary Methods), indeed,  $M_7$  (blue solid line in Fig. 6b) was able to reproduce the bidirectional MP shifts that were much closer to those observed experimentally<sup>16</sup> (black squares in Fig. 6b).

## Discussion

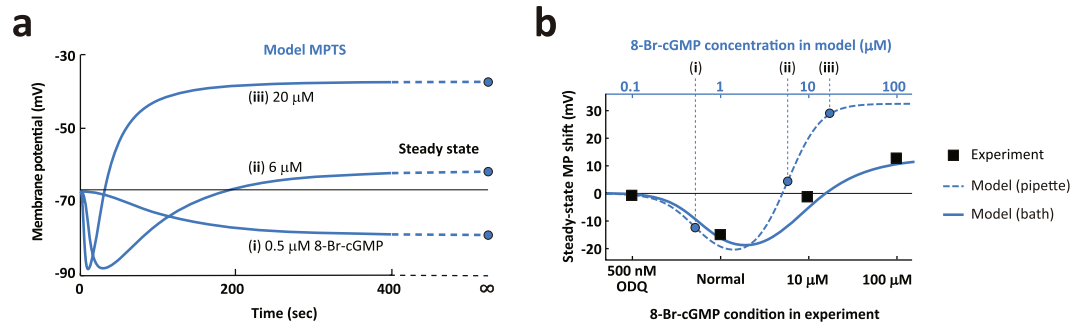
We present a computational analysis that reveals an essential molecular signaling pathway within the core system of MP shifts recorded from a growth cone in response to an external stimulus that directs growth cone turning. We show a novel integrated reverse-engineering of the system comprised of different physical quantities and Bayesian framework methods that accommodate the large cell-to-cell variability and small number of data sets that otherwise hinder the biophysical modeling. By implementing the Bayesian framework, we specifically show the optimization of peripheral parameters for individual cells that overcomes the cell-to-cell variability, and the core system parameters for all given data sets to extract an unknown molecular pathway. Thus, our parameter categorization is especially useful for extracting common molecular pathways involved in electrophysiological responses in a cell.

The model plausibility expressed by Bayesian evidence<sup>29,30</sup> evaluates the overall possible parameter ranges, unlike the Akaike information criterion<sup>42</sup> (AIC; Supplementary Fig. 7a) and the Bayesian information criterion<sup>43</sup> (BIC; Supplementary Fig. 7b), and the maximum log likelihood (Supplementary Fig. 7c). AIC and BIC did not show clearly divided criteria values as shown in Fig. 3a,b. These model selections successfully function when the number of data is large enough. The maximum log likelihood clearly divided likelihood values as shown in Fig. 3a,b.





**Figure 5.** Model predictability of untrained 8-Br-cGMP-induced MPTS. **(a)** Schematic illustration of the procedure for testing model predictability for the late-phase of MPTS induced by 10 μM 8-Br-cGMP. The initial-phase (250 sec) of the left-out  $i$ -th MPTS was extracted to estimate the MAP peripheral parameters ( $\phi_i$ ), and the remaining  $i$ -th Msas used to test the model predictability. The MAP core system parameters were estimated from the initial and late phase of the remaining MPTS ( $\theta_i^{MAP}$ ;  $n = 15$ ). **(b)** Sample MPTS predicted by model  $M_1$  (blue dotted line) and  $M_7$  (blue solid line) by the prediction procedure in **a**. **(c)** Matrix representation of log-likelihood of the predicted MPTS when repeating  $i = 1, \dots, 16$  and summing them. **(d)** Same as **a**, but 5 μM 8-Br-cGMP MPTS was predicted with  $\phi_i$  given the initial-phase MPTS (up to 250 sec) and  $\theta_{all}^{MAP}$  by all the 10 μM 8-Br-cGMP MPTS. Black and white bars represent experimental MPTS. MPTS in the white regions (after 250 sec, the late-phase) were never used for the parameter estimation; these MPTSs were completely untrained data sets even for the peripheral parameters. **(e)** Sample MPTS predicted by model  $M_1$  (blue dotted line) and  $M_7$  (blue solid line) by the prediction procedure in **d**. **(f)** Matrix representation of log-likelihood of the predicted MPTS when repeating  $i = 1, \dots, 11$  and summing them.



**Figure 6.** Reproducibility of cGMP-dependent MP shifts. **(a)** Model MPTS by the model  $M_7$  induced by different 8-Br-cGMP concentration: (i) 0.5, (ii) 6, and (iii) 20  $\mu\text{M}$  with  $\theta_{all}^{MAP}$  and  $\phi_{all}^{MAP}$  (Supplementary Table 1). Each MPTS reaches the steady-state at about 300 sec. We defined the model's steady-state MP shift as the difference between the MPs at times zero and infinity, namely,  $\hat{V}(t \rightarrow \infty) - \hat{V}(t = 0)$ , where the core system parameters ( $\theta_{all}^{MAP}$ ) from the 10  $\mu\text{M}$  8-Br-cGMP dataset were used. For the peripheral parameters,  $V_K$  is a mere constant bias in  $\hat{V}(t)$  (see Supplementary Methods) and, therefore, was canceled out,  $\tau_S$  was ignorable at times zero and infinity, and  $A_Z$  and  $A_W$  were replaced by their means in  $\phi_{all}^{MAP}$ . **(b)** Model of 8-Br-cGMP concentration dependent MP shifts. Prediction of the cGMP-dependency of the steady-state MP shifts induced by the direct pipette stimulation (inject 8-Br-cGMP through the recording pipette; blue dashed line). Mimicking of the bath application condition of 8-Br-cGMP (blue solid line) by correcting the 8-Br-cGMP permeation process from outside to inside of the cell (Supplementary Methods). Experimental data of steady-state MP shifts under the bath application condition (black square with lower x-axis; ref.<sup>16</sup>) are shown for comparison. Normal condition indicates a cGMP level when Sema3A was applied in the turning assay; ODQ (1H-(1,2,4) oxadiazolo(4,3-a) quinoxalin-1-one, a soluble guanylyl cyclase inhibitor) condition reduces growth-cone cGMP levels.

However, the evaluation of the likelihood function is performed at one specific parameter set (maximum likelihood parameters) and hence it is unclear whether high likelihood is also given around this parameter set. Thus, it allows reasonable model selection even when a small number of data sets are given. Normal approaches to deduce a signaling pathway, such as a biophysical modeling, although the model accuracy may be higher, demands a large number of data sets to estimate not only the specific values of the core system parameters, but also the distributions of the peripheral parameters. Collecting a large number of biological data sets is laborious and time-consuming, and data sets such as the MPTSs in this work are difficult to measure. Our computational methodology demonstrates the feasibility of extracting a hidden core molecular system from a small number of data sets and with a large cell-to-cell variability. Thus, our computational analysis, at the least, has the capacity of ranking the possibility of the biomolecular system, which significantly reduces the laborious task and extensive time required by conventional experimentation.

Our computational analyses reveal that PKG-mediated CIC inhibition is an essential pathway that acts in concert with CNGC-mediated CIC activation and PKG-mediated NaC activation, as demonstrated in model  $M_7$ . In support of our model prediction, biochemical studies have shown that PKG, indeed, is a regulator of the Mitogen-activated Protein Kinase (MAPK), such as ERK and p38<sup>44</sup>, both of which inhibit calcium-dependent CIC<sup>45</sup>. Regarding the parameter values, the bidirectional cGMP-dependent MP shifts (Fig. 6b) showed that CNGC-mediated CIC activation is due to a high-affinity of cGMP ( $K_x = 1.15$  in Fig. 3e), whereas PKG-mediated NaC activation is due to a low-affinity of cGMP ( $K_y = 16.61$  in Fig. 3e). The bidirectional phenomenon based on a difference in the dissociation constants of positive and negative regulators has been shown in the molecular system of synaptic plasticity, for example, in the phosphorylation of  $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors that occurs in competition between the kinase and phosphatase<sup>46</sup>. Our study suggests a novel regulatory mechanism of the bidirectional cGMP-dependent MP shifts in growth cone guidance in which PKG not only facilitates membrane depolarization<sup>16</sup>, but also simultaneously inhibits hyperpolarization. Thus, PKG-mediated CIC inhibition could facilitate the dynamic range of MP shifts by opposing a stimulatory input at a small range of 0.1–10  $\mu\text{M}$  cGMP concentration that contributes to the overall dynamic range of the growth cone turning response.

## Methods

**Data preprocessing.** Analyzed experimental data of membrane potential (MP) were recorded from growth cones of cultured *Xenopus* spinal neurons as described previously<sup>16</sup>. The recorded MP time series (MPTS) contain significant noises and we considered three types of noises (Supplementary Fig. 1): 1. spike-like system noise, 2. step-like temporal change that likely due to experimental artifacts, and 3. small observation noises, like thermal fluctuation. In the analysis, we used the sampled MPTS of 300–900 data points (one sec interval; green dots in Fig. 1c) from the raw data (>10000 data points; black line in Fig. 1c) that removed most of the spike-like system noise. The step-like artifacts were complemented with a straight line connecting the onset point and the end point of the temporal change (one sec interval; blue points in Supplementary Fig. 1). On the other hand, we disregarded the small observation noise, as it is too small to remove. Instead, we used scaling model fitness (difference between the data point and the mathematical model). We modeled the small observation noise for each MPTS by a white Gaussian with the s.d. ( $\sigma_j$  in Fig. 2c), which was estimated from the difference between the sampled MPTS

data points and the smoothed data points (red line in Supplementary Fig. 1. Such noise model was applied for standardizing error between observed and model data.

**MP dataset.** We have previously recorded MPTS under following stimulation and culture conditions<sup>16</sup>:

1. 10  $\mu$ M 8-Br-cGMP in the recording pipette, as a stimulant (control,  $n = 7$ ; Fig. 1b), and in the presence of either chloride channel (ClC) blocker (DNDS;  $n = 5$ ; Supplementary Fig. 2a), or sodium channel (NaC) blocker (STX;  $n = 4$ ; Supplementary Fig. 2b) in the culture bath.
2. 5  $\mu$ M 8-Br-cGMP in the recording pipette, as a stimulant (control,  $n = 2$ ; Supplementary Fig. 2c), and in the presence of either chloride channel (ClC) blocker (DNDS;  $n = 5$ ; Supplementary Fig. 2d), or sodium channel (NaC) blocker (STX;  $n = 4$ ; Supplementary Fig. 2e) in the culture bath.
3. 10  $\mu$ M 8-Br-cGMP in the recording pipette as a stimulant in the presence of PKG inhibitor (KT5823;  $n = 5$ ; Supplementary Fig. 3a).
4. 5  $\mu$ M netrin-1, as a stimulant (control,  $n = 5$ ; Supplementary Fig. 3b).

The 10  $\mu$ M 8-Br-cGMP-stimulated datasets that include control, DNDS, and STX, were used for system identification, as they provide the largest number of datasets. The remaining datasets were used for the validation test.

**Molecular signalling pathways.** Previously, we have shown<sup>16</sup> that during bath application of pharmacological drugs to cultured neurons in the presence of 8-Br-cGMP stimulation: 1. MP shifts to depolarization in the presence of DNDS, the ClC blocker; 2. MP shifts to hyperpolarization in the presence of STX; and 3. Application of a PKG inhibitor, KT5823, caused sustained hyperpolarization, supporting that ClCs are required for hyperpolarization; NaCs are required for depolarization; and PKG activity is required for depolarization (Supplementary Fig. 5). It has also been demonstrated that the cGMP-induced hyperpolarization is, in part, due to the activation of CNGCs via cGMP directly activating the channels<sup>14,47</sup>, which ultimately activates the hyperpolarizing Cl channels (ClCs)<sup>31,32</sup>. Likewise, the cGMP-activated PKG, is known to be a regulator of the Mitogen-activated Protein Kinase (MAPK) such as p38<sup>44</sup>, which activates a TTX-resistant sodium channel, Nav1.8<sup>33,34</sup>. Thus, we incorporated these known pathways of NaC and ClC activation by PKG and CNGC, respectively, into our model (Fig. 1d,e).

**Bayesian formulation for the parameter estimation.** Because the parameters of growth cone volumes and ion channel densities affect the MP shifts, we applied a Bayesian framework. The model parameters were categorized into three classes: core system and peripheral parameters, and experimentally derived parameters (Supplementary Table 1).

The core system parameter set, which includes the biochemical reaction rates, Hill coefficients, and means of  $A_{Cl}$  and  $A_{Na}$  ( $A_{Cl0}$  and  $A_{Na0}$ , respectively), was estimated from the total MP data sets, as it is a common system for all neurons.

On the other hand, the peripheral parameter set,  $\phi = \{V_K, A_{Cl}, A_{Na}, \tau_S\}$ , which is highly dependent on characteristics of individual neurons (e.g., growth cone volume, channel densities) was estimated for each MP time series. The experimental condition parameter set,  $c = \{\eta_{Cl}, \eta_{Na}, S_{max}\}$ , represents the experimentally derived parameters, such as pharmacological condition and applied 8-Br-cGMP concentration.

By the Bayesian approach, the model parameters were estimated from the experimental data under constraints given by prior distributions (mostly left-truncated Gaussians; see Supplementary Table 1). The Bayes' theorem delivers the posterior distribution of the model parameters as,

$$p(\phi, \theta | V, M) = \frac{p(V | \phi, \theta) p(\phi, \theta | M)}{E(V, M)} \quad (1)$$

$$E(V, M) = \iint p(V | \phi, \theta, M) p(\phi, \theta | M) d\phi d\theta. \quad (2)$$

Here, the experimental condition parameter set,  $c$ , is omitted for easy visibility. The evidence,  $E(V, M)$ , is the model evaluation criteria (Fig. 2d) determined by the MP dataset,  $V$ , and the model,  $M$ . The parameters of the prior distributions,  $p(\phi, \theta | M)$  (Fig. 2b), were listed in Supplementary Table 1. The likelihood, which represents the fitness level to the given dataset, is given by the product of the Gaussians (Fig. 2c),

$$\begin{aligned} p(V | \phi, \theta) &= \prod_{i=1}^I p(V_i | \phi, \theta) = \prod_{i=1}^I \prod_{t=1}^{T_i} \frac{1}{\sqrt{2\pi} \sigma_i} \exp\left(-\frac{\Delta V_i^2(t)}{2\sigma_i^2}\right) \\ &= \left(\frac{1}{\sqrt{2\pi} \sigma_i}\right)^I \exp\left[-\frac{1}{2} \sum_{i=1}^I \sum_{t=1}^{T_i} \left(\frac{\Delta V_i(t)}{\sigma_i}\right)^2\right], \end{aligned} \quad (3)$$

where  $\Delta V_i(t)$  is defined as the MP difference at the time  $t$  (Fig. 2c),  $\Delta V_i(t) = V_i(t) - \hat{V}(t | \phi, \theta)$ , indicating the error between the observed MP,  $V_i(t)$ , and the model MP,  $\hat{V}(t | \phi, \theta)$  (Eq. (S15) in Supplementary Methods), where  $i$  is the data index ( $I = 16$  and  $T_i = 300 \sim 900$ ). The s.d. of the  $i$ -th MP time series,  $\sigma_i$ , was the size of observation noise, which was computed during the data preprocessing. Model fitness expressed as likelihood is defined in Fig. 2c, and is obtained by taking the natural logarithm of Eq. (3), ignoring constant parameters. The practical calculation of the evidence was performed by Markov Chain Monte Carlo (MCMC) simulation (Supplementary Methods).

**Computation software and time.** All computations were performed with Matlab (MathWorks) and its parallel computing toolbox. We used up to 1,000 core CPUs for Monte Carlo simulation. The calculation of a single evidence took about a day per model (Fig. 3a), and that of a single likelihood took from a half day to a day per validation (Figs 3d, 4b–e, and 5c,f).

**Data availability.** All data and code used to perform analyses reported herein are available from the corresponding author at reasonable request.

## References

- Watson, J. D. & Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).
- Field, G. D. *et al.* Functional connectivity in the retina at the resolution of photoreceptors. *Nature* **467**, 673–677, <https://doi.org/10.1038/nature09424> (2010).
- Gabitto, M. I. *et al.* Bayesian Sparse Regression Analysis Documents the Diversity of Spinal Inhibitory Interneurons. *Cell* **165**, 220–233, <https://doi.org/10.1016/j.cell.2016.01.026> (2016).
- Sakumura, Y. *et al.* Diagnosis by Volatile Organic Compounds in Exhaled Breath from Lung Cancer Patients Using Support Vector Machine Algorithm. *Sensors* **17**, <https://doi.org/10.3390/s17020287> (2017).
- Goodman, C. S. Mechanisms and molecules that control growth cone guidance. *Annu Rev Neurosci* **19**, 341–377, <https://doi.org/10.1146/annurev.ne.19.030196.002013> (1996).
- Yu, T. W. & Bargmann, C. I. Dynamic regulation of axon guidance. *Nat Neurosci* **4**(Suppl), 1169–1176, <https://doi.org/10.1038/nn748> (2001).
- Song, H. & Poo, M. The cell biology of neuronal navigation. *Nat Cell Biol* **3**, E81–88, <https://doi.org/10.1038/35060164> (2001).
- Nishiyama, M. *et al.* Cyclic AMP/GMP-dependent modulation of Ca<sup>2+</sup> channels sets the polarity of nerve growth-cone turning. *Nature* **423**, 990–995 (2003).
- Gomez, T. M. & Zheng, J. Q. The molecular basis for calcium-dependent axon pathfinding. *Nat Rev Neurosci* **7**, 115–125 (2006).
- Forbes, E. M., Thompson, A. W., Yuan, J. & Goodhill, G. J. Calcium and cAMP levels interact to determine attraction versus repulsion in axon guidance. *Neuron* **74**, 490–503, <https://doi.org/10.1016/j.neuron.2012.02.035> (2012).
- Song, H. *et al.* Conversion of neuronal growth cone responses from repulsion to attraction by cyclic nucleotides. *Science* **281**, 1515–1518 (1998).
- Xiang, Y. *et al.* Nerve growth cone guidance mediated by G protein-coupled receptors. *Nat Neurosci* **5**, 843–848, <https://doi.org/10.1038/nn899> (2002).
- Tojima, T., Itofusa, R. & Kamiguchi, H. The nitric oxide-cGMP pathway controls the directional polarity of growth cone guidance via modulating cytosolic Ca<sup>2+</sup> signals. *J Neurosci* **29**, 7886–7897, <https://doi.org/10.1523/JNEUROSCI.0087-09.2009> (2009).
- Togashi, K. *et al.* Cyclic GMP-gated CNG channels function in Sema3A-induced growth cone repulsion. *Neuron* **58**, 694–707 (2008).
- Tessier-Lavigne, M. & Goodman, C. S. The molecular biology of axon guidance. *Science* **274**, 1123–1133 (1996).
- Nishiyama, M., von Schimmelmann, M. J., Togashi, K., Findley, W. M. & Hong, K. Membrane potential shifts caused by diffusible guidance signals direct growth-cone turning. *Nat Neurosci* **11**, 762–771 (2008).
- Selimkhanov, J. *et al.* Systems biology. Accurate information transmission through dynamic biochemical signaling networks. *Science* **346**, 1370–1373, <https://doi.org/10.1126/science.1254933> (2014).
- Filippi, S. *et al.* Robustness of MEK-ERK Dynamics and Origins of Cell-to-Cell Variability in MAPK Signaling. *Cell reports* **15**, 2524–2535, <https://doi.org/10.1016/j.celrep.2016.05.024> (2016).
- Lenive, O., PD, W. K. & MP, H. S. Inferring extrinsic noise from single-cell gene expression data using approximate Bayesian computation. *BMC systems biology* **10**, 81, <https://doi.org/10.1186/s12918-016-0324-x> (2016).
- Barnes, C. P., Silk, D., Sheng, X. & Stumpf, M. P. Bayesian design of synthetic biological systems. *Proc Natl Acad Sci USA* **108**, 15190–15195, <https://doi.org/10.1073/pnas.1017972108> (2011).
- Liepe, J. *et al.* A framework for parameter estimation and model selection from experimental data in systems biology using approximate Bayesian computation. *Nat Protoc* **9**, 439–456, <https://doi.org/10.1038/nprot.2014.025> (2014).
- Brown, E. N., Frank, L. M., Tang, D., Quirk, M. C. & Wilson, M. A. A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *J Neurosci* **18**, 7411–7425 (1998).
- Mortimer, D. *et al.* Bayesian model predicts the response of axons to molecular gradients. *Proc Natl Acad Sci USA* **106**, 10296–10301, <https://doi.org/10.1073/pnas.0900715106> (2009).
- Paninski, L. *et al.* A new look at state-space models for neural data. *J Comput Neurosci* **29**, 107–126, <https://doi.org/10.1007/s10827-009-0179-x> (2010).
- Babtie, A. C. & Stumpf, M. P. H. How to deal with parameters for whole-cell modelling. *J R Soc Interface* **14**, <https://doi.org/10.1098/rsif.2017.0237> (2017).
- Basso, K. *et al.* Reverse engineering of regulatory networks in human B cells. *Nat Genet* **37**, 382–390, <https://doi.org/10.1038/ng1532> (2005).
- Santra, T. A Bayesian framework that integrates heterogeneous data for inferring gene regulatory networks. *Front Bioeng Biotechnol* **2**, 13, <https://doi.org/10.3389/fbioe.2014.00013> (2014).
- Engelhardt, B., Kschischo, M. & Frohlich, H. A Bayesian approach to estimating hidden variables as well as missing and wrong molecular interactions in ordinary differential equation-based mathematical models. *J R Soc Interface* **14**, <https://doi.org/10.1098/rsif.2017.0332> (2017).
- MacKay, D. J. Bayesian interpolation. *Neural computation* **4**, 415–447 (1992).
- Toni, T., Welch, D., Strelkowa, N., Ipsen, A. & Stumpf, M. P. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J R Soc Interface* **6**, 187–202 (2009).
- Duran, C., Thompson, C. H., Xiao, Q. & Hartzell, H. C. Chloride channels: often enigmatic, rarely predictable. *Annu Rev Physiol* **72**, 95–121, <https://doi.org/10.1146/annurev-physiol-021909-135811> (2010).
- Boudes, M. & Scamps, F. Calcium-activated chloride current expression in axotomized sensory neurons: what for? *Frontiers in molecular neuroscience* **5**, 35, <https://doi.org/10.3389/fnmol.2012.00035> (2012).
- Mikami, M. & Yang, J. Short hairpin RNA-mediated selective knockdown of NaV1.8 tetrodotoxin-resistant voltage-gated sodium channel in dorsal root ganglion neurons. *Anesthesiology* **103**, 828–836 (2005).
- Hudmon, A. *et al.* Phosphorylation of sodium channel Na(v)1.8 by p38 mitogen-activated protein kinase increases current density in dorsal root ganglion neurons. *J Neurosci* **28**, 3190–3201, <https://doi.org/10.1523/JNEUROSCI.4403-07.2008> (2008).
- Hodgkin, A. L. & Huxley, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* **117**, 500–544 (1952).
- Zheng, J. Q., Wan, J. J. & Poo, M. M. Essential role of filopodia in chemotropic turning of nerve growth cone induced by a glutamate gradient. *J Neurosci* **16**, 1140–1149 (1996).
- Yao, J., Sasaki, Y., Wen, Z., Bassell, G. J. & Zheng, J. Q. An essential role for beta-actin mRNA localization and translation in Ca<sup>2+</sup>-dependent growth cone guidance. *Nat Neurosci* **9**, 1265–1273, <https://doi.org/10.1038/nn1773> (2006).

38. Homann, U. & Thiel, G. The number of K(+) channels in the plasma membrane of guard cell protoplasts changes in parallel with the surface area. *Proc Natl Acad Sci USA* **99**, 10215–10220 (2002).
39. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings in the International Joint Conference on Artificial Intelligence* **14**, 1137–1145 (1995).
40. Wang, G. X. & Poo, M. M. Requirement of TRPC channels in netrin-1-induced chemotropic turning of nerve growth cones. *Nature* **434**, 898–904, <https://doi.org/10.1038/nature03478> (2005).
41. Daley, R. *Atmospheric Data Analysis*. 457 (Cambridge University Press, 1991).
42. Akaike, H. A new look at the statistical model identification. *IEEE transactions on automatic control* **19**, 716–723 (1974).
43. Schwarz, G. Estimating the dimension of a model. *The annals of statistics* **6**, 461–464 (1978).
44. Li, Z., Zhang, G., Feil, R., Han, J. & Du, X. Sequential activation of p38 and ERK pathways by cGMP-dependent protein kinase leading to activation of the platelet integrin alphaIIb beta3. *Blood* **107**, 965–972, <https://doi.org/10.1182/blood-2005-03-1308> (2006).
45. Keely, S. J. & Barrett, K. E. p38 mitogen-activated protein kinase inhibits calcium-dependent chloride secretion in T84 colonic epithelial cells. *Am J Physiol Cell Physiol* **284**, C339–348, <https://doi.org/10.1152/ajpcell.00144.2002> (2003).
46. Lisman, J. A mechanism for the Hebb and the anti-Hebb processes underlying learning and memory. *Proc Natl Acad Sci USA* **86**, 9574–9578 (1989).
47. Kaupp, U. B. & Seifert, R. Cyclic nucleotide-gated ion channels. *Physiol Rev* **82**, 769–824, <https://doi.org/10.1152/physrev.00008.2002> (2002).

## Acknowledgements

We thank T. Araki, K. Watanabe, T. Takenouchi, S. Takeda, T. Nomura, R. Yoshida, S. Koyama, K. Kunida, and Warren Jelinek for helpful discussions and comments. This research was supported in part by NIH CRCNS, MEXT and JSPS KAKENHI (25127713), and NAIST Interdisciplinary Frontier Research Project.

## Author Contributions

M.N., K.H., and Y.S. designed research; T.Y. performed research; M.N. and K.H. contributed experimental data; T.Y., S.O., H.C.J., K.I., S.I., and Y.S. analyzed data; and T.Y., M.N., S.I., K.H., and Y.S. wrote the paper.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-22506-3>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018