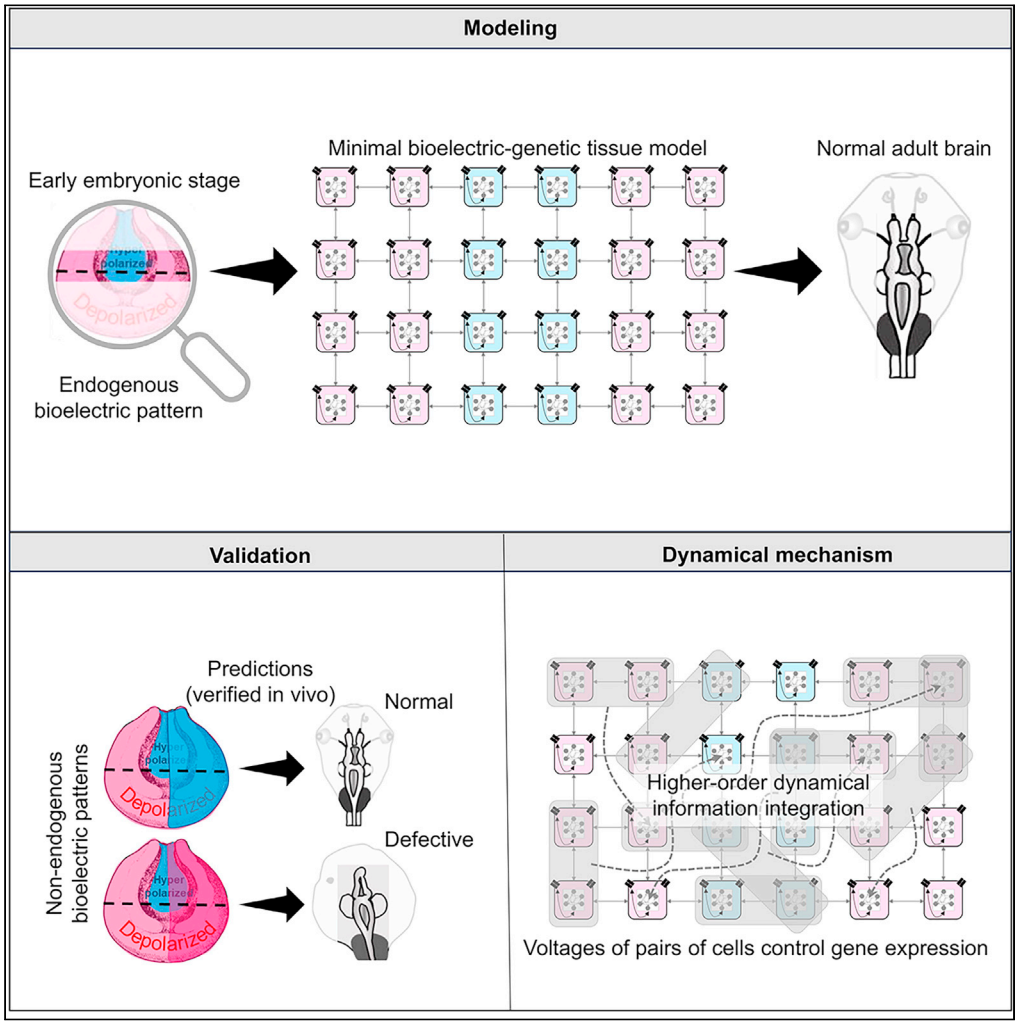
iScience

CellPress
OPEN ACCESS

Article

Information integration during bioelectric regulation of morphogenesis of the embryonic frog brain



Santosh Manicka, Vaibhav P. Pai, Michael Levin

michael.levin@tufts.edu

Highlights

A minimal model shows how cells can sense large-scale patterns of cell voltage

Model makes predictions about the outcome of new bioelectric patterns

Model predictions are verified by experiments in Xenopus brain development

Higher-order information integration is seen in voltage-transcription dynamics

Manicka et al., iScience 26, 108398
December 15, 2023
https://doi.org/10.1016/j.isci.2023.108398

iScience

CellPress
OPEN ACCESS

Article

Information integration during bioelectric regulation of morphogenesis of the embryonic frog brain

Santosh Manicka,[1] Vaibhav P. Pai,[1] and Michael Levin[1,2,3,*]

SUMMARY

Spatiotemporal patterns of cellular resting potential regulate several aspects of development. One key aspect of the bioelectric code is that transcriptional and morphogenetic states are determined not by local, single-cell, voltage levels but by specific distributions of voltage across cell sheets. We constructed and analyzed a minimal dynamical model of collective gene expression in cells based on inputs of multicellular voltage patterns. Causal integration analysis revealed a higher-order mechanism by which information about the voltage pattern was spatiotemporally integrated into gene activity, as well as a division of labor among and between the bioelectric and genetic components. We tested and confirmed predictions of this model in a system in which bioelectric control of morphogenesis regulates gene expression and organogenesis: the embryonic brain of the frog *Xenopus laevis*. This study demonstrates that machine learning and computational integration approaches can advance our understanding of the information-processing underlying morphogenetic decision-making, with a potential for other applications in developmental biology and regenerative medicine.

INTRODUCTION

Embryonic development is a remarkable process, in which a large number of cells cooperate toward invariant large-scale anatomical shapes.[1,2] Central to this phenomenon (in most organisms) is long-range order: not only local identity for each cell in a hardwired mosaic but high levels of tolerance to noise, ability to handle perturbations, and active cell-cell interactions that maintain growth and form.[1,3,4] Robust mechanisms ensure that overall proportions among organs, symmetry, and topological and geometrical relationships are maintained across the whole organism.[5–8] Of particular interest are the context-sensitive responses that enable transcriptional events based on large-scale (non-cell-autonomous) information that signals completion of morphogenetic processes or errors in tissue-level order.[2,9] However, while cell-level mechanisms linking signals to cell fate are beginning to be well understood, the rate-limiting step for many approaches in regenerative medicine and morphogenic bioengineering is gaining control over tissue- and organ-level collective dynamics.

Morphogenesis is orchestrated by several physical modalities that enable cells to coordinate their actions toward specific anatomical outcomes. In addition to biochemical gradients[10–15] and biomechanical forces,[16–18] it is now evident that bioelectric networks[19–21] (spatiotemporal patterns of membrane voltage across cell fields) produce spatial gradient information that is instructive for growth and form in a number of well-studied contexts. For example, endogenous bioelectric prepatterns have been functionally implicated in the location and size of organs such as the eye,[22] invertebrate wing,[23–26] vertebrate appendage,[27–29] and face.[29–31] Bioelectric prepatterns also polarize the left-right[32–35] and dorsoventral[36,37] axes and are involved in size control,[38–42] embryonic compartmentalization,[43–47] and stem cell function.[48–50] Transduction mechanisms for bioelectric signals,[51] downstream transcriptional machinery,[52] and interfaces between electrical and mechanical events[53,54] have begun to be characterized. The key role of bioelectric signals and their potential as therapeutic targets is underscored by computational models that have now come online,[55–60] which enable rational manipulation of ion channel interfaces (via drugs, light, or mRNA misexpression) to induce ectopic organ formation,[22,61] trigger complex regenerative response,[62–64] and reverse tumorigenesis[65,66] in model systems *in vivo*. Some of these modulators are beginning to be applied to human medicine as electroceuticals.[67–76]

A number of mechanisms for converting membrane voltage changes in an individual cell into changes in that cell's behaviors have been characterized as cell-autonomous transduction mechanisms.[51] However, it remains unclear how *tissue*-level bioelectric information (distribution of cell resting potentials across cell fields) is processed to achieve large-scale morphological outcomes. For example, specific distributions of membrane voltage pattern across cells indicate the location, size, and organ identity of hearts, eyes, wings, and other structures.[22–24,77–79] This requires downstream transcriptional cascades to be triggered by specific multicellular patterns of voltage, as has been shown for Sox2, Pax6, MSX-1, Notch, Wnt/β-catenin, and BMP,[64,80,81] as well as a range of other transcripts identified in transcriptomic

[1]Allen Discovery Center at Tufts University, Medford, MA 02155, USA
[2]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA
[3]Lead contact
*Correspondence: michael.levin@tufts.edu
https://doi.org/10.1016/j.isci.2023.108398

analyses of voltage-modified tissues.[52] Because complex organogenesis and regeneration cannot be solved by purely local information at the single-cell level, it is essential to understand how the regional spatial pattern of voltages across a *group* of cells is converted into the diverse patterns of gene expression required for normal morphology.[1,9] Thus, models of developmental bioelectricity must move to considering the more complex and interesting systems-level question of how a collective of cells processes the information encoded in bioelectrical patterns across distance to orchestrate a differentiated pattern of transcriptional responses.[57,59,82,83]

The relationships between multicellular bioelectric patterns and downstream transcriptional cascades are not easily intuited, and the parameter space for experimentally testing all possibilities is vast. Especially, there is a need for rigorous models that show how composite transcriptional-bioelectric circuits can recognize correct and incorrect patterns during regulative development; generative models of this type serve as crucial hypotheses about the dynamics that are sufficient to perform the task and can make testable predictions. We therefore developed a framework in which machine learning could be deployed to search for empirically testable, minimal, predictive models that integrate simplified bioelectric pattern and transcriptional changes (Figures 1 and 2). We then applied this methodology to the specific instance of modeling bioelectric control of *Xenopus* embryonic brain development,[84–86] where such a relation between bioelectric pattern and transcriptional changes has been experimentally identified.[80,87,88]

Previous modeling work in this area includes multicellular genetic networks that either do not comprise any bioelectric components,[89,90] are highly detailed multicellular bioelectric-gene networks containing several biophysical components that render them too computational complex for using machine learning for parameter space exploration,[56] or are simple bioelectric-gene networks possessing only one gene per cell.[91] Moreover, prior work on predictive modeling did not analyze the information-processing dynamics that the models employ to map voltage patterns to transcriptional patterns—an issue more broadly relevant to the evolution of patterning mechanisms. Here, we adopted a bioelectric-genetic network modeling approach that shows the minimal mechanisms and relationships sufficient to accomplish a specific morphogenetic task. Our bioelectric networks are modeled with just two generic ion channels, whereas the genetic networks comprise several genes and intercellular and intracellular interconnections (Figures 1F and 2A). Importantly, this enabled us to analyze the information-processing mechanisms that bridge the dynamical gap between bioelectric patterns into collective gene expression.

We considered the nascent *Xenopus laevis* brain, because it is a complex organ commonly used to model disorders with biomedical relevance[84,85,92–96] and because its normal morphogenesis has already been shown to include a bioelectric component.[22,80,97,98] In this sytem, a complex interplay between secreted factors, morphogens, and transcription factors leads to neural induction and formation of the neural plate.[15] Briefly, secreted proteins from the mesoderm block the action of morphogens (particularly BMPs and Wnts) in the adjacent ectodermal cells. Blocking of BMPs and Wnts along with FGF activation leads to the induction of neural transcription factors (such as Fox, Sox, Zic, and Irx families) that make these ectodermal cells refractory to morphogens and drive neural induction and neural plate formation.[12,15,99–104] For the bioelectric pattern, beginning at embryonic stage 16, a dramatic contrast develops between the transmembrane voltage potential at the neural plate (hyperpolarized) and in the surrounding ectoderm (depolarized), establishing a bioelectric pre-pattern that regulates subsequent large-scale brain development[80,97] (Figures 1A and 1B). Crucially, it is the long-range difference between voltage in these regions that is required for normal development, not the absolute values of specific regions. This bioelectric contrast pattern regulates expression of canonical brain development transcription factors (e.g., *Notch*, *Otx2*, *Pax6*, *Emx*, *Xbf1/Foxg1*, *Sox2*) as well as neural progenitor cell proliferation and apoptosis, driving normal development of the brain[80,87,97] (Figures 1A and 1B). Disrupting this normal bioelectric contrast pattern disrupts the differential expression of these transcription factors, resulting in abnormal large-scale brain morphology and function[80,97] (Figures 1C and 1D). In addition, neuroteratogens such as nicotine and ethanol and genetic or biochemical defects such as a disrupted notch signal cause severe brain morphology defects and impaired function by disrupting this bioelectric contrast pattern[87,88,105,106] (Figures 1C and 1D). Remarkably, local and long-range interventions that restore the normal bioelectric pattern in the presence of these disrupters rescue large-scale brain morphology and function, showing the incredible regulatory importance of this bioelectric pattern in embryonic brain development[80,87,88,105,106] (Figure 1E). Conversely, interfering with normal expression of Notch or Pax6 also disrupts the bioelectric contrast pattern, demonstrating feedback from gene expression to the large-scale bioelectric contrast pattern, at least for these two genes.[22,80,105] Furthermore, recreating the bioelectric contrast pattern in other regions of the embryo is sufficient to re-specify those regions to produce ectopic gene expression and in turn ectopic brain tissues, even outside of the head region, demonstrating the instructive functional role of this large-scale voltage pattern for brain development.[80]

Although the bioelectric pattern that regulates brain morphology in this system is known, it is not understood how these *multicellular* patterns could be collectively interpreted and executed across the field of cells in the neural plate and the surrounding ectoderm. Here, we sought to answer a general question: how are such multicellular tissue-level bioelectric patterns converted into specific genetic expression patterns in cells, leading to correct large-scale brain morphology? This is a question that affects numerous contexts in addition to the frog brain (reviewed in[107]), and our model is focused on that part of the process: from large-scale bioelectric pattern to a transcriptional output (Figures 1F and 1G). This approach does not depend on the specific identity of the responding genes or their relationship to other downstream genes, but our model is consistent with the known molecular details of neural tube patterning in frog.

Our goal is to take the first step toward exploring the high-level information-processing aspects of how multicellular bioelectric patterns could lead to specific gene expression changes in cells that then lead to large-scale morphogenesis. Our model is developed in the context of experimental observations in Xenopus embryonic development that changes in bioelectric pattern spread across the neural plate result in specific changes in neural patterning transcription factor genes such as Otx2, Pax6, Foxg1, Emx1, Sox2, and Notch. We are not seeking to simulate exact molecular level details (which is impossible, given the current state of knowledge) but to reveal generic, yet sufficiently biorealistic, dynamics that could account for this phenomenon in a range of contexts. Therefore, we adopted a modeling approach that

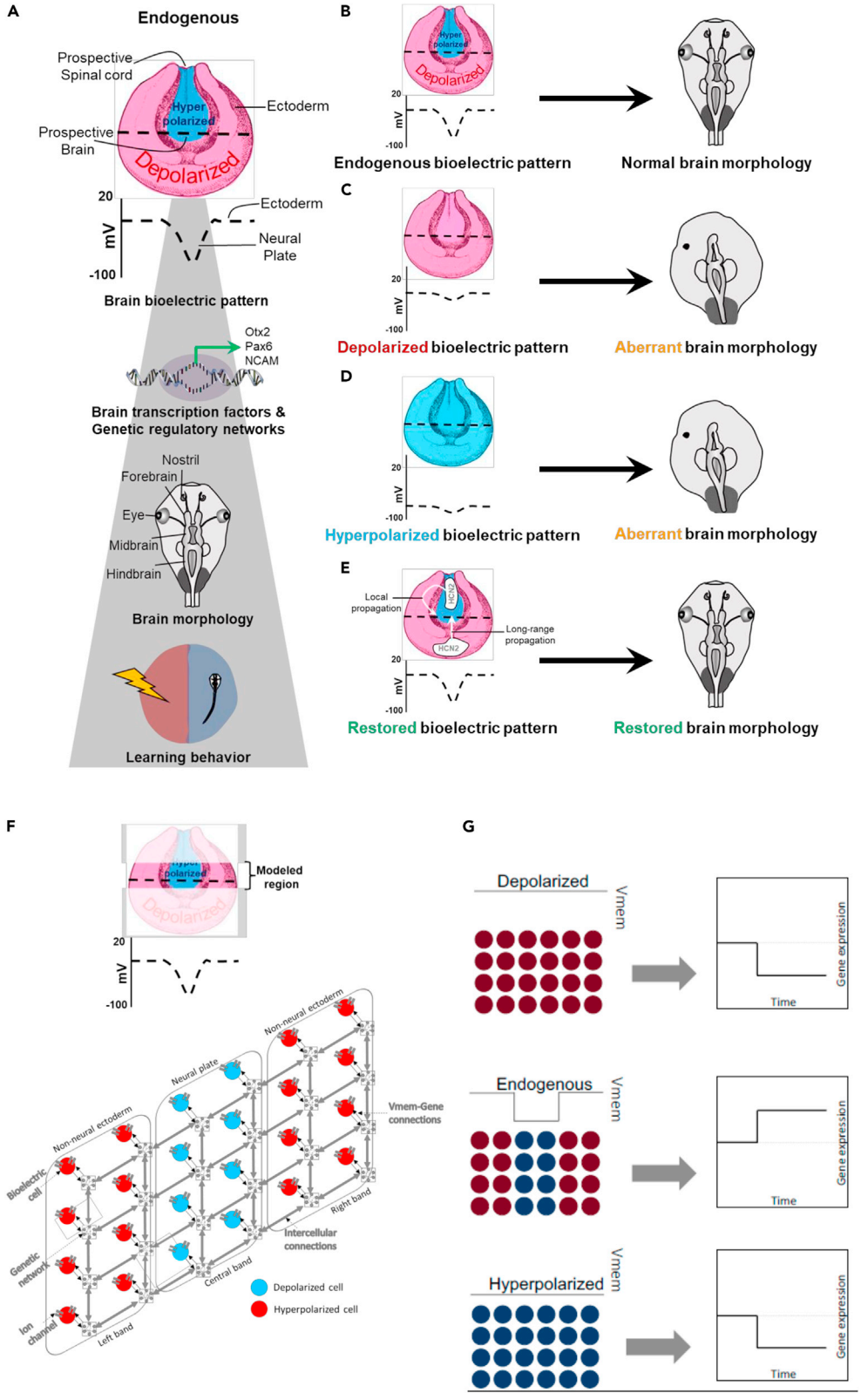

A
Endogenous
Prospective Spinal cord
Prospective Brain
Hyper polarized
Depolarized
Ectoderm
Ectoderm
Neural Plate
20
mV
-100
Brain bioelectric pattern
Otx2
Pax6
NCAM
Brain transcription factors & Genetic regulatory networks
Nostril
Forebrain
Eye
Midbrain
Hindbrain
Brain morphology
Learning behavior
B
Endogenous bioelectric pattern
Normal brain morphology
C
Depolarized bioelectric pattern
Aberrant brain morphology
D
Hyperpolarized bioelectric pattern
Aberrant brain morphology
E
Local propagation
Long-range propagation
HCN2
Restored bioelectric pattern
Restored brain morphology
F
Modeled region
Non-neural ectoderm
Neural plate
Vmem-Gene connections
Bioelectric cell
Genetic network
Ion channel
Intercellular connections
Right band
Central band
Left band
Depolarized cell
Hyperpolarized cell
G
Depolarized
Vmem
Gene expression
Time
Endogenous
Hyperpolarized

Figure 1. Schematic summary of the endogenous voltage prepattern's control of embryonic brain patterning in *Xenopus* and an overview of the neural plate circuit model of this process

(A) Summary of our previous studies[80,87,88,97] showing that the *difference* (dashed black line) in voltage across the hyperpolarized neural plate and relatively depolarized surrounding ectoderm is essential for driving correct gene expression pattern, large-scale brain morphology, and normal learning behavior.

(B) Endogenous embryonic voltage prepattern resulting in normal brain morphology.[80,87,88,97]

(C) Embryonic manipulation of **ion** fluxes, exposure to teratogens (nicotine and ethanol), or genetic/biochemical disruption (Notch disruption) leading to depolarization of the neural plate erases the critical voltage difference between neural plate and ectoderm, resulting in mispatterned gene expression, and defects in large-scale brain morphology and learning behavior.[80,87,88,97,105,106]

(D) Embryonic manipulation of ion fluxes leading to hyperpolarization of surrounding ectoderm also erases the critical voltage difference between neural plate and ectoderm, resulting in mispatterned gene expression and abnormal large-scale brain morphology.[80,97]

(E) Both local (neural plate) and distant (ectodermal) interventions (manipulating ion fluxes by channel misexpression or drugs targeting ion channels) that restore this critical voltage difference between neural plate and ectoderm, restore gene expression, large-scale brain morphology, and learning behavior.[80,87,88,97,105,106]

(F) A schematic of our model. The overt structure of the model is a 2-dimensional lattice approximately representing the neural plate tissue (shown in the illustration). The circles indicate the individual cells of the tissue, with the colors indicating different voltages corresponding to the endogenous voltage prepattern (blue means depolarized and red, hyperpolarized). Each cell possesses a generic hyperpolarizing channel and depolarizing channel (not shown) and a generic gene regulatory network (gray networks enclosed in squares). The cells exchange gene products and proteins with neighboring cells through the IN-generic intercellular connections (thick gray lines with bidirectional arrowheads) representing paracrine signaling pathways. These collectively represent a plethora of paracrine signaling pathways such as calcium, cAMP, serotonin, integrin, Notch, etc. In this current first iteration model, we generalize them all under paracrine signals as the model is consistent with them all and does not depend on their precise nature. Of note, the IN does not directly allow the exchange of ions but it does so only indirectly via the voltage→gene→gene→voltage pathway (Figure 2A). The voltage and the gene network of a cell influence each other via dedicated connections (thin black lines with unidirectional arrowheads): the voltage acts as an external input to a subset of the genes, affecting their expression and, in turn, the expressions of a subset of the genes can dynamically alter the conductance of the ion channels. The cells in the model are symmetrical, that is, they do not possess a preferred polarity such as that imposed by planar cell polarity. The 4x6 dimension of the lattice shown here is the size that was used during training.

(G) Schematic of the input-output map of the pattern discrimination problem; the three input patterns on the left correspond to the embryo schematics displayed in C, B, and D, respectively, with the depolarized (red) and hyperpolarized (blue) voltages corresponding to about −5.2 mV (−10 mV) and −52 mV (−55 mV) in the model (empirical observations[80]).

combines the predictive power of generic artificial neural networks and the explanatory power of specific biophysical models. In this model, the inputs (bioelectric patterns) are mapped to the outputs (gene expression patterns representing overt brain development outcomes) via generic mechanisms designed by machine learning; specific biological features are also incorporated as model constraints (Figures 1F and 2A). This modeling strategy is inspired by the so-called biologically informed and physics-informed neural networks,[108,109] which are conventional artificial neural networks subject to additional physical and biological constraints. As we sought to understand the interplay between the bioelectrical and gene-regulatory variables, our model assumes simplified bioelectric and gene regulatory mechanisms and their interactions in a static fixed network (cell division and migration are not explicitly modeled). The model also assumes generic features that embryos generally share, such as topographic and symmetrical intercellular connectivity, a common gene regulatory network that all cells share, etc.

We adopted an approach that is intermediate between "forward" and "reverse" modeling—by starting from known biological features (Table 1) and leaving the unknown properties to automatic reverse engineering.[110,111] While the overall transcriptional outcomes of correct and incorrect bioelectrical patterns in the neural plate have been characterized,[88,97,105,112] detailed information about the regulatory networks comprising relevant genes and the regulatory connections between voltage and gene expression is limited, if any.[113–116] Thus, we sought to analyze a minimal, yet biologically realistic, model that would reveal the dynamics sufficient to observe the known phenomena of biophysical genetic regulation of outcomes in the nascent tadpole brain. Adopting a hybrid approach also enabled us to leverage the benefits of both by facilitating the exploration of high-level information-processing mechanisms afforded by the forward approach as well as specific quantitative predictions enabled by the reverse approach.[110]

As a first step to address the question of how subsequent gene expression of cells distinguishes correct versus incorrect voltage distributions across cell fields, we elucidated the information-processing mechanism relating the bioelectric and gene expression patterns of a high-performing model that machine learning discovered (Figure 2B) from among a suite of 40 initially randomly parameterized models. We focused our analysis on a single model, as we sought to demonstrate and analyze an example of sufficient design principles rather than a broad study of necessary conditions. While no model can formally prove particular mechanisms to be the cause of empirical observations, models that show dynamics sufficient[117] to reproduce the dynamics observed in a biological system are essential to begin the continuous dialogue between modeling and experimentation that leads to improved functional understanding and control of complex system-level biology.[111,117] Our model was formulated with the goal of continued improvement, as specific data come in from future testing of its implications.

Our analysis shows that the mechanism instantiated in our class of models not only exhibits invariance to the scale of the modeled tissue but also a simplified higher-order spatiotemporal information integration mechanism that maps the spatial bioelectric pattern to the gene expression pattern. The tissue exhibited functional specialization, with cells at the voltage inflection points of the bioelectric pattern being important for correct pattern recognition and collective cellular decision-making. The genes within the individual cells were responsive to the voltages of faraway cells over longer periods of time, indicating their roles in integrating information across the tissue. This character further manifested as a division of labor among the genes in regard to their differential sensitivities to the various spatial scales of the bioelectric pattern. Our analysis led to predictions concerning the previously unknown role of the polarization ratio of the tissue in the determination

Table 1. Model features and their biological analogs

Model feature	Biological analog
(a) Two-dimensional plane topology	Approximates the somewhat flat structure of the neural plate
(b) Topographic (lattice) connectivity	Represents connectivity only among neighboring cells characteristic of biological tissues
(c) Bidirectionally connected symmetric intercellular networks (IN)	Represents paracrine cell signaling networks[90,172,173] by which neighboring cells exchange gene products, with the additional assumption that there's no preferred directionality imposed by features such as planar cell polarity
(d) Simplified "equivalent-circuit" bioelectric machinery within each cell	Represents the aggregate bioelectric components of a cell, consisting of a pair of generic hyperpolarizing and depolarizing ion channels
(e) A gene regulatory network (GRN), comprising a set of at least six generic genes, shared by all cells	Represents a regulatory network of generic genes, with the number six derived from the present knowledge that there are at least six genes known to be sensitive to the voltage patterns of the *Xenopus* neural plate, namely *Notch*, *Otx2*, *Pax6*, *Emx*, *Xbf1/Foxg1*, and *Sox2*
(f) Bidirectional connectivity between the cellular voltage and the genes of each cell	Represents the influence of the bioelectric pattern on gene expression, acting as an "external input",[80,87,97] and, in turn, the influence of genes on the maintenance of the endogenous bioelectric pattern by modulating the ion channel conductances.[22,80,105]

of the morphological outcomes that we later confirmed *in vivo*. Our work demonstrates that the bioelectric regulation of neural morphogenesis in *Xenopus* embryos is not a local process but one that involves the integration of information across the tissue (including non-neural cells), and it showcases one way of implementing this dynamical process. Importantly, it demonstrates the possibility of regulating the higher-dimensional genetically centered morphogenetic process via lower-dimensional bioelectric interfaces that leverages the coarse-graining offered by the membrane potential.[118] Taken together, our results demonstrate the tractability of a combined *in silico/in vivo* approach for identifying system-level features of a multi-modal morphogenetic process involving both biophysical and genetic components.

RESULTS

Machine learning discovered a model that solves the voltage pattern recognition problem

To elucidate the conditions *sufficient* to solve the pattern discrimination problem, we chose a single high-performing model for subsequent analysis and examined the evolution of its ability to discriminate between correct and incorrect voltage patterns (see Methods).

The connectivity and associated parameters of the chosen model are depicted in Figure 3, which shows a simplified two-cell version of the full model (Tables S1–S8). The lines connecting the two cells comprise about 20% of the strongest connections: 10% of the most positive and 10% of the most negative. The strongest connections (thick edges in the figure) are (1) those that run from the voltage to a few genes, providing reliable transmission of information about the input voltage to the genes and (2) those connecting a small subset of genes both within a cell and between cells. These observations translate into differences in timescales of voltage and gene expression dynamics; here, voltage changes occur at a relative slower rate compared with gene expression (Figure S14). This makes sense in light of the fact that the voltage pattern serves as the input, whereas gene expression is the output: for the input to be reliably translated into the output, it should change relatively slowly,[119,120] despite the recurrent connectivity between them. This also reflects biological observations, as non-neural embryonic bioelectric signals happen on the timescale of several hours (neural plate bioelectric pattern happens from stage 15 to stage 18[80,121]), which is slower than the regulation of many genes. We next analyze in detail the behavior and dynamics of this model.

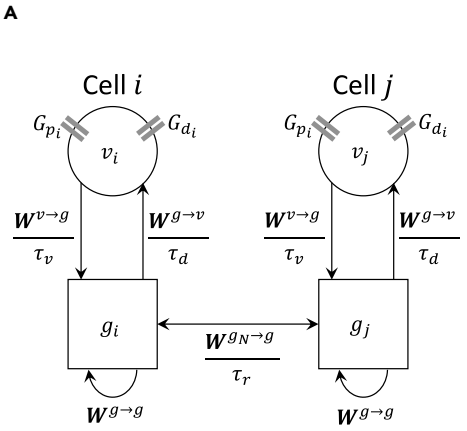
On average, genes in the model successfully discriminate between the correct and incorrect voltage patterns: when averaged over all cells and all genes, expression is increased for the correct voltage pattern and decreased for the incorrect pattern (Figure 4A). The average magnitudes of expression are relatively low, about +0.5 for activation and −0.2 for repression, compared with the ideal values of +1.0 and −1.0

$$C\frac{dv_i}{dt} = -\frac{G_{p_i}(v_i - E_p)}{1 + e^{\frac{z(v_i - v_{th})}{v_T}}} - \frac{G_{d_i}(v_i - E_d)}{1 + e^{-\frac{z(v_i - v_{th})}{v_T}}}$$

$$\frac{dG_{d_i}}{dt} = -G_{d_i} + \frac{\boldsymbol{W}^{g\to v}}{\tau_d}\sigma(\boldsymbol{g_i} + \boldsymbol{b_g})$$

$$\frac{d\boldsymbol{g_i}}{dt} = \frac{1}{\boldsymbol{\tau_g}}\left[-\boldsymbol{g_i} + \boldsymbol{W}^{g\to g}\,\sigma(\boldsymbol{g_i} + \boldsymbol{b_g}) + \frac{\boldsymbol{W}^{g_N\to g}}{\tau_r}\sum_{j\in N_i}\sigma(\boldsymbol{g_j} + \boldsymbol{b_g}) + \frac{\boldsymbol{W}^{v\to g}}{\tau_v}\sigma(e^u v_i + b_v)\right]$$

Where:
n_c = number of cells
n_g = number of genes in a cell
N_i = the set of neighboring cells of cell i
v_i = Vmem of cell i
G_{p_i} = the conductance of the polarizing ion channel of cell i
G_{d_i} = the conductance of the depolarizing ion channel of cell i
C = the capacitance of a cell
$\boldsymbol{g_i} = (g_{i,1}, g_{i,2}, \ldots, g_{i,n_g})$ is the vector of the gene expressions in cell i, with the corresponding vector of timeconstants $\boldsymbol{\tau_g}$
$\boldsymbol{W}^{g\to v}$ = 1 x n_g matrix of weights of the connections flowing from the genes to the Vmem, with timeconstant τ_d
$\boldsymbol{W}^{g\to g}$ = n_g x n_g asymmetric matrix of weights of the connections flowing among the genes within a cell (GRN)
$\boldsymbol{W}^{g_N\to g}$ = n_g x n_g asymmetric matrix of weights of the connections flowing among the genes between neighboring cells (iGRN), with timeconstant τ_r
$\boldsymbol{W}^{v\to g}$ = n_g x 1 matrix of weights of the connections flowing from the genes to the Vmem, with timeconstant τ_v
$\boldsymbol{b_g}$ = n_g x 1 vector of the biases of gene expression
b_v = bias of the Vmem
u = Vmem gain
$\sigma(.)$ = standard sigmoid transformation



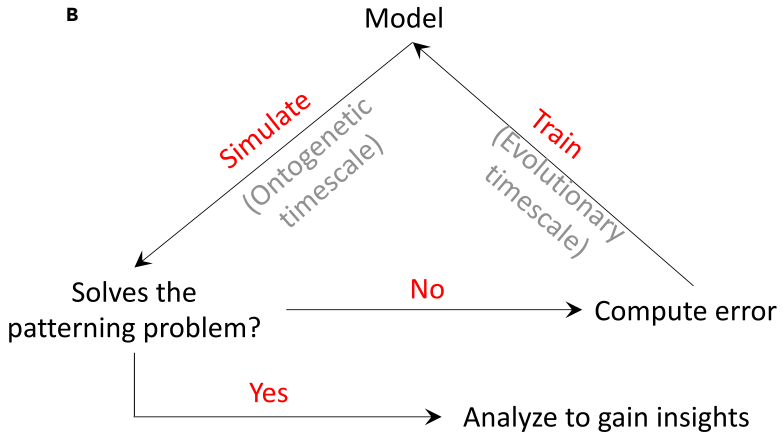


Figure 2. Overview of the mathematical details of the neural plate circuit model and the process of identifying its parameters (training)

(A) The equations defining the model, with the variables and parameters mapped to a simplified two-cell version (the rest of the model is symmetrical to this version). The parameters (red) are learned using machine-learning techniques. The bioelectric constants, namely, C, v_{th}, v_0, and v_T, are described in ref.[88] and in Tables S1–S8.

(B) The model is trained using a combination of genetic algorithm and gradient descent. The final model referred to throughout this paper is a product of a training loop that starts with a population of randomly parametrized models, simulates all of them, and then modifies them using genetic algorithm and gradient descent to improve their ability to solve the pattern discrimination problem.

respectively. These magnitudes yield a performance score of 0.64, significantly higher than the 0.5 score expected for a randomly parameterized model; for comparison, the minimum observed score was 0.5 and the 95th percentile score was about 0.58 (Figure S1). Importantly, the average changes in gene expression in response to the correct and incorrect voltage patterns clearly diverge over time. When we look at individual genes, this separation is largest for gene 6, as indicated by the corresponding discrimination score (Figure 4A; bottom right inset). For this reason, we term gene 6 the "discriminator gene" and focus on its behavior in the analyses below.

Interestingly, the temporal dynamics of the voltage pattern in this model also recapitulate the temporal evolution of the bioelectric pattern *in vivo* during development even though the model was not specifically trained to develop it (solid black lines in Figure 4B). At the beginning, our model exhibits a flat depolarized pattern across all cells; then, fairly early in the simulation, the cells in the center (modeling the neural plate) hyperpolarize to produce the characteristic voltage contrast seen *in vivo*, before reverting to the flat depolarized pattern about halfway through the simulation. This corresponds to what we see in this region between developmental stages 14 and 20 *in vivo*.[80] These observations suggested that even though the quantitative performance of the model is not ideal, its quality warranted further analysis.

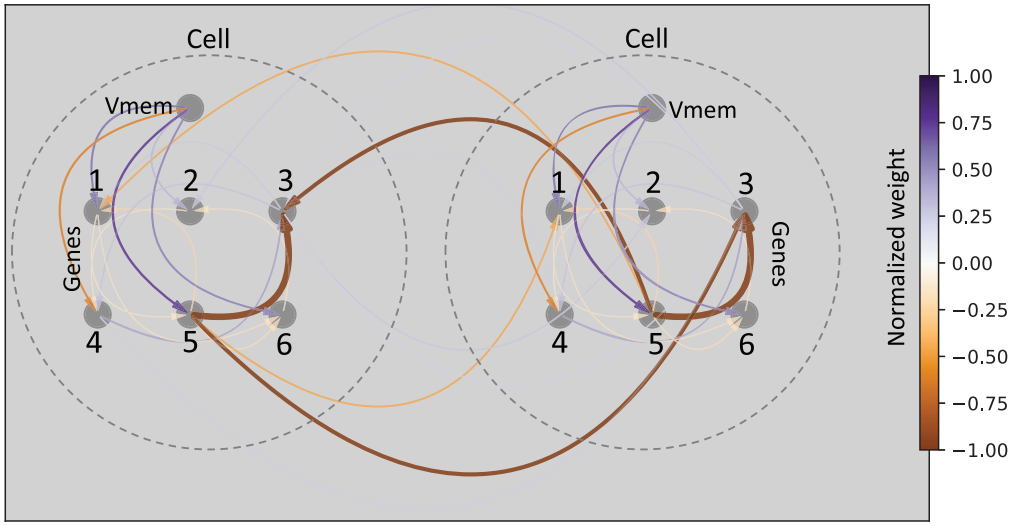


Figure 3. Architecture of the chosen high-performing model

Depicted here is the connectivity of a two-cell version of the full model; the rest of the network is symmetrical to this version (full model parameters described in Tables S1–S8). For clarity, we show only about 20% of the strongest connections (10% of the most positive in purple and 10% of the most negative in orange/brown); edge weights are linearly normalized to the range (−1, 1). The most prominent connections run from the voltage to gene 5 (purple), with a positive weight, and among genes 3 and 5 both within a cell and between cells (brown), with a negative weight. Connections running from genes to voltage do exist in the full model, but they were filtered out due to their weak weights.

Emergent scaling of the chosen model: The pattern discrimination problem was solved in larger tissues despite lack of explicit selection for such capability

One central aspect of biological control mechanisms is that they often have properties that were not specifically selected for, by virtue of generic (inherent) properties of networks.[8,122,123] We were interested in uncovering any novel emergent features in our model, focusing especially on size control. The ability to establish correct overall pattern despite differences in tissue size or available cell number,[124–128] a key competency of developmental and regenerative morphogenesis, is still poorly understood in models based on reaction-diffusion networks[56,129,130] due to their limited scaling robustness.

Interestingly, our model successfully solved the pattern discrimination problem even in a larger tissue containing 180 cells (compared with the original 24 cells), even though it was not specifically trained for that purpose (Figure 5). In other words, the qualitative behavior of the discriminator gene is preserved, where the activity in the flanking columns is relatively lower compared with the columns closer to the center in both models, suggesting that the dynamics of the original model have scaled up (Figure S3A). The model was found to scale to tissues as large as about 400 cells but the quality of scaling drops for much larger tissues where the genes in the flanking columns deactivate since information about the pattern does not propagate to the margins of the tissue in time (Figure S3B). The Xenopus neural plate can be roughly estimated to contain 43 X 43 cells. Thus, the bioelectric pattern which contains neural plate cells and surrounding ectoderm cells can be estimated to be ~83 cells wide,[131,132] which is squarely within the scaling range of our model. This suggests that the pattern can scale in a similar manner to a tissue that is larger or smaller than the *Xenopus* neural plate tissue, within certain limits.

High-level information-processing mechanisms employed to solve the pattern discrimination problem

We next performed a detailed analysis of the model to decipher some of the high-level information-processing mechanisms and organizational principles it employs to solve the pattern discrimination problem. In the following, we specifically asked the following: (1) Does functional specialization exist among the components (cells, genes, modules, etc.)? (2) Does a division of labor exist in terms of the scales (distributed or local) at which information-processing is performed? (3) Are there high-level relationships among the components that facilitate pattern discrimination? (4) How is the voltage pattern processed in the tissue? All of these are precursors to a mature understanding of collective, multimodal patterning mechanisms necessary for progress in basic evolutionary developmental biology and regenerative medicine and morphogenetic engineering.

Cells surrounding voltage pattern transition points are more important than other cells

To determine whether different cells in the tissue play different roles in the collective recognition of the endogenous voltage pattern, we performed an *in silico* cell "knockout" analysis. Specific cells were frozen by clamping the states of the corresponding genes at zero and dropping their biases to negative infinity, thereby rendering them ineffectual both as input and output, and the performance of the model was then computed.

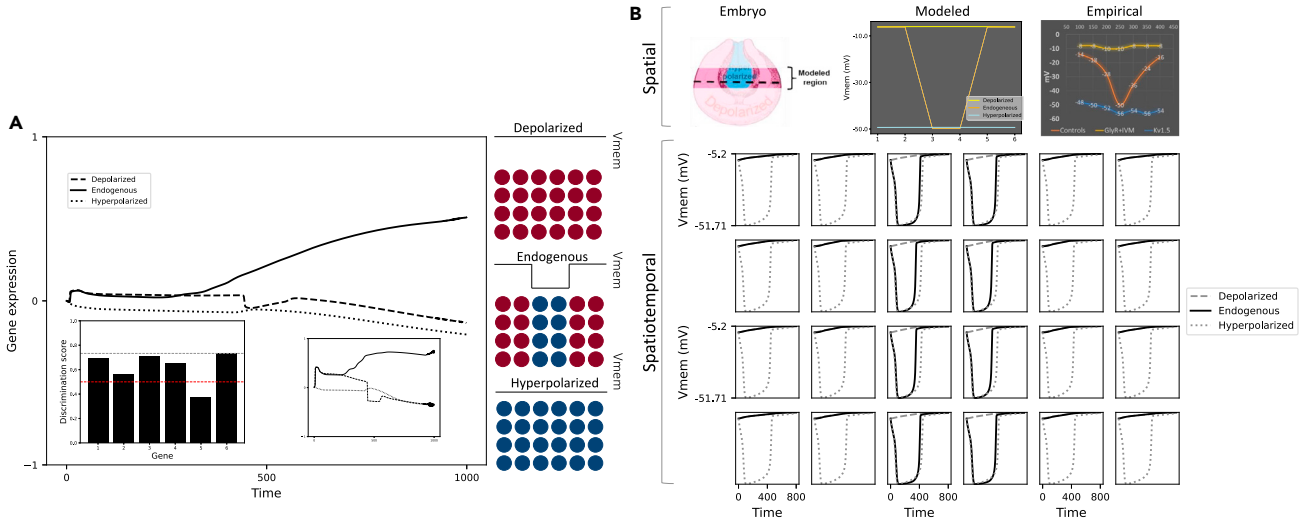


Figure 4. The model not only discriminates between the correct and the incorrect voltage patterns by mapping them to distinct gene expression values but also recapitulates empirically observed development of the voltage pattern

(A) Timeseries of the mean normalized gene expression, averaged over all cells and genes, for the three different input conditions depicted on the right: "Depolarized" (all cells are initially depolarized), "Endogenous" (the left and right two columns are initially depolarized with the middle two columns hyperpolarized), or "Hyperpolarized" (all cells are initially hyperpolarized). Insets: (left) discrimination scores of the six individual genes. The red dashed horizontal line indicates the discrimination score expected from a randomly parametrized model. (Right) Gene expression timeseries for gene 6, the gene with the highest discrimination score; the scales of the axes match the scales of the main plot. More detailed timeseries resolved at the level of cells and genes are presented in Figure S2.

(B) Top: comparison of the modeled and empirical spatial voltage profiles of a horizontal cross-section of the embryo (schematic) at a point in time when the voltages corresponding to the three different input conditions (treatments) have fully developed (~200th timestep in the model). Bottom: the spatiotemporal voltage profile of the model depicting the timeseries of Vmem for each cell under each of the three input conditions. Each square in the 4x6 layout of the tissue refers to an individual cell.

Knocking out the cells surrounding the voltage transition points of the endogenous pattern—at the boundaries of the central band where there are large changes in the spatial profile of the voltage pattern—resulted in a greater drop in performance (Figure 6). In other words, cells closer to the voltage transition points are more important in recognizing the pattern compared with the cells further away. Specifically, the four cells at the center of the 4x6 tissue are the most significant contributors to performance (Figure 6A). By virtue of scale robustness, this property also extends to the larger 10x18 tissue where small 2x3 patches of cells were knocked out. There again, the cells flanking the center of the tissue are the most important (Figure 6B). Although there are slight differences between the smaller and the larger tissues with regard to the exact spatial regions that are most important, the general principle holds: cells coinciding with changes in the pattern contribute more to its discrimination. A potential explanation for these observations is that the voltage transition points comprise the only difference between the endogenous pattern and the uniform patterns, suggesting that this strategy may computationally optimal.

A simplified higher-order control mechanism characterizes the relationship between the endogenous voltage pattern and gene expression

To characterize the high-level relationships between the endogenous voltage pattern and tissue-level gene expression, we attempted to reconstruct the dynamic behavior of the discriminator gene in terms of its dynamic sensitivities to voltage patterns captured in the Jacobian and Hessian tensors (Figure 7; Equations 3 and 4); reconstruction of the activities of the other genes are shown in Figure S9. Due to the 4-fold reflection symmetry of the 2D lattice-structured tissue (Figure 1F), we computed the sum of sensitivities of the discriminator gene, corresponding to the endogenous input pattern, only for the cells in the top left quadrant of the tissue (mathematical details in the figure caption). The extent to which these quantities can reconstruct the dynamic timeseries of gene expression reflects the amount of control that voltage exerts over gene expression.

We found that the dynamic profile of the discriminator gene expression corresponding to the endogenous input pattern could be almost fully recovered with just the Hessian (Figure 7), with a match significantly higher than random expectation (Figure S8). On the other hand, the Jacobian alone produced a reconstruction with a lower quality of match that was not significantly different from random expectation (Figure S8). Taken together, these observations suggest that the voltage pattern exerts control over the discriminator gene expression almost exclusively at the second-order level. In other words, it is the voltages of pairs of cells, rather than of single cells, that influence gene expression. A possible explanation is that pattern discrimination requires a comparison of voltage of different cells, which the Hessian presumably aids by containing the differential information. The aforementioned observations also suggest that the voltage pattern

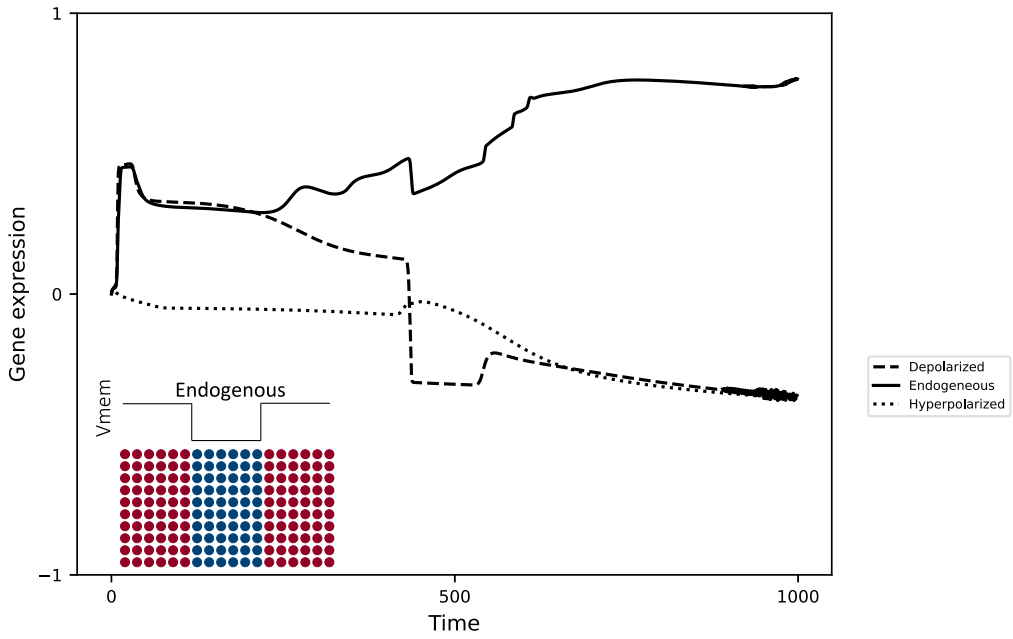


Figure 5. The best fit model flexibly scales to larger tissues when solving the voltage pattern discrimination problem

The timeseries shows the mean expression, averaged over all cells, of the discriminator gene in a scaled up 10x18 model, initialized with the endogenous voltage pattern (inset; left and right six columns are depolarized, and the middle six columns are hyperpolarized), showing that it successfully discriminates between the correct and incorrect voltage patterns. More detailed timeseries resolved at the cell of the level and for larger tissues are presented in Figure S3.

controls tissue-level gene expression in an effectively feedforward manner (a simple analog of this is shown in Figure S6 and analyzed in Figure S7) despite the complex recurrent connectivity between them (Figure 4; also demonstrated by the fact that severing the feedback connections from the genes to the voltage partially affects the pattern discrimination ability of the model, as shown in Figure S12). In other words, the bioelectric pattern exerts a "canalized" control over the discriminator gene's expression by effectively linearizing the complex nonlinearity of the network—an observation that cannot be discerned from an inspection of the static connectivity of the tissue (Figure 3) alone.

Long-distance influence and spatiotemporal integration characterizes bioelectric control

To further characterize the information-processing mechanisms underlying the bioelectric control of tissue-level gene expression, we analyzed the spatiotemporal organization of Hessian networks (Figure 8, Methods subsection "causal integration analysis") underlying the discriminator gene expression timeseries shown in Figure 7 that was reconstructed using approximation methods (Equations 3 and 4). In the context of the embryo, a Hessian network represents the extent of changes in gene expression following small perturbations in the voltages of a pair of cells after a time interval spanning a length of τ. Thus, the Hessian networks are dynamic in nature since the causal influence among the cells changes over time; in contrast, the underlying apparent tissue connectivity (Figure 1F) is static in nature. We found that the Hessian networks are characterized by symmetry breaking and long-distance influence. Furthermore, the resemblance between them and the original network declines over time during the simulation, as evidenced by the fact that the locality of the influencing cell is broken, despite the symmetric and topographic nature of the underlying connectivity. This is due to the fact that the influences are dynamic and oscillatory in that they start out locally during the early time points ($t = 5$), then switch to the other bands ($t = 50 - 410$) and then back to being more local ($t = 495$). During this process, the signs of the influence also change. For example, the influence exerted by all the bands tends to be mostly negative up until $t = 125$, but around $t = 200$ they exert differential influence, in that the middle band continues to exert negative influence, whereas the marginal bands switch to a positive influence. This time point also roughly happens to be the decision-making point where the gene expressions diverge for the correct and the incorrect input patterns (Figure 4A). Overall, the oscillatory scanning-like behavior of the influence suggests that information is integrated both over space and time before a decision is made—a hypothesis that is further supported by the observation that a similar mechanism is exhibited by the Jacobian as well (Figure S10). A possible explanation is that since the genes in a cell have immediate access only to the cell's own voltage, it would need to integrate the voltage information from across the tissue before the tissue as a whole detects the class of the input pattern and makes a collective decision.

Division of labor: Different genes are attuned to different spatial scales of the tissue

We next analyzed the scales of sensitivity of each gene by computing the difference between the Jacobian (first-order) sensitivity to the whole tissue and that of the single cell that contains it, averaged over the cells in the top left quadrant of the tissue (Figure 9;

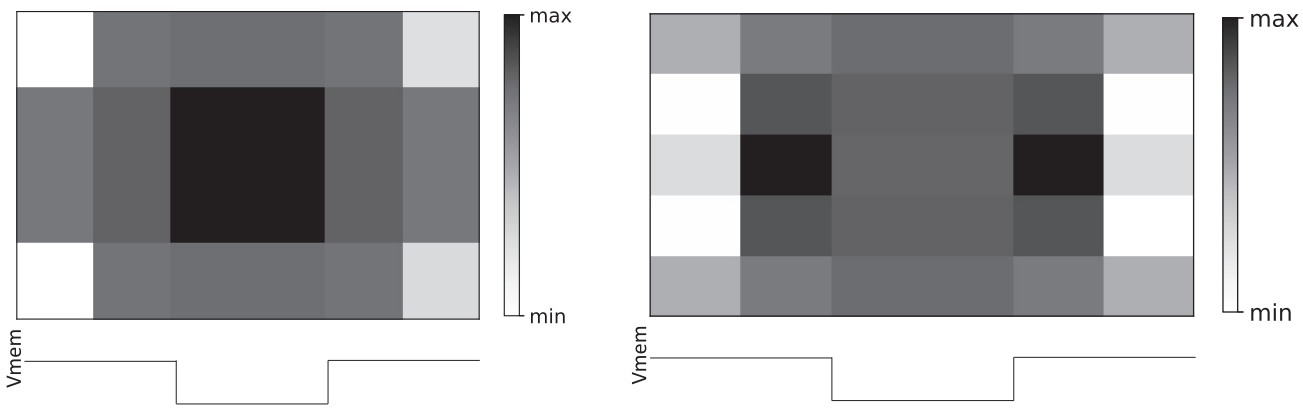


Figure 6. Differential change in model performance due to knocking out individual or patches of cells in the tissue

The darker the shade of a block, the higher the drop in performance due to knocking out that block. A cell or a patch was "knocked out" by freezing the states of its corresponding genes at zero and dropping their biases to negative infinity, thereby rendering them ineffectual both as inputs and outputs.

(A) A 4x6 tissue where individual cells were knocked out.

(B) A 10x18 tissue where 2-cell x 3-cell patches (each block is a patch) were knocked out.

mathematical details in figure caption). We found that there is a unique apportioning of the gene sensitivities among the various spatial scales of the voltage pattern: some genes (2 and 5) are more attuned to the pattern at the tissue level, some (genes 3 and 4) to the single-cell level, and others (genes 1 and 6) display a fine balance between the two scales. This distribution of labor suggests that different genes play different functional roles tied to their respective spatial sensitivities. Indeed, we found that knocking out genes 2 or 5 affected detection of the non-endogenous uniform patterns but not the detection of the endogenous voltage pattern, whereas knocking out genes 4 or 6 affected detection of the latter (Figure S11). This implies that genes 2 and 5 are specialized for the homogeneous (incorrect) patterns, whereas genes 4 and 6 are specialized for the heterogeneous patterns. A possible explanation that ties these observations to the genes' spatial sensitivity scales is that detection of homogeneity requires a tissue-level sensitivity, but heterogeneity requires a relatively smaller scale of sensitivity in this model. It appears that genes 2 and 5 sense whether the voltages across all cells in tissue are equal, whereas the other genes are sensing local differences in voltages. The tissue is deemed heterogeneous if a local patch is heterogeneous, whereas the tissue is homogeneous only if all local patches are homogeneous, hence the differences in the corresponding scales of sensitivities.

Model predicts preservation or partial disruption of outcome in response to non-endogenous voltage patterns with altered polarization ratios

We already know that incorrect voltage patterns can disrupt gene expression and morphogenesis *in vivo*.[30,31,80,87,88] For example, manipulating ion channels to produce a uniformly hyperpolarized pattern in this region altered forebrain markers gene expression and resulted in stage 45 tadpoles with deformed forebrain, midbrain, and eyes (Figures 2 and 3 from ref.[80]). Hence, we next asked whether our model can provide insight into the robustness of this system: which changes in the voltage pattern do and do not interfere with correct brain morphological patterning? To that end we sought to test predictions suggested by our analysis of the model's information-processing mechanisms.

The analyses in the previous section suggest that an asymmetric context dependency exists among the cells: the extent to which the voltage of a cell influences the gene expression in other cells depends on the type of polarization of the other cells. Specifically, there is a greater dependency on depolarized cells compared with hyperpolarized cells, on average, as seen in Figure 8 and quantified in Figure S5. In other words, depolarized cells exert a greater influence than hyperpolarized cells on the gene expression of other cells regardless of their polarization. This suggests that the proportion of depolarized to hyperpolarized cells in the tissue may be one of the factors that determines how the voltage pattern is processed. Consequently, we tested two types of modified endogenous patterns: a "half-and-half" pattern and a "sharpened" pattern. In the half-and-half pattern, one-half of the tissue is hyperpolarized and the other half is depolarized (in essence a step function pattern), whereas in the sharpened pattern the overall voltage pattern is maintained but fewer cells in the central band are hyperpolarized compared with the normal pattern.

The gene activity in response to the half-and-half pattern was similar to that of the endogenous pattern on average, with a mean asymptotic activity of about 0.7 and 0.9 for the endogenous pattern and a mean overall activity of 0.5 in both cases, thereby predicting either a normal brain morphology or one with very minor defects (Figure 10A). On the other hand, the gene activity in response to the sharpened pattern was lower on average, with a mean asymptotic activity of about 0.4 as compared with 0.9 for the endogenous pattern and mean overall activities of about 0.3 and 0.5, respectively, thereby predicting a brain morphology with minor or substantive defects (Figure 10B). A more nuanced analysis is presented in Figures S4 and S14. We next tested these predictions *in vivo*.

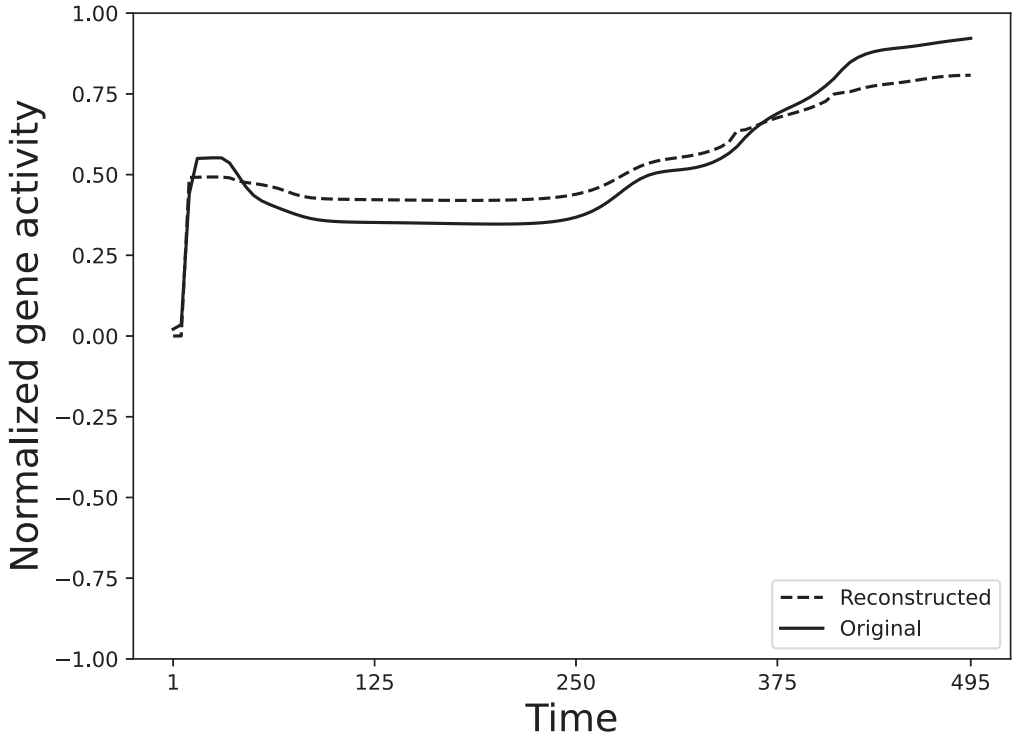


Figure 7. Analytical reconstruction of the gene expression timeseries using only the second-order influence of the voltage pattern on gene expression

A comparison between the normalized original expression timeseries of the discriminator gene (dashed) for the endogenous input pattern and the corresponding reconstruction (solid) obtained from the second-order (Hessian) influence of the endogenous voltage pattern reveals their similarity. The timeseries was reconstructed using the equation $\widehat{g_6(t+1)} = \widehat{g_6(t)} + \sum_{j \in TL} \sum_{l=1}^{n_c} \sum_{k=1}^{n_c} c_H[\delta] \frac{\partial^2 g_{6,j}(t)}{\partial G_{p_l} \partial G_{p_k}(0)}$, where *TL* refers to the set of cells in the top left quadrant of the tissue and c_H was optimized to fit the observed normalized $\{\widehat{g_6(t)}\}$ with the initial condition $\widehat{g_6(0)} = 0$. Only the cells in the 2x3 top left quadrant of the full 4x6 tissue (dashed boxes) were considered because the model is bilaterally symmetrical both about the vertical and horizontal axes.

Empirical testing of prediction one: Altering the voltage pattern regulating brain development in only one-half of *Xenopus* embryo does not disrupt endogenous brain development

Our previous work showed that disrupting the normal membrane voltage pattern in the neural plate and surrounding ectoderm resulted in significant abnormalities in large-scale brain morphology and function[80,87,88,97,105,106] (Figures 1C and 1D). Hence our model's first prediction (Figure 10A) that inducing an abnormal step function bioelectric pattern will either not disrupt normal brain development or cause minor (likely undetectable) defects, was surprising. To test this prediction, we used a well-established strategy for altering voltage patterns in *Xenopus* embryos by overexpressing ion channels via mRNA microinjection.[33,42,80,87,88,105,133,134] Ion channel mRNA was titrated to the lowest levels that produced phenotypes. As in our previous studies, there were no signs of general toxicity, ill health, or off-target effects (midline patterning, overall growth rate, scale proportions, and behavior were normal[135,136]).

The neural plate and surrounding ectoderm are largely derived from the two dorsal blastomeres of the four-celled *Xenopus* embryo[137] (Figure 11A). To convert the endogenous voltage contrast pattern to a step function pattern, we co-microinjected *Kv1.5* (a voltage-gated potassium channel that hyperpolarizes[138]) with lineage-tracer *β-galactosidase* mRNA into a single dorsal blastomere (Figures 11A–11C). Uninjected embryos and embryos injected only with *β-galactosidase* mRNA served as controls. Imaging of ~stage 15 embryos using voltage reporter dyes confirmed that *Kv1.5 + β-galactosidase* mRNA microinjection hyperpolarized only the injected half of the neural plate and surrounding ectoderm, changing the normal V-shaped voltage pattern seen in control embryos to a step function pattern (Figures 11B and 11C). We then evaluated brain morphology and confirmed our microinjection tissue targeting at stage 45. Our perturbations were specifically targeted to central nervous tissue (brain, eye, and spinal cord) as confirmed by specific localization of β-galactosidase lineage label in these tissues (Figure 11G), and we did not observe any extraneous morphological defects outside CNS in the injected tadpoles. At stage 45, control tadpoles showed normal brain patterning[85] (Figures 11D and 11H). Surprisingly, as predicted by the model, the *Kv1.5 + β-galactosidase* mRNA-injected tadpoles also exhibited normal brain patterning (Figures 11E and 11H). To demonstrate that this counter-expectation result is not specific to Kv1.5 channels but is due to the altered voltage pattern, we repeated the experiment with a different channel, Kir4.1, which is also known to hyperpolarize cells.[35,139] Microinjecting *Kir4.1* mRNA in the same protocol also resulted in stage 45 tadpoles with normal eye and brain patterning (Figure 11H).

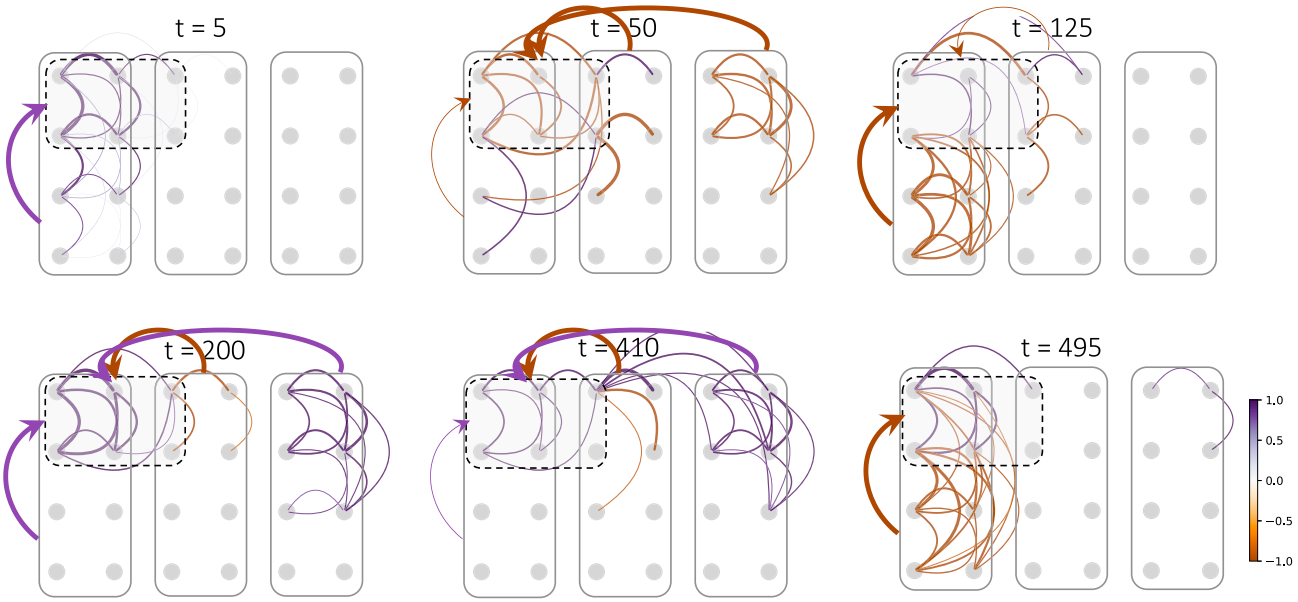


Figure 8. Spatiotemporal dynamics of the second-order control of gene expression by the endogenous voltage pattern

Each undirected edge represents a pair of cells whose voltages collectively influence the expression of an average discriminator gene in the top left quadrant. The weight of these edges indicates the degree of influence (second-order derivative) exerted by the voltages of the corresponding pairs of cells on the discriminator gene, whereas their colors indicate whether the influence is positive or negative (strong purple means strongly positive and strong orange/brown means strongly negative). These edges represent the summation term in the equation $\widehat{g_6(t+1)} = \widehat{g_6(t)} + \sum_{j \in TL} \sum_{l=1}^{n_c} \sum_{k=1}^{n_c} c_H[6] \frac{\partial^2 g_{6,j}(t)}{\partial G_{P_l} \partial G_{P_k}(0)}$, where TL refers to the set of cells in the top left quadrant of the tissue and c_H was optimized to fit the observed normalized $\{\widehat{g_6(t)}\}$ with the initial condition $\widehat{g_6(0)} = 0$. Only the cells in the 2x3 top left quadrant of the full 4x6 tissue (dashed boxes) were considered because the model is bilaterally symmetrical both about the vertical and horizontal axes. For clarity, only about 10% of the observed edges are shown. The arrows running from one band to another serve as a visual illustration of the average weighted and signed influence exerted by the three bands on the target top left quadrant of the pattern.

These results confirm the model's first prediction (Figure 10A) that certain specific deviations from the normal endogenous voltage pattern responsible for brain development do not result in large-scale brain morphology defects.

Empirical testing of prediction two: Sharpening the voltage contrast pattern by reducing the number of neural plate hyperpolarized cells in the endogenous pattern results in brain morphology defects

To test the model's second prediction (Figure 10B), we used the same strategy, but in this case injecting a well-characterized dominant-negative K_{ATP} construct (DNKir6.1p), which is known to inhibit endogenous K_{ATP} channels and cause depolarization.[22,33,140] Again our perturbations were specifically targeted to central nervous tissue (brain, eye, and spinal cord) as seen by specific localization of β-galactosidase in these tissues in the tadpole (Figure 12G), and we did not observe any extraneous morphological defects outside CNS in the injected tadpoles. In comparison to control embryos, which exhibit the characteristic voltage contrast pattern and normal brain morphology[85] (Figures 12D and 12H), *DN-K_{ATP} + β-galactosidase* mRNA microinjection depolarized only the injected half of the neural plate and surrounding ectoderm (Figures 12B and 12C), decreasing the number of hyperpolarized cells in the neural plate to create a narrower central band of hyperpolarized cells while maintaining the overall voltage contrast pattern (Figures 12B and 12C). A reasonable expectation (based on previously published manipulations of the voltage pattern *in vivo*) would be that this compressed, but qualitatively similar, pattern would still be sufficient to direct normal brain morphology. However, our results defied this expectation and instead confirmed the model's novel prediction. At stage 45, *DN-K_{ATP} + β-galactosidase* mRNA-injected tadpoles had significant brain morphology defects, but only on the injected side (Figures 12E and 12H). β-galactosidase staining of stage 45 tadpoles confirmed intended tissue targeting (Figures 12F–12H).

These results confirm the model's second prediction (Figure 10B) that, in addition to the shape of the voltage pattern, the proportion of depolarized and hyperpolarized cells that form the pattern is also a crucial factor in regulating embryonic brain development.

DISCUSSION

Knowledge gap: Mapping multicellular voltage pattern to gene expression

A central goal of developmental biology is to understand how multiple modalities (e.g., biophysical, transcriptional, and biochemical) interact to reliably produce the correct target morphology of complex organs. Bioelectric signaling is becoming increasingly recognized as an

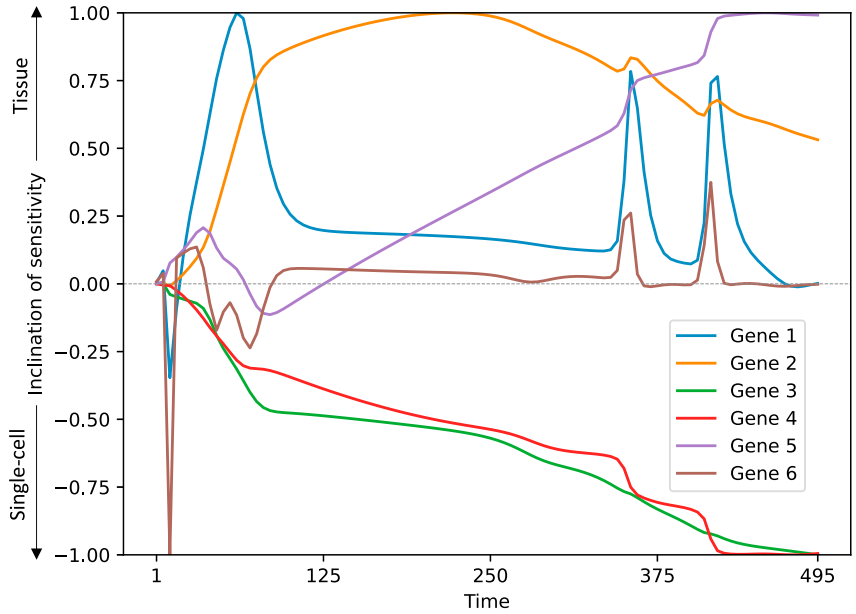


Figure 9. A division of labor exists among the genes in terms of their sensitivities to the different spatial scales of the voltage pattern

The difference between the first-order (Jacobian) directional derivatives of the gene activities computed with respect to the voltage at the tissue level and the single-cell level tends to be positive for genes $i = 2$ and 5, negative for genes 3 and 4, and mostly balanced but interspersed with positive spurts for genes 1 and 6. This quantity, known as the net scale sensitivity of a gene, is defined as $S(i) = \sum_{j \in TL} \sum_{l=1}^{n_c} \frac{1}{\sqrt{n_c}} \frac{\partial g_{l,j}(t)}{\partial G_{p_j}(0)} - \sum_{j \in TL} \frac{\partial g_{i,j}(t)}{\partial G_{p_j}(0)}$. Here, the first term is a directional derivative and computes the net sensitivity to the tissue, and the second term refers to the single cell. The term $1/\sqrt{n_c}$ is a normalization term for computing the directional derivative along the direction of the unit vector, thus enabling the comparison between the two scales of the tissue and the single cell. The corresponding inclinations of sensitivity are toward the tissue level, single-cell level, and an intermediate level, respectively. Due to the symmetry of the model, only the genes contained by the cells in the 2x3 top left quadrant of the full 4x6 tissue were considered for these calculations.

important component of developmental control mechanisms,[19–21] especially as concerns craniofacial patterning.[30,31] Great strides have been made in dissecting cellular level transduction mechanisms and specific transcripts downstream of voltage-mediated signaling. However, much remains to be learned about how multicellular bioelectric patterns are interpreted by cellular collectives and bioelectric patterns interface with downstream genetic, biochemical, and biomechanical signals.[57,59,60,82,88,141]

Evolution utilizes bioelectric patterns as a code,[142–144] because, like other biological codes, bioelectric signaling involves the mapping of arbitrary signals to specific outcomes. Specific voltage patterns are now known to be associated with organ-level fates, able to induce for example tails,[64] heads,[81] eyes,[22] and a variety of other structures in ectopic locations. In all these cases, the triggering stimulus is not a specific gene product that intrinsically is associated with that organ structure, but a physiological state that is arbitrary in the sense that it could have been mapped to any outcome just as easily. Indeed, specific ion channels and pumps, and even ionic species, can be swapped out at will, as long as the resulting electrophysiological states are the same (e.g., ref.[63]). Thus, much as with DNA, epigenetic, and other codes, the assignment of trigger to outcome is set by the interpretation machinery, not by intrinsic (biophysical or genetic) properties of the stimulus. Thus, efforts to understand and exploit this for biomedical purposes can reasonably be viewed as attempting to crack the bioelectric code[107]: to work out the mapping of bioelectric input states (the symbols) to the anatomical/morphological outcomes they entail *in vivo*.

Critically, it is now seen that during development and regeneration, the triggering stimulus for specific responses is not the voltage of a single cell, but a spatial distribution of resting potentials. Thus, a key question concerns how downstream morphogenetic decisions (such as formation of eyes,[22] patterning of the anterior-posterior[61,81] and left-right[32,134] axes, and size control[39–42]) are triggered by voltage patterns spread across cell fields (not simply the cell-autonomous control of transcription by the cell's own V_{mem}). If diverse aspects of organ size and shape, driven by cell behaviors (migration, differentiation, apoptosis, and gene expression), are triggered by *specific multicellular patterns* of membrane potential, how can cells read a large-scale pattern as input and determine whether it is correct or not (Figure 13)?

A model of generic bioelectric pattern recognition by transcriptional readouts

We sought to develop a model of the interplay between voltage control mechanisms and gene-regulatory networks. Complementing loss-of-function studies identifying elements *necessary* for this to occur, we wanted to show the first constructivist model that illustrates what dynamics are *sufficient* for native gene-regulatory circuits to reach the correct state downstream of recognizing the appropriate large-scale organ prepattern. The *Xenopus* brain is an ideal context for such work because both the correct and incorrect patterns of bioelectric signals in

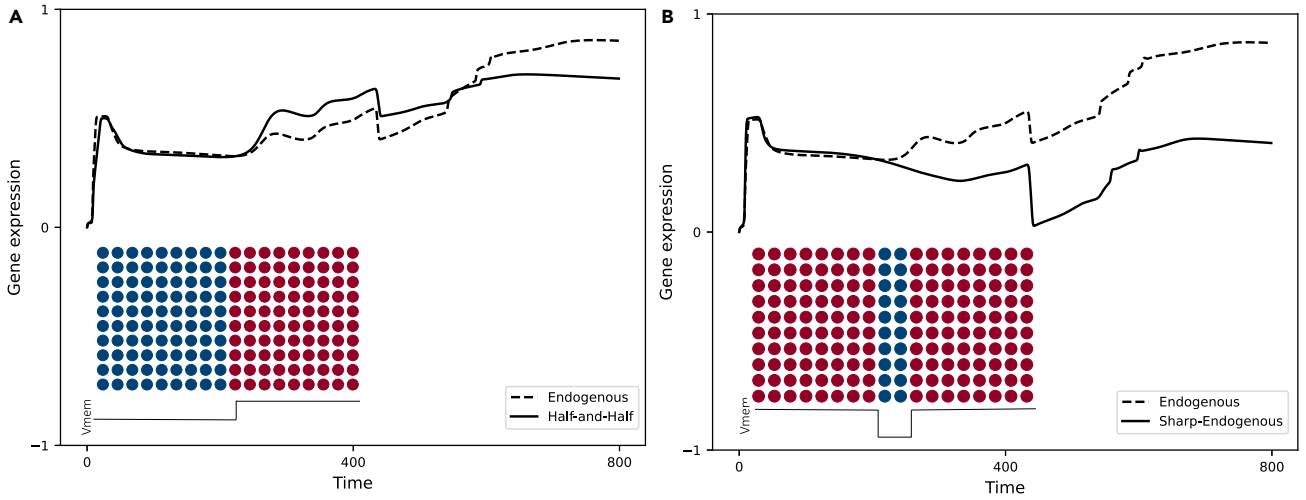


Figure 10. Model predictions about the responses to two types of non-endogenous input voltage patterns with different polarization ratios simulated in a 10x18 tissue

Average activity levels of the discriminator gene in response to (A) a voltage pattern with a less skewed ratio of polarization in comparison to the endogenous pattern: a half-and-half (step function) voltage pattern where the left half is hyperpolarized and the right half is depolarized (inset); (B) a voltage pattern with a more skewed ratio of polarization in comparison to the endogenous pattern: a sharpened voltage pattern where the overall pattern is maintained but the hyperpolarized neural plate (central band) is only a third of its original normal size (inset).

the nascent brain,[80,87,88,97,105] and at least some of the transcriptional targets required for brain development downstream,[116,145,146] are known.

Our model is generically applicable to any scenario in which specific voltage patterns induce changes in gene expression. It does not depend on the molecular identity of the genes or signaling methods and thus can be applied beyond the tadpole brain. However, it was formulated so as to be completely consistent with the large amount of existing knowledge about brain patterning. For example, the downstream genes being controlled are known to include Otx2, Pax6, Foxg1, Emx1, Sox2, and Notch. Likewise, the paracrine signaling mechanism shown in the model (Figures 1, 2, and 3) can be instantiated by known signals such as calcium, cAMP, serotonin, integrin, and Notch. There are additional elements such as morphogens[14,15,100] that are crucial and evolutionarily conserved components of neural patterning, which are not included here because the details of their connection to bioelectric regulation of brain patterning are not yet available. However, bioelectric control of morphogen movement has been seen in other systems, and this can be readily added to our model in specific cases where data exist.

Emergent features of the model not directly selected for: An analysis of model dynamics

The minimal recurrent dynamical model described here, designed using supervised machine learning techniques, recapitulates the experimentally observed phenomenon of bioelectric control of morphogenesis in the nascent *Xenopus* embryo: only certain appropriately shaped voltage contrast patterns across the neuroectoderm tissue result in the correct development of the brain. Even though we trained a small model, consisting of only 24 cells, its behavior scaled up to larger tissues with up to about 400 cells—a biologically realistic number. Interestingly, this suggests that the learned dynamics are somewhat general, that is, not specific to the particular network size or shape seen in training, though within limitations. Moreover, this means that any insights derived from the analysis of the original smaller model would be applicable to a range of larger tissue sizes. The cap on the scaling ability of the model may be due to absence of an active scaling mechanism such as the "expander-repressor" scheme[147,148] and may simply have been a result of a buffering mechanism that allows tolerance to delay in signal propagation from distant regions of the tissue. This may be due to the fact that the input bioelectric pattern is already scaled (by the experimenter) and may itself implicitly supply most of the scaling information.

Analysis of the information-processing mechanisms underlying the model's pattern discrimination behavior revealed that it employs a feedforward-like control mechanism (Figure 7) where the spatial voltage pattern influences a relatively more dominant control on gene expression, rather than the other way around, despite the complex feedback connectivity between the bioelectric patterns and the genes (Figure 3). Further analysis revealed a unique oscillatory mechanism with which different regions of the tissue influence gene expressions in distant cells at different points in time (Figure 8). The statistical properties of this mechanism predicted that other spatial voltage patterns with atypical ratios of the number of hyperpolarized to depolarized cells in the neural plate tissue could also result in a normal brain (Figure 10). These predictions were later experimentally verified (Figures 11 and 12).

Our model demonstrates how bioelectric control of gene expression can extend beyond the scope of a single cell[51] by showing that the macroscale shape of bioelectric patterns across a tissue can control gene expression in the (microscale) cells comprising it. In other words, our

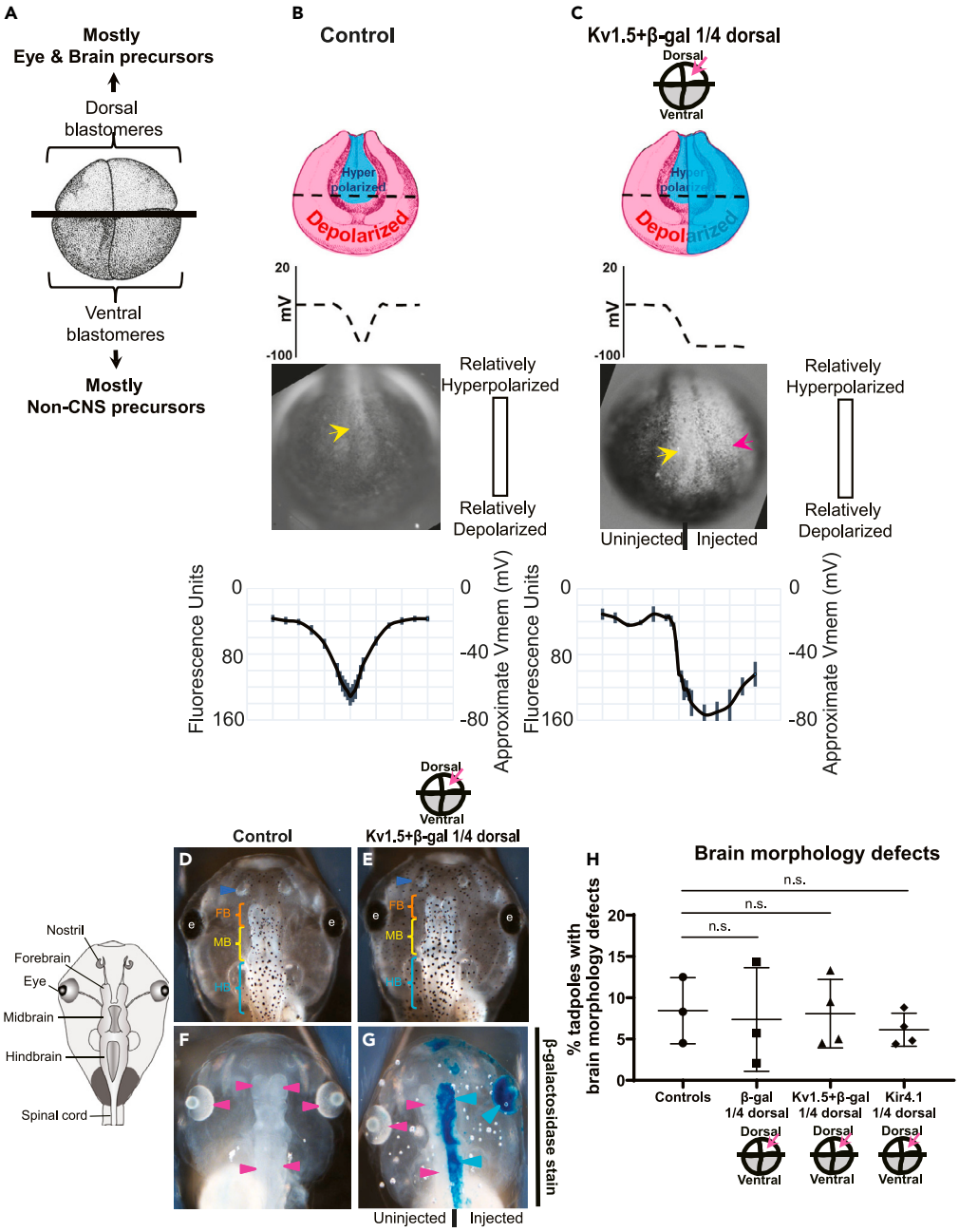


Figure 11. Perturbation of voltage in one-half of an embryo, creating a step function voltage pattern, does not cause defects in brain morphology

(A) Illustration of *Xenopus* embryo at stage 3 (four-cell), indicating two dorsal blastomeres as main precursors of eye and brain and two ventral blastomeres as main precursors of non-neural tissues.[137,171]

(B) Normal voltage pattern in control (uninjected) stage ~15 *Xenopus* embryos. Top: illustration of expected membrane voltage distribution in stage 15 embryo and expected voltage pattern along the dashed line; bottom: representative image of CC2-DMPE:DiBAC voltage-reporter-dyes-stained embryos, showing characteristic hyperpolarization in the neural plate (yellow arrows) and surrounding depolarized ectoderm[80,87,88] (N = 6) and quantification of CC2-DMPE:DiBAC$_4$ images along the dashed line indicated in the illustration along with electrophysiology-based membrane voltage approximations. Data are mean+/− SD.

(C) Step function voltage pattern at stage ~15 in embryos microinjected with *Kv1.5 + β-galactosidase* mRNA in one dorsal blastomere at four-cell stage. Top: illustration of expected membrane voltage distribution in injected stage 15 embryo and expected voltage pattern along the dashed line; bottom: representative image of CC2-DMPE: DiBAC voltage-reporter-dyes-stained *embryos*, showing characteristic hyperpolarization in neural plate (yellow arrow) but hyperpolarized ectoderm only on the injected side (red arrow) (N = 6) and quantification of CC2-DMPE:DiBAC$_4$ images along the dashed line indicated in the illustration along with electrophysiology-based membrane voltage approximations. Data are mean+/− SD.

Figure 11. *Continued*

(D–G) Representative images of stage 45 tadpoles. (D, E) Tadpoles from uninjected (D) or injected (E) embryos. Blue arrowheads indicate intact nostrils, orange brackets indicate intact forebrain (FB), yellow brackets indicate intact midbrain (MB), cyan brackets indicate intact hindbrain (HB) and intact eyes (e) (F, G) β-Galactosidase expression assessed using X-Gal (blue) in bleached tadpoles. In injected tadpoles (G), X-gal staining is evident in the eye and brain on the injected side (blue arrowheads) but not the uninjected side (magenta arrowheads), validating our targeting of mRNA microinjection. There was no X-gal staining (magenta arrowheads) in uninjected tadpoles (F) (N > 10 tadpoles per experimental group).

(H) Quantification of brain morphology defects in stage 45 tadpoles under different injection conditions demonstrates no significant differences among any of the injection conditions. All injections were made into the right dorsal blastomere at 4-cell stage as indicated in the schematic under each graph. Percentage of tadpoles with brain defects for each experimental group are as follows: Controls: 8%, β-galactosidase: 7%, Kv1.5 + β-galactosidase: 8%, and Kir4.1: 6%. Data are mean ± SD, n.s. = non-significant (one-way ANOVA with Tukey's post-hoc test for n = 3 independent experiments with N > 50 embryos per treatment group per experiment).

model shows how gene expression can be a function of not the local voltage of individual cells but the global multicellular voltage pattern. This fact is made evident in the observation in Figure 4A that while a fully depolarized input voltage pattern results in deactivation of the genes, the endogenous pattern activates them, even though the polarization levels of the left and the right bands of the tissue are the same in both patterns (depolarized). Our analysis revealed interpretable spatiotemporal information-integration mechanisms that the model employs to both recognize and discriminate between the various input voltage patterns, as depicted in Figures 6, 7, 8 and 9. Importantly, the model was not micro-managed to implement all of the features we found; our analysis of this model revealed numerous surprising emergent features that had not been specifically included in the design, thus illustrating the power of such approaches for uncovering emergent aspects of even simple mechanisms.[149,150] Further experimental testing of these predictions of such emergent features is needed to verify if they indeed are truly being employed in various morphogenetic contexts.

One of the unique features of the mechanism is functional specialization. There are two ways in which functional specialization manifests in our model. First, certain cell positions are more important in interpreting bioelectric patterns than others (Figure 6). This suggests that collective cell decision-making in this context occurs via a partially distributed system, which allows the system to be robust to external perturbations and at the same time be highly plastic, since change in a few key cell positions can drastically change the interpretation of patterns. This is supported by our *in vivo* experiments that demonstrate that while drastic changes in the bioelectric pattern still result in normal brain morphology (Figure 12), other relatively minor changes could significantly alter the outcomes (Figure 13). Second, different genes are sensitive to different spatial scales of the bioelectric pattern, ranging from the level of the tissue to that of the single cell (Figure 9). Even though such division of labor may be inevitable in a decentralized information-processing system, it is the particular way in which labor may be divided among the genes that we note is interesting.

The ratio of depolarized to hyperpolarized cells in the tissue is another important property of the voltage pattern that determines the outcome of the bioelectric regulation of gene expression (Figures 10 and S5). That is, the extent to which the voltage of a cell influences the gene expression in other cells depends on the type of polarization of other cells. This is validated by our *in vivo* experiment where a slight change in proportion of polarized cells while keeping the overall bioelectric pattern same results in major brain development deformities. Particularly of interest is the observation that the depolarized (non-neural ectoderm) cells are more important in this collective cell decision-making process than the hyperpolarized (neural ectoderm) cells. This means that developmental patterning of any tissue or organ is as much dependent on cells in the surrounding embryo as on the cells making up that tissue and organ. In hindsight, this makes sense as no embryonic tissue develops in a vacuum but instead is part of a whole organism, and all tissue developments have to be somewhat coordinated for the proper development of the whole organism.

The other characteristic features of the information-processing mechanism employed by our model included long-distance influence and a simplified voltage-gene control. Long-distance influence (Figure 8) allows the genes in a cell to be responsive to the states of faraway cells.[151] In particular, some of these interactions bring cells that are distant in physical space, closer in physiological (voltage) space—a phenomenon that one could view as a generalization of Hebbian learning. Not only is the responsiveness of the genes not local but also the extent of the non-locality dynamically varies. This adds another layer of complexity to gene expression and regulation especially during active morphogenesis, where gene expression at any given time can no longer be considered as a static readout of the state of that cell but should be viewed as a culmination of several influences, local and distant, perhaps similar to that of input integration on a neuron. The simplified voltage-gene control mechanism (Figure 7) allows the voltage pattern to effectively control the discriminator gene in a direct fashion, thus masking the recurrent connectivity among the genes (a simple analog of this setting is described in Methods and analyzed in Figures S6 and S7). This feature is analogous to the concept of "canalization" that is known to often endow biological systems with the ability to shield itself from random mutations.[152] In the context of complex dynamical systems, it translates into a phenomenon where only a subset of the nodes or pathways in the network actually direct the system to its final attractor state for a given initial state.[153,154] Our model also exhibits a similar canalizing relationship between the voltage pattern and the discriminator gene expression in that the former is sufficient to determine the asymptotic expression of the latter. Moreover, we have described this relationship in formal terms and characterized it as existing at the second-order level (Figure 7), as indicated by a dominant Hessian tensor. This mathematical understanding of canalization in the specific context of the relationship between bioelectricity and gene expression is crucial to advance our understanding of the bioelectric control of morphogenesis in *Xenopus* and possibly other biological systems. These hypotheses can be experimentally tested *in vivo* using a temporally dynamic gene expression monitoring system during embryonic development.

The aforementioned discussions identify the critical features of the model as dynamical phenomena such as functional specialization, sensitivity, and control. This raises the question of whether there may be structural motifs underlying some of these dynamical mechanisms.

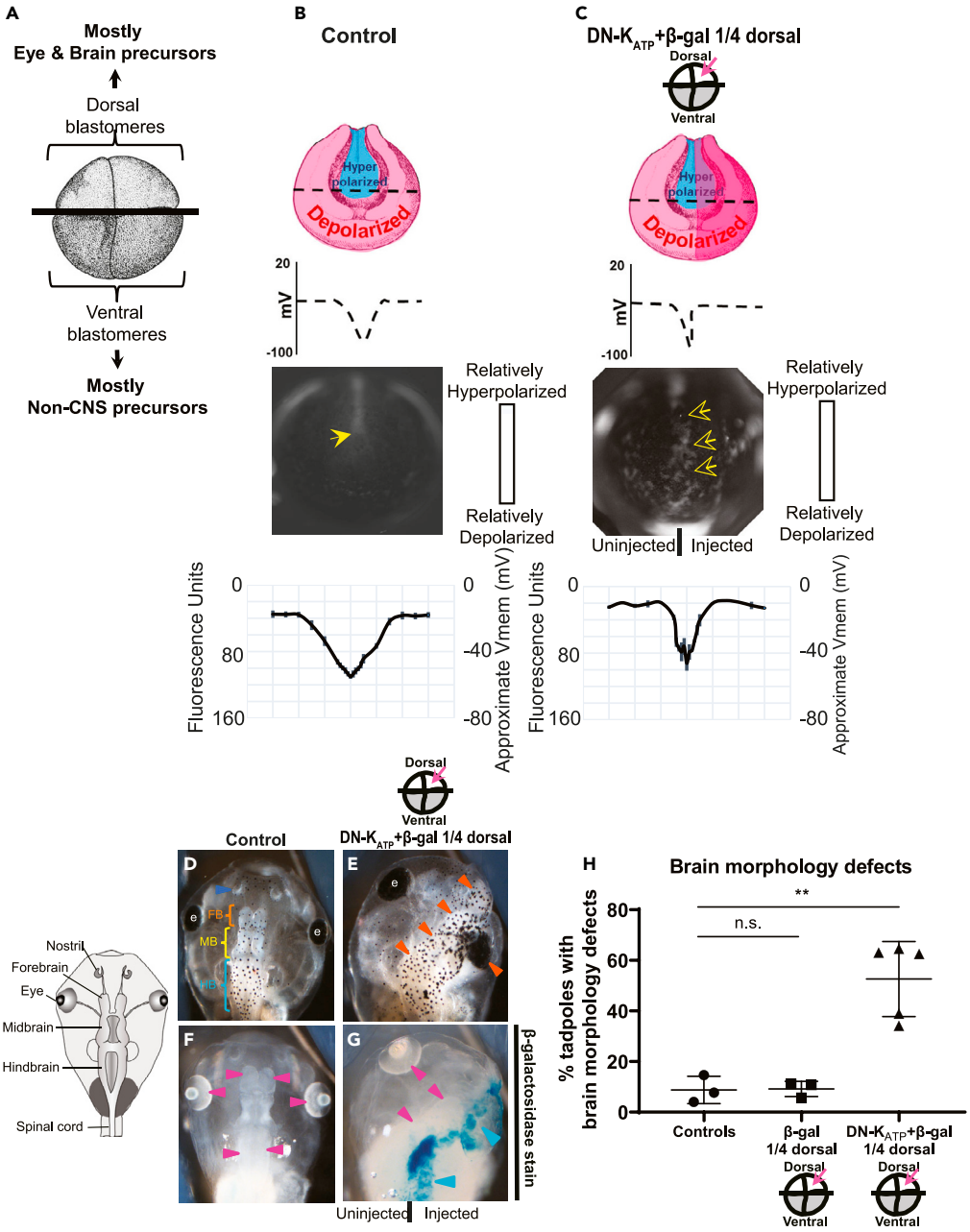


Figure 12. Perturbation of voltage to decrease neural plate hyperpolarized cells while minimally changing the overall voltage pattern leads to significant brain morphology defects

(A) Illustration of *Xenopus* embryo at stage 3 (four-cell) indicating two dorsal blastomeres as main precursors of eye and brain and two ventral blastomeres as main precursors of non-neural tissues.[137,171]

(B) Endogenous voltage pattern in control (uninjected) stage ~15 embryos. Top: illustration of expected membrane voltage distribution in stage 15 embryo and expected voltage pattern along the dashed line; bottom: representative image of embryo stained with CC2-DMPE:DiBAC$_4$ voltage reporter dyes, showing characteristic hyperpolarization in the neural plate (yellow arrows) and surrounding depolarized ectoderm[80,87,88] (N = 6) and quantification of CC2-DMPE:DiBAC$_4$ images along the dashed line indicated in the illustration along with electrophysiology-based membrane voltage approximations. Data are mean+/− SD.

(C) Altered voltage pattern in stage ~15 embryos microinjected with *DN-K$_{ATP}$ + β-galactosidase* mRNA in one dorsal blastomere at four-cell stage. Top: illustration of expected membrane voltage distribution in injected stage 15 embryo and expected voltage pattern along the dashed line; bottom: representative image of injected embryo stained with CC2-DMPE: DiBAC$_4$ membrane voltage reporter dyes showing reduced region of hyperpolarization in neural plate and narrower hyperpolarization region (empty yellow arrow) (N > 6 embryos per experimental group) and quantification of CC2-DMPE:DiBAC$_4$ images along the dashed line indicated in the illustration along with electrophysiology-based membrane voltage approximations. Data are mean+/− SD.

Figure 12. *Continued*

(D–G) Representative images of stage 45 tadpoles. (D, E) Tadpoles from uninjected (D) or microinjected with *DN-K_{ATP} + β-galactosidase* mRNA (E) embryos. Blue arrowheads indicate intact nostrils; orange, yellow, and cyan brackets indicate intact forebrain (FB), midbrain (MB), and hindbrain (HB), respectively. (e) indicates intact eyes, and orange arrowheads indicate mispatterned brain and eye. (F, G) β-Galactosidase expression assessed using X-Gal (blue) in bleached tadpoles either left uninjected (controls) or co-injected with *DN-K_{ATP} + β-galactosidase* mRNA. There was no X-gal staining (magenta arrowheads) in uninjected tadpoles (F) In injected tadpoles (G), β-galactosidase was observed in the eye and brain on the injected side (blue arrowheads) but not the uninjected side (magenta arrowheads), confirming expected targeting of microinjected mRNAs. (N > 10 tadpoles per experimental group).

(H) DN-K_{ATP} + β-galactosidase injection significantly increased the percentage of stage 45 tadpoles with brain morphology defects: uninjected or β-galactosidase controls—both 9%; DN-K_{ATP} + β-galactosidase—53%. All injections were made into the right dorsal blastomere at four-cell stage as indicated in the schematic under each graph. Data are mean ± SD, **p < 0.01, n.s. = non-significant (one-way ANOVA with Tukey's post-hoc test for n = 3 independent experiments with N > 50 embryos per treatment group per experiment).

Figure 3 highlights some of the characteristic features of the two-cell network involving recurrent connections with a mixture of positive and negative regulation that may be posited to facilitate intercellular communication in a stable manner. However, such structural patterns in a two-cell context cannot possibly convey the complete mechanism, as that would require many more than two cells. More importantly, what ultimately steers collective decision-making in our model is the dynamics that leverages structural patterns and not the structures themselves. For example, the asymmetry of the initial voltage pattern plays an important role by breaking the intercellular structural symmetry (which remains fixed regardless of the input voltage patterns) to make appropriate decisions. In other words, at least in the case of this model, analysis of the dynamics rather than structure is more important for understanding collective decision-making.

Surprising features of the model's predictions were confirmed empirically

Our model generated several unexpected predictions that were tested *in vivo* (Figures 10, 11, 12, and 13). For instance, altering the bilaterally symmetrical endogenous bioelectric pattern, which was hitherto thought to be required for normal brain development,[87,88] by reducing the number of depolarized cells and disrupting the symmetry, still led to normal brain development on *both* sides (Figures 10A and 11). This confirms the ability of the combined bioelectric-GRN system to collectively achieve outcomes despite perturbations by leveraging correct information on one side to accommodate informational defects on the other side.

On the other hand, increasing the proportion of polarized cells while maintaining the symmetry of the pattern led to major brain development defects. This is also counter-intuitive because the overall bioelectric pattern is maintained and so it should have led to normal brain patterning. These observations make sense in light of the fact that the depolarized cells in our model are more influential than the hyperpolarized cells (Figure S5): there is just enough of them in the first case to not suppress the signals passed by the hyperpolarized cells, whereas an overwhelming number of them in the second case renders the system as effectively comprising only depolarized cells. They moreover indicate the robustness of this system to perturbations: as long as critical information on integration nodes is maintained, the correct target morphology will be achieved. These results show that what might seem an obvious pattern to an observer scientist might not correctly reflect the underlying salience to the cell group, given the computations and interpretations being made by the cell collective. In other words, the cell collective's interpretation and decision-making cannot be intuited by static experimental observations. This underscores the importance of making models like this and analyzing them for effective triggers of change (and input invariants), which may not be otherwise obvious. Herein lies the relevance of models like these in gaining insights in complex biological processes.

Future work based on these results

Our model also generates many hypotheses for future experiments that may lead to a better understanding of the phenomenon studied here. For example, our model predicts an asymmetric relationship between voltage and genes in that the connectivity from the former to the latter is stronger than the reverse direction (Figures 3 and S13). We hypothesize that this feature facilitates the reliable translation of the upstream bioelectric pattern of the neural plate into the downstream gene expression despite the recurrent connectivity between them. This hypothesis is partly supported by our previous *in vivo* observation that creating ectopic neural plate pattern in *Xenopus* embryos results in ectopic gene expression and ectopic brain tissue even outside the head region.[80] Future experiments that characterize the relationship between voltage and genes in *Xenopus* embryonic tissues could attempt to verify this prediction.

One immediate impact of this current iteration of modeling is that we now have an idea of the control properties of gene expression changes regulated by neural plate bioelectric patterns. Future work can thus screen for these characteristics in expression pattern of members of Fox, Sox, Zic, and Irx families (known to be important in neural patterning) to determine which of these genes might be regulated by the neural plate bioelectric pattern. Similarly, once we know the specific integration mechanism connecting the bioelectric signals with paracrine signals, we can then search for the plethora of paracrine signals (such as calcium, cAMP, serotonin, integrin, Notch, etc.) that show a similar integration pattern; that mechanism becomes a likely candidate for paracrine mechanism involved in bioelectric signaling, which of course needs to be tested *in vivo*. This approach can accelerate work by focusing on a subset of possible mechanisms to be empirically tested for involvement in bioelectric signaling.

Another future experiment is inspired by the causal integration analysis of our model that shows that information about the bioelectric pattern is integrated across space and time into gene activity (Figure 8). Specifically, the genes in a cell are differentially influenced by the voltages of cells whose locations depend nonlinearly on the stage of development (Figure 8). For example, the most influential cells at

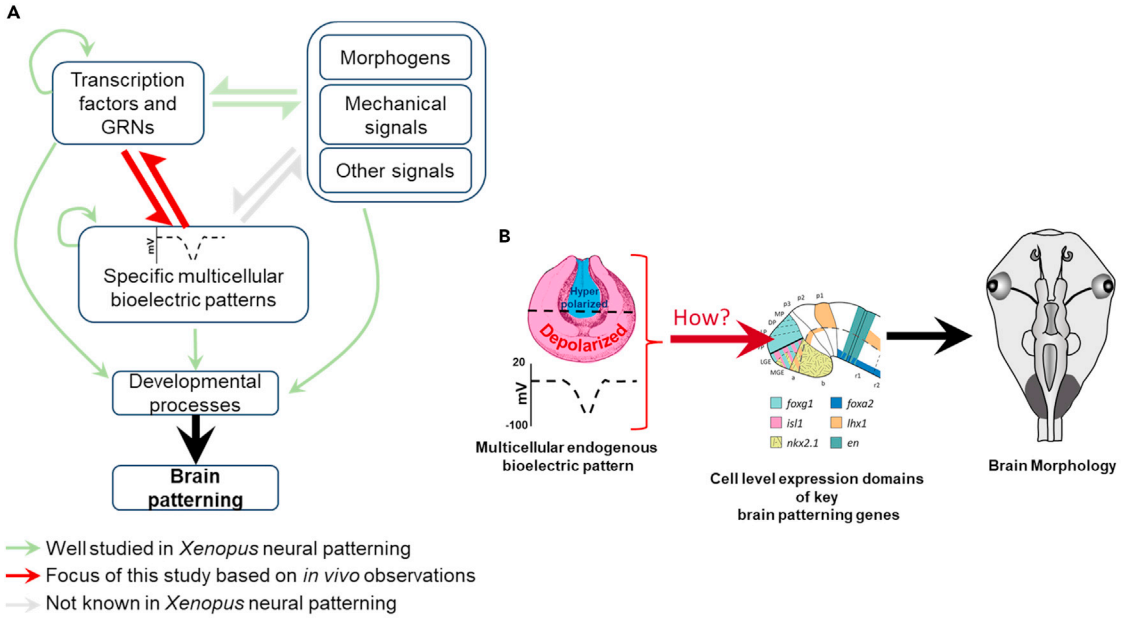


Figure 13. An overview of the current knowledge and gaps in our understanding of *Xenopus* brain development

(A) A schematic of the general state of our knowledge in embryonic brain patterning and the area of focus of this study.

(B) A schematic of the functional relationship in which a spatial pattern (the characteristic bell curve voltage distribution, left panel) is transduced into specific gene expression decisions (middle panel, taken from with permission from ref.[84]) that is required to implement mature brain anatomy (right panel).

latest developmental stages could be the ones that are close by rather than those that are further away (Figure 8, $t = 125$, for instance). Experiments that attempt to understand the dynamics of bioelectric regulation during development could characterize the nonlinear pattern of this spatiotemporal influence and match its characteristics with that predicted by our model. Finally, our model predicts that it is the voltages of pairs of cells, rather than those of single cells, that significantly influences gene expression (Figure 7). This is expected since correct detection of a voltage pattern requires comparisons of the voltages of pairs or even larger groups of cells at the minimum; detection of the whole pattern would be an outcome of the integration of those comparisons. Future experiments could study the effect of perturbations of the voltages of pairs or even groups of cells on tissue-wide gene expression. If verified, it would imply that bioelectric regulation is indeed of a higher-order nature, and it would inspire more experiments involving multidimensional perturbations that might help decipher the combinatorial nature of the bioelectric code and biological regulation in general. In future iterations of this model (after collecting insights from experimental investigations) we will incorporate bioelectric-morphogen and bioelectric-mechanical signals during neural patterning in the model. This will allow us to make predictions in a wider area of neural patterning serving an even wider community of scientists working in this area. This approach shows that the modeling work can be of immense biological relevance without needing every single molecular level detail in the model.

Broader aspects of the dynamics uncovered by the model

Our results complement the current focus on molecular pathways in which events propagate at a single scale: for example, proteins interacting with other proteins and genes regulating other genes within a cell. Here, we see how specific genes can be triggered by multicellular patterns. This, and the surprising aspects of long-range impact and stability we uncover, may have implications for gene therapy. For instance, it may be possible to regulate certain genes not directly but indirectly through large-scale features such as bioelectric voltage patterns present at some distance.[88,97,155–158]

Our model also shares functional features with a class of problems faced by nervous systems: for example, to respond to a visual stimulus, behavioral modules must be triggered by complex patterns of inputs in the retina, not the states of individual retinal cells. The tissue bioelectric system has been proposed to be an ancient version of a complex set of circuits which began by controlling the traversal of body configuration through anatomical morphospace[3,159] and eventually evolved to solve the same navigational problem in conventional 3D behavioral spaces.[159] The proposal that tissue developmental bioelectric signaling may be the evolutionary origin of the brain, and of its multiscale, top-down pattern recognition capacities,[1,9] is consistent with the idea that morphogenesis performs some of the same computational tasks.[160]

Given the important role that electrophysiology has in regulating the scale of signaling across biology, we can ask if the physical boundaries of a given cell in our multicellular model (Figure 1F) are relevant and to what extent. Are there larger or smaller *effective*

boundaries cutting across clusters of cells and genes that are more meaningful for the pattern discrimination problem? It has been hypothesized, in the context of cancer and morphogenesis, that when active biological units such as cells come together in informationally connected groups,[161–163] the computational boundary demarcating a coherent "individual" could scale up from single units to the collective. Large-scale phenomena such as morphogenesis have been described as behaviors of a collective intelligence in anatomical morphospace.[3,164,165] Could our neuroectoderm tissue similarly possess functional boundaries larger than that of a single cell in ways that could facilitate pattern discrimination at the tissue level? Our analysis has already provided hints to support this view, where small clusters of cells seem to act as coherent *modules* in terms of the constituent cells' similarity in behavior during the spatiotemporal integration process (Figure 8). Moreover, more than one such instructive causal nexus exists in the tissue—these may be separated in physical space but interact with each other at certain times (Figure 8)—an indication that they may be closer in "physiological space" even if distant in physical space.

These results weave together the components of an exciting emerging field, including developmental biophysics, dynamical systems, and machine learning. Future advances in computational perspectives on the activity of multi-scale biological systems will reveal both mechanistic and cognition-based perspectives that optimize insight, prediction, and control. These frameworks will be an essential component of the future roadmap to exploit the innate competencies of biological systems as information-processing agents, thus overcoming the many fundamental barriers facing purely bottom-up molecular approaches. Complementary fusion of both approaches will greatly potentiate the discovery of interventions for regenerative medicine and synthetic bioengineering.

Limitations of the study

Our model, an example of digital embryogeny,[166–170] was designed not to make claims about specific gene products, but rather to show a general design principle sufficient to explain the observed tissue-wide, multimodal collective behavior. Such models will be made even more bio-realistic in the future by training the model from scratch with real-time dynamic physiomic and transcriptomic datasets that will come online in subsequent years. Our analysis offers one possible set of mechanisms that the neural ectoderm might be employing; others are not ruled out, though experimental validation of their predictions provides compelling support. It may be interesting in the future to map out the space of all possible mechanisms in a way that could even throw light on the evolutionary aspects of the system.

Our model may further be limited by its minimal and phenomenological nature, in that the model does not precisely map to the known biological components in a comprehensive way; rather it assumes simplified representations of the biological features. As a result, it does not generate precise quantitative predictions about the voltage and gene expression values that could be experimentally verified. Therefore, our model may not be helpful in predicting the outcomes of precise interventions at the level of known individual genes. Still, the model proved to be useful by virtue of its ability to generate high-level predictions about eventual morphological outcomes.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Model
  - Model simulation
  - Pattern discrimination problem
  - Machine learning search and model selection
  - Performance measures
  - Causal integration analysis
  - Reconstructing gene activity timeseries using causal derivatives
  - Microinjections
  - Membrane voltage imaging with $DiBAC_4(3)$:CC2-DMPE
  - β-Galactosidase Enzymatic detection
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.108398.

ACKNOWLEDGMENTS

We thank Erin Switzer and Rakela Colon for *Xenopus* husbandry and general laboratory assistance. We thank Joan Lemire for assistance with cloning of cDNAs into frog expression vectors. We thank many members of the Levin lab for helpful discussions, as well as Susan Lewis and Julia Poirier for assistance with the manuscript. We gratefully acknowledge grants from the TWCF (TWCF0606) and the John Templeton Foundation (Grant 62212).

AUTHOR CONTRIBUTIONS

S.M. performed computational model development and analysis; V.P.P. performed *in vivo* experiments; S.M., V.P.P, and M.L. designed experiments, interpreted data, and wrote the manuscript.

DECLARATION OF INTERESTS

M.L. is a scientific co-founder of a company, Morphoceuticals, which seeks to develop therapeutics based on bioelectrical control mechanisms. M.L. is also a scientific co-founder of a company, Astonishing Labs, which is interested in decision-making and collective intelligence of cell groups *in vivo*. The Levin lab has sponsored research agreements with both companies covering other projects.

Received: January 23, 2023
Revised: July 18, 2023
Accepted: November 2, 2023
Published: November 4, 2023

REFERENCES

1. Pezzulo, G., and Levin, M. (2016). Top-down models in biology: explanation and control of complex living systems above the molecular level. J. R. Soc. Interface *13*, 20160555.
2. Harris, A.K. (2018). The need for a concept of shape homeostasis. Biosystems *173*, 65–72.
3. Fields, C., and Levin, M. (2022). Competency in Navigating Arbitrary Spaces as an Invariant for Analyzing Cognition in Diverse Embodiments. Entropy *24*, 819.
4. Clawson, W.P., and Levin, M. (2022). Endless forms most beautiful 2.0: teleonomy and the bioengineering of chimaeric and synthetic organisms. Biol. J. Linn. Soc. Lond. *139*, 457–486.
5. Vollmer, J., Casares, F., and Iber, D. (2017). Growth and size control during development. Open Biol. *7*, 170190.
6. Mirth, C.K., Anthony Frankino, W., and Shingleton, A.W. (2016). Allometry and size control: what can studies of body size regulation teach us about the evolution of morphological scaling relationships? Curr. Opin. Insect Sci. *13*, 93–98.
7. Thompson, B.J. (2021). From genes to shape during metamorphosis: a history. Curr. Opin. Insect Sci. *43*, 1–10.
8. Newman, S.A. (2019). Inherency of Form and Function in Animal Development and Evolution. Front. Physiol. *10*, 702.
9. Pezzulo, G., and Levin, M. (2015). Re-membering the body: applications of computational neuroscience to the top-down control of regeneration of limbs and other complex organs. Integr. Biol. *7*, 1487–1517.
10. Huizar, F., Soundarrajan, D., Paravitorghabeh, R., and Zartman, J. (2020). Interplay between morphogen-directed positional information systems and physiological signaling. Dev. Dynam. *249*, 328–341.
11. Gordon, N.K., Chen, Z., Gordon, R., and Zou, Y. (2020). French flag gradients and Turing reaction-diffusion versus differentiation waves as models of morphogenesis. Biosystems *196*, 104169.
12. Zagorski, M., Tabata, Y., Brandenberg, N., Lutolf, M.P., Tkačik, G., Bollenbach, T., Briscoe, J., and Kicheva, A. (2017). Decoding of position in the developing neural tube from antiparallel morphogen gradients. Science *356*, 1379–1383.
13. Inomata, H. (2017). Scaling of pattern formations and morphogen gradients. Dev. Growth Differ. *59*, 41–51.
14. Stoeckli, E.T. (2009). Morphogens and Neural Development. In Encyclopedia of Neuroscience, M.D. Binder, N. Hirokawa, and U. Windhorst, eds. (Springer Berlin Heidelberg), pp. 2397–2401.
15. Kicheva, A., and Briscoe, J. (2023). Control of Tissue Development by Morphogens. Annu. Rev. Cell Dev. Biol. *39*, 91–121.
16. Miller, C.J., and Davidson, L.A. (2013). The interplay between cell signalling and mechanics in developmental processes. Nat. Rev. Genet. *14*, 733–744.
17. Davidson, L.A. (2012). Epithelial machines that shape the embryo. Trends Cell Biol. *22*, 82–87.
18. Davidson, L.A., Joshi, S.D., Kim, H.Y., von Dassow, M., Zhang, L., and Zhou, J. (2010). Emergent morphogenesis: elastic mechanics of a self-deforming tissue. J. Biomech. *43*, 63–70.
19. Levin, M. (2021). Bioelectric signaling: Reprogrammable circuits underlying embryogenesis, regeneration, and cancer. Cell *184*, 1971–1989.
20. Harris, M.P. (2021). Bioelectric signaling as a unique regulator of development and regeneration. Development *148*, dev180794.
21. Bates, E. (2015). Ion Channels in Development and Cancer. Annu. Rev. Cell Dev. Biol. *31*, 231–247.
22. Pai, V.P., Aw, S., Shomrat, T., Lemire, J.M., and Levin, M. (2012). Transmembrane voltage potential controls embryonic eye patterning in Xenopus laevis. Development *139*, 313–323.
23. George, L.F., Pradhan, S.J., Mitchell, D., Josey, M., Casey, J., Belus, M.T., Fedder, K.N., Dahal, G.R., and Bates, E.A. (2019). Ion Channel Contributions to Wing Development in Drosophila melanogaster. G3 (Bethesda) *9*, 999–1008.
24. Dahal, G.R., Pradhan, S.J., and Bates, E.A. (2017). Inwardly rectifying potassium channels influence Drosophila wing morphogenesis by regulating Dpp release. Development *144*, 2771–2783.
25. Zhou, Y., Wong, C.O., Cho, K.J., van der Hoeven, D., Liang, H., Thakur, D.P., Luo, J., Babic, M., Zinsmaier, K.E., Zhu, M.X., et al. (2015). SIGNAL TRANSDUCTION. Membrane potential modulates plasma membrane phospholipid dynamics and K-Ras signaling. Science *349*, 873–876.
26. Emmons-Bell, M., and Hariharan, I.K. (2021). Membrane potential regulates Hedgehog signalling in the Drosophila wing imaginal disc. EMBO Rep. *22*, e51861.
27. Ferreira, F., Luxardi, G., Reid, B., and Zhao, M. (2016). Early bioelectric activities mediate redox-modulated regeneration. Development *143*, 4582–4594.
28. Ferreira, F., Raghunathan, V., Luxardi, G., Zhu, K., and Zhao, M. (2018). Early redox activities modulate Xenopus tail regeneration. Nat. Commun. *9*, 4296.
29. Belus, M.T., Rogers, M.A., Elzubeir, A., Josey, M., Rose, S., Andreeva, V., Yelick, P.C., and Bates, E.A. (2018). Kir2.1 is important for efficient BMP signaling in mammalian face development. Dev. Biol. *444* (*Suppl 1*), S297–S307.
30. Vandenberg, L.N., Morrie, R.D., and Adams, D.S. (2011). V-ATPase-dependent ectodermal voltage and pH regionalization are required for craniofacial morphogenesis. Dev. Dynam. *240*, 1889–1904.
31. Adams, D.S., Uzel, S.G.M., Akagi, J., Wlodkowic, D., Andreeva, V., Yelick, P.C., Devitt-Lee, A., Pare, J.F., and Levin, M. (2016). Bioelectric signalling via potassium channels: a mechanism for craniofacial

dysmorphogenesis in KCNJ2-associated Andersen-Tawil Syndrome. J. Physiol. *594*, 3245–3270.
32. Levin, M., Thorlin, T., Robinson, K.R., Nogi, T., and Mercola, M. (2002). Asymmetries in H+/K+-ATPase and cell membrane potentials comprise a very early step in left-right patterning. Cell *111*, 77–89.
33. Aw, S., Koster, J.C., Pearson, W., Nichols, C.G., Shi, N.Q., Carneiro, K., and Levin, M. (2010). The ATP-sensitive K(+)-channel (K(ATP)) controls early left-right patterning in *Xenopus* and chick embryos. Dev. Biol. *346*, 39–53.
34. Morokuma, J., Blackiston, D., and Levin, M. (2008). KCNQ1 and KCNE1 K+ channel components are involved in early left-right patterning in *Xenopus* laevis embryos. Cell. Physiol. Biochem. *21*, 357–372.
35. Aw, S., Adams, D.S., Qiu, D., and Levin, M. (2008). H,K-ATPase protein localization and Kir4.1 function reveal concordance of three axes during early determination of left-right asymmetry. Mech. Dev. *125*, 353–372.
36. Stern, C.D. (1987). Control of epithelial polarity and induction in th early chick embryo. In Mesenchymal-epithelial interactions in neural development, S. Wolff and Berry, eds. (Springer-Verlag), pp. 91–100.
37. Stern, C.D. (1982). Experimental reversal of polarity in chick embryo epiblast sheets in vitro. Exp. Cell Res. *140*, 468–471.
38. Beane, W.S., Morokuma, J., Lemire, J.M., and Levin, M. (2013). Bioelectric signaling regulates head and organ size during planarian regeneration. Development *140*, 313–322.
39. Yi, C., Spitters, T.W., Al-Far, E.A.D.A., Wang, S., Xiong, T., Cai, S., Yan, X., Guan, K., Wagner, M., El-Armouche, A., and Antos, C.L. (2021). A calcineurin-mediated scaling mechanism that controls a K(+)-leak channel to regulate morphogen and growth factor transcription. Elife *10*, e60691.
40. Lanni, J.S., Peal, D., Ekstrom, L., Chen, H., Stanclift, C., Bowen, M.E., Mercado, A., Gamba, G., Kahle, K.T., and Harris, M.P. (2019). Integrated K+ channel and K+Cl- cotransporter functions are required for the coordination of size and proportion during development. Dev. Biol. *456*, 164–178.
41. Daane, J.M., Lanni, J., Rothenberg, I., Seebohm, G., Higdon, C.W., Johnson, S.L., and Harris, M.P. (2018). Bioelectric-calcineurin signaling module regulates allometric growth and size of the zebrafish fin. Sci. Rep. *8*, 10391.
42. Perathoner, S., Daane, J.M., Henrion, U., Seebohm, G., Higdon, C.W., Johnson, S.L., Nüsslein-Volhard, C., and Harris, M.P. (2014). Bioelectric signaling regulates size in zebrafish fins. PLoS Genet. *10*, e1004080.
43. Schotthöfer, S.K., and Bohrmann, J. (2020). Analysing bioelectrical phenomena in the Drosophila ovary with genetic tools: tissue-specific expression of sensors for membrane potential and intracellular pH, and RNAi-knockdown of mechanisms involved in ion exchange. BMC Dev. Biol. *20*, 15.
44. Weiß, I., and Bohrmann, J. (2019). Electrochemical gradients are involved in regulating cytoskeletal patterns during epithelial morphogenesis in the Drosophila ovary. BMC Dev. Biol. *19*, 22.
45. Weiß, I., and Bohrmann, J. (2019). Electrochemical patterns during Drosophila oogenesis: ion-transport mechanisms generate stage-specific gradients of pH and membrane potential in the follicle-cell epithelium. BMC Dev. Biol. *19*, 12.
46. Lautemann, J., and Bohrmann, J. (2016). Relating proton pumps with gap junctions: colocalization of ductin, the channel-forming subunit c of V-ATPase, with subunit a and with innexins 2 and 3 during Drosophila oogenesis. BMC Dev. Biol. *16*, 24.
47. Krüger, J., and Bohrmann, J. (2015). Bioelectric patterning during oogenesis: stage-specific distribution of membrane potentials, intracellular pH and ion-transport mechanisms in Drosophila ovarian follicles. BMC Dev. Biol. *15*, 1.
48. Fennelly, C., and Soker, S. (2019). Bioelectric Properties of Myogenic Progenitor Cells. Bioelectricity *1*, 35–45.
49. van Vliet, P., de Boer, T.P., van der Heyden, M.A.G., El Tamer, M.K., Sluijter, J.P.G., Doevendans, P.A., and Goumans, M.J. (2010). Hyperpolarization induces differentiation in human cardiomyocyte progenitor cells. Stem Cell Rev. Rep. *6*, 178–185.
50. Sundelacruz, S., Levin, M., and Kaplan, D.L. (2009). Role of membrane potential in the regulation of cell proliferation and differentiation. Stem Cell Rev. Rep. *5*, 231–246.
51. Levin, M., Pezzulo, G., and Finkelstein, J.M. (2017). Endogenous Bioelectric Signaling Networks: Exploiting Voltage Gradients for Control of Growth and Form. Annu. Rev. Biomed. Eng. *19*, 353–387.
52. Pai, V.P., Martyniuk, C.J., Echeverri, K., Sundelacruz, S., Kaplan, D.L., and Levin, M. (2016). Genome-wide analysis reveals conserved transcriptional responses downstream of resting potential change in Xenopus embryos, axolotl regeneration, and human mesenchymal cell differentiation. Regeneration (Oxf) *3*, 3–25.
53. Silver, B.B., Zhang, S.X., Rabie, E.M., and Nelson, C.M. (2021). Substratum stiffness tunes membrane voltage in mammary epithelial cells. J. Cell Sci. *134*, jcs256313.
54. Silver, B.B., Wolf, A.E., Lee, J., Pang, M.F., and Nelson, C.M. (2020). Epithelial tissue geometry directs emergence of bioelectric field and pattern of proliferation. Mol. Biol. Cell *31*, 1691–1702.
55. Cervera, J., Pietak, A., Levin, M., and Mafe, S. (2018). Bioelectrical coupling in multicellular domains regulated by gap junctions: A conceptual approach. Bioelectrochemistry *123*, 45–61.
56. Pietak, A., and Levin, M. (2017). Bioelectric gene and reaction networks: computational modelling of genetic, biochemical and bioelectrical dynamics in pattern regulation. J. R. Soc. Interface *14*, 20170425.
57. Riol, A., Cervera, J., Levin, M., and Mafe, S. (2021). Cell Systems Bioelectricity: How Different Intercellular Gap Junctions Could Regionalize a Multicellular Aggregate. Cancers *13*, 5300.
58. Cervera, J., Levin, M., and Mafe, S. (2021). Morphology changes induced by intercellular gap junction blocking: A reaction-diffusion mechanism. Biosystems *209*, 104511.
59. Cervera, J., Ramirez, P., Levin, M., and Mafe, S. (2020). Community effects allow bioelectrical reprogramming of cell membrane potentials in multicellular aggregates: Model simulations. Phys. Rev. E *102*, 052412.
60. Cervera, J., Meseguer, S., Levin, M., and Mafe, S. (2020). Bioelectrical model of head-tail patterning based on cell ion channels and intercellular gap junctions. Bioelectrochemistry *132*, 107410.
61. Beane, W.S., Morokuma, J., Adams, D.S., and Levin, M. (2011). A chemical genetics approach reveals H,K-ATPase-mediated membrane voltage is required for planarian head regeneration. Chem. Biol. *18*, 77–89.
62. Adams, D.S., Tseng, A.S., and Levin, M. (2013). Light-activation of the Archaerhodopsin H(+)-pump reverses age-dependent loss of vertebrate regeneration: sparking system-level controls in vivo. Biol. Open *2*, 306–313.
63. Adams, D.S., Masi, A., and Levin, M. (2007). H+ pump-dependent changes in membrane voltage are an early mechanism necessary and sufficient to induce Xenopus tail regeneration. Development *134*, 1323–1335.
64. Tseng, A.S., Beane, W.S., Lemire, J.M., Masi, A., and Levin, M. (2010). Induction of vertebrate regeneration by a transient sodium current. J. Neurosci. *30*, 13192–13200.
65. Chernet, B.T., Adams, D.S., Lobikin, M., and Levin, M. (2016). Use of genetically encoded, light-gated ion translocators to control tumorigenesis. Oncotarget *7*, 19575–19588.
66. Chernet, B.T., and Levin, M. (2013). Transmembrane voltage potential is an essential cellular parameter for the detection and control of tumor development in a Xenopus model. Dis. Model. Mech. *6*, 595–607.
67. Mathews, J., Kuchling, F., Baez-Nieto, D., Diberardinis, M., Pan, J.Q., and Levin, M. (2022). Ion Channel Drugs Suppress Cancer Phenotype in NG108-15 and U87 Cells: Toward Novel Electroceuticals for Glioblastoma. Cancers *14*, 1499.
68. Gentile, S. (2016). hERG1 potassium channel in cancer cells: a tool to reprogram immortality. Eur. Biophys. J. *45*, 649–655.
69. Kale, V.P., Amin, S.G., and Pandey, M.K. (2015). Targeting ion channels for cancer therapy by repurposing the approved drugs. Biochim. Biophys. Acta *1848*, 2747–2755.
70. Arcangeli, A., and Becchetti, A. (2010). New Trends in Cancer Therapy: Targeting Ion Channels and Transporters. Pharmaceuticals *3*, 1202–1224.
71. Wulff, H., Castle, N.A., and Pardo, L.A. (2009). Voltage-gated potassium channels as therapeutic targets. Nat. Rev. Drug Discov. *8*, 982–1001.
72. Djamgoz, M.B.A., Coombes, R.C., and Schwab, A. (2014). Ion transport and cancer: from initiation to metastasis. Philos. Trans. R. Soc. Lond. B Biol. Sci. *369*, 20130092.
73. Yang, M., and Brackenbury, W.J. (2013). Membrane potential and cancer progression. Front. Physiol. *4*, 185.
74. Brackenbury, W.J. (2012). Voltage-gated sodium channels and metastatic disease. Channels *6*, 352–361.
75. Payne, S.L., Levin, M., and Oudin, M.J. (2019). Bioelectric Control of Metastasis in Solid Tumors. Bioelectricity *1*, 114–130.
76. Oudin, M.J., and Weaver, V.M. (2016). Physical and Chemical Gradients in the Tumor Microenvironment Regulate Tumor Cell Invasion, Migration, and Metastasis. Cold Spring Harb. Symp. Quant. Biol. *81*, 189–205.

77. Pitcairn, E., Harris, H., Epiney, J., Pai, V.P., Lemire, J.M., Ye, B., Shi, N.Q., Levin, M., and McLaughlin, K.A. (2017). Coordinating heart morphogenesis: A novel role for hyperpolarization-activated cyclic nucleotide-gated (HCN) channels during cardiogenesis in Xenopus laevis. Commun. Integr. Biol. *10*, e1309488.
78. Pai, V.P., Willocq, V., Pitcairn, E.J., Lemire, J.M., Paré, J.F., Shi, N.Q., McLaughlin, K.A., and Levin, M. (2017). HCN4 ion channel function is required for early events that regulate anatomical left-right patterning in a nodal and lefty asymmetric gene expression-independent manner. Biol. Open *6*, 1445–1457.
79. Dahal, G.R., Rawson, J., Gassaway, B., Kwok, B., Tong, Y., Ptácek, L.J., and Bates, E. (2012). An inwardly rectifying K+ channel is required for patterning. Development *139*, 3653–3664.
80. Pai, V.P., Lemire, J.M., Paré, J.F., Lin, G., Chen, Y., and Levin, M. (2015). Endogenous gradients of resting potential instructively pattern embryonic neural tissue via Notch signaling and regulation of proliferation. J. Neurosci. *35*, 4366–4385.
81. Durant, F., Bischof, J., Fields, C., Morokuma, J., LaPalme, J., Hoi, A., and Levin, M. (2019). The Role of Early Bioelectric Signals in the Regeneration of Planarian Anterior/Posterior Polarity. Biophys. J. *116*, 948–961.
82. Cervera, J., Levin, M., and Mafe, S. (2020). Bioelectrical Coupling of Single-Cell States in Multicellular Systems. J. Phys. Chem. Lett. *11*, 3234–3241.
83. Carvalho, J. (2023). A computational model of cell membrane bioelectric polarization and depolarization, connected with cell proliferation, in different tissue geometries. J. Theor. Biol. *557*, 111338.
84. Exner, C.R.T., and Willsey, H.R. (2021). Xenopus leads the way: Frogs as a pioneering model to understand the human brain. Genesis *59*, e23405.
85. Pratt, K.G., and Khakhalin, A.S. (2013). Modeling human neurodevelopmental disorders in the Xenopus tadpole: from mechanisms to therapeutic targets. Dis. Model. Mech. *6*, 1057–1065.
86. Ta, A.C., Huang, L.-C., McKeown, C.R., Bestman, J.E., Van Keuren-Jensen, K., and Cline, H.T. (2022). Temporal and spatial transcriptomic dynamics across brain development in Xenopus laevis tadpoles. G3 *12*, jkab387.
87. Pai, V.P., Pietak, A., Willocq, V., Ye, B., Shi, N.Q., and Levin, M. (2018). HCN2 Rescues brain defects by enforcing endogenous voltage pre-patterns. Nat. Commun. *9*, 998.
88. Pai, V.P., Cervera, J., Mafe, S., Willocq, V., Lederer, E.K., and Levin, M. (2020). HCN2 Channel-Induced Rescue of Brain Teratogenesis via Local and Long-Range Bioelectric Repair. Front. Cell. Neurosci. *14*, 136.
89. Smart, M., and Zilman, A. (2023). Emergent properties of collective gene-expression patterns in multicellular systems. Cell Reports Physical Science *4*, 101247.
90. Komurov, K. (2012). Modeling community-wide molecular networks of multicellular systems. Bioinformatics *28*, 694–700.
91. Cervera, J., Meseguer, S., and Mafe, S. (2016). The interplay between genetic and bioelectrical signaling permits a spatial regionalisation of membrane potentials in model multicellular ensembles. Sci. Rep. *6*, 35201.
92. Macken, W.L., Godwin, A., Wheway, G., Stals, K., Nazlamova, L., Ellard, S., Alfares, A., Aloraini, T., AlSubaie, L., Alfadhel, M., et al. (2021). Biallelic variants in COPB1 cause a novel, severe intellectual disability syndrome with cataracts and variable microcephaly. Genome Med. *13*, 34.
93. Blum, M., and Ott, T. (2018). Xenopus: An Undervalued Model Organism to Study and Model Human Genetic Disease. Cells Tissues Organs *205*, 303–313.
94. Kaltenbrun, E., Tandon, P., Amin, N.M., Waldron, L., Showell, C., and Conlon, F.L. (2011). Xenopus: An emerging model for studying congenital heart disease. Birth Defects Res. A Clin. Mol. Teratol. *91*, 495–510.
95. Lasser, M., Pratt, B., Monahan, C., Kim, S.W., and Lowery, L.A. (2019). The Many Faces of Xenopus: Xenopus laevis as a Model System to Study Wolf-Hirschhorn Syndrome. Front. Physiol. *10*, 817.
96. Ott, T., Kaufmann, L., Granzow, M., Hinderhofer, K., Bartram, C.R., Theiß, S., Seitz, A., Paramasivam, N., Schulz, A., Moog, U., et al. (2019). The Frog Xenopus as a Model to Study Joubert Syndrome: The Case of a Human Patient With Compound Heterozygous Variants in PIBF1. Front. Physiol. *10*, 134.
97. Pai, V.P., Lemire, J.M., Chen, Y., Lin, G., and Levin, M. (2015). Local and long-range endogenous resting potential gradients antagonistically regulate apoptosis and proliferation in the embryonic CNS. Int. J. Dev. Biol. *59*, 327–340.
98. Pai, V.P., Vandenberg, L.N., Blackiston, D., and Levin, M. (2012). Neurally Derived Tissues in Xenopus laevis Embryos Exhibit a Consistent Bioelectrical Left-Right Asymmetry. Stem Cell. Int. *2012*, 353491.
99. Stern, C.D. (2005). Neural induction: old problem, new findings, yet more questions. Development *132*, 2007–2021.
100. Lee, H.K., Lee, H.S., and Moody, S.A. (2014). Neural transcription factors: from embryos to neural stem cells. Mol. Cell. *37*, 705–712.
101. Balaskas, N., Ribeiro, A., Panovska, J., Dessaud, E., Sasai, N., Page, K.M., Briscoe, J., and Ribes, V. (2012). Gene regulatory logic for reading the Sonic Hedgehog signaling gradient in the vertebrate neural tube. Cell *148*, 273–284.
102. Briscoe, J., Chen, Y., Jessell, T.M., and Struhl, G. (2001). A hedgehog-insensitive form of patched provides evidence for direct long-range morphogen activity of sonic hedgehog in the neural tube. Mol. Cell *7*, 1279–1291.
103. Delás, M.J., Kalaitzis, C.M., Fawzi, T., Demuth, M., Zhang, I., Stuart, H.T., Costantini, E., Ivanovitch, K., Tanaka, E.M., and Briscoe, J. (2023). Developmental cell fate choice in neural tube progenitors employs two distinct cis-regulatory strategies. Dev. Cell *58*, 3–17.e8.
104. Placzek, M., and Briscoe, J. (2018). Sonic hedgehog in vertebrate neural tube development. Int. J. Dev. Biol. *62*, 225–234.
105. Pai, V.P., and Levin, M. (2022). HCN2 channel-induced rescue of brain, eye, heart and gut teratogenesis caused by nicotine, ethanol and aberrant notch signalling. Wound Repair Regen. *30*, 681–706.
106. Pai, V.P., and Adams, D.S. (2019). Preventing Ethanol-Induced Brain and Eye Morphology Defects Using Optogenetics. Bioelectricity *1*, 260–272.
107. Levin, M., and Martyniuk, C.J. (2018). The bioelectric code: An ancient computational medium for dynamic control of growth and form. Biosystems *164*, 76–93.
108. Lagergren, J.H., Nardini, J.T., Baker, R.E., Simpson, M.J., and Flores, K.B. (2020). Biologically-informed neural networks guide mechanistic modeling from sparse experimental data. PLoS Comput. Biol. *16*, e1008462.
109. Raissi, M., Perdikaris, P., and Karniadakis, G.E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys. *378*, 686–707.
110. Meyer, P., and Saez-Rodriguez, J. (2021). Advances in systems biology modeling: 10 years of crowdsourcing DREAM challenges. Cell Syst. *12*, 636–653.
111. Gunawardena, J. (2014). Models in biology: 'accurate descriptions of our pathetic thinking'. BMC Biol. *12*, 29.
112. Pai, V.P., Lemire, J.M., Paré, J.F., Lin, G., Chen, Y., and Levin, M. (2015). Endogenous Gradients of Resting Potential Instructively Pattern Embryonic Neural Tissue via Notch Signaling and Regulation of Proliferation. J. Neurosci. *35*, 4366–4385.
113. Trevers, K.E., Lu, H.-C., Yang, Y., Thiery, A., Strobl, A.C., Anderson, C., Pálinkášová, B., de Oliveira, N.M., de Almeida, I.M., Khan, M.A., et al. (2023). A gene regulatory network for neural induction. eLife *12*, e73189.
114. Maharana, S.K., and Schlosser, G. (2018). A gene regulatory network underlying the formation of pre-placodal ectoderm in Xenopus laevis. BMC Biol. *16*, 1–18.
115. Beccari, L., Marco-Ferreres, R., and Bovolenta, P. (2013). The logic of gene regulatory networks in early vertebrate forebrain patterning. Mech. Dev. *130*, 95–111.
116. Hovland, A.S., Rothstein, M., and Simoes-Costa, M. (2020). Network architecture and regulatory logic in neural crest development. Wiley Interdiscip. Rev. Syst. Biol. Med. *12*, e1468.
117. Brodland, G.W. (2015). How computational models can help unlock biological systems. Semin. Cell Dev. Biol. *47–48*, 62–73.
118. Levin, M. (2014). Molecular bioelectricity: how endogenous voltage potentials control cell behavior and instruct pattern regulation in vivo. Mol. Biol. Cell *25*, 3835–3850.
119. Pinotsis, D.A., and Miller, E.K. (2022). Beyond dimension reduction: Stable electric fields emerge from and allow representational drift. Neuroimage *253*, 119058.
120. Pinotsis, D.A., and Miller, E.K. (2023). In vivo ephaptic coupling allows memory network formation. Cereb. Cortex *33*, 9877–9895.
121. Pai, V.P., Lemire, J.M., Chen, Y., Lin, G., and Levin, M. (2015). Local and long-range endogenous resting potential gradients antagonistically regulate apoptosis and proliferation in the embryonic CNS. Int. J. Dev. Biol. *59*, 327–340.
122. Serra, R., Villani, M., Barbieri, A., Kauffman, S.A., and Colacci, A. (2010). On the dynamics of random Boolean networks subject to noise: attractors, ergodic sets and cell types. J. Theor. Biol. *265*, 185–193.
123. Solé, R.V., Fernández, P., and Kauffman, S.A. (2003). Adaptive walks in a gene network model of morphogenesis: insights into the

Cambrian explosion. Int. J. Dev. Biol. *47*, 685–693.
124. Cooke, J., and Webber, J.A. (1985). Dynamics of the Control of Body Pattern in the Development of Xenopus-Laevis .1. Timing and Pattern in the Development of Dorsoanterior and Posterior Blastomere Pairs, Isolated at the 4-Cell Stage. J. Embryol. Exp. Morphol. *88*, 85–112.
125. Cooke, J. (1981). Scale of body pattern adjusts to available cell number in amphibian embryos. Nature *290*, 775–778.
126. Oviedo, N.J., Newmark, P.A., and Sánchez Alvarado, A. (2003). Allometric scaling and proportion regulation in the freshwater planarian *Schmidtea mediterranea*. Dev. Dynam. *226*, 326–333.
127. Werner, S., Stückemann, T., Beirán Amigo, M., Rink, J.C., Jülicher, F., and Friedrich, B.M. (2015). Scaling and regeneration of self-organized patterns. Phys. Rev. Lett. *114*, 138101.
128. Fankhauser, G. (1945). Maintenance of normal structure in heteroploid salamander larvae, through compensation of changes in cell size by adjustment of cell number and cell shape. J. Exp. Zool. *100*, 445–455.
129. Pietak, A., Bischof, J., LaPalme, J., Morokuma, J., and Levin, M. (2019). Neural control of body-plan axis in regenerating planaria. PLoS Comput. Biol. *15*, e1006904.
130. Rasolonjanahary, M., and Vasiev, B. (2016). Scaling of morphogenetic patterns in reaction-diffusion systems. J. Theor. Biol. *404*, 109–119.
131. Cooke, J. (1973). Properties of the primary organization field in the embryo of Xenopus laevis: IV. Pattern formation and regulation following early inhibition of mitosis. Development *30*, 49–62.
132. Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key numbers in molecular and cell biology. Nucleic Acids Res. *38*, D750–D753.
133. Vandenberg, L.N., Adams, D.S., and Levin, M. (2012). Normalized shape and location of perturbed craniofacial structures in the Xenopus tadpole reveal an innate ability to achieve correct morphology. Dev. Dynam. *241*, 863–878.
134. Adams, D.S., Robinson, K.R., Fukumoto, T., Yuan, S., Albertson, R.C., Yelick, P., Kuo, L., McSweeney, M., and Levin, M. (2006). Early, H+-V-ATPase-dependent proton flux is necessary for consistent left-right patterning of non-mammalian vertebrates. Development *133*, 1657–1671.
135. Adams, D.S., and Levin, M. (2013). Endogenous voltage gradients as mediators of cell-cell communication: strategies for investigating bioelectrical signals during pattern formation. Cell Tissue Res. *352*, 95–122.
136. Blackiston, D., Adams, D.S., Lemire, J.M., Lobikin, M., and Levin, M. (2011). Transmembrane potential of GlyCl-expressing instructor cells induces a neoplastic-like conversion of melanocytes via a serotonergic pathway. Dis. Model. Mech. *4*, 67–85.
137. Moody, S.A. (1987). Fates of the blastomeres of the 32-cell-stage Xenopus embryo. Dev. Biol. *122*, 300–319.
138. Strutz-Seebohm, N., Gutcher, I., Decher, N., Steinmeyer, K., Lang, F., and Seebohm, G. (2007). Comparison of potent Kv1.5 potassium channel inhibitors reveals the molecular basis for blocking kinetics and binding mode. Cell. Physiol. Biochem. *20*, 791–800.
139. Fakler, B., Bond, C.T., Adelman, J.P., and Ruppersberg, J.P. (1996). Heterooligomeric assembly of inward-rectifier K+ channels from subunits of different subfamilies: Kir2.1 (IRK1) and Kir4.1 (BIR10). Pflugers Arch. *433*, 77–83.
140. Schwappach, B., Zerangue, N., Jan, Y.N., and Jan, L.Y. (2000). Molecular basis for K(ATP) assembly: transmembrane interactions mediate association of a K+ channel with an ABC transporter. Neuron *26*, 155–167.
141. Cervera, J., Pai, V.P., Levin, M., and Mafe, S. (2019). From non-excitable single-cell to multicellular bioelectrical states supported by ion channels and gap junction proteins: Electrical potentials as distributed controllers. Prog. Biophys. Mol. Biol. *149*, 39–53.
142. Barbieri, M. (2019). A general model on the origin of biological codes. Biosystems *181*, 11–19.
143. Barbieri, M. (2018). What is code biology? Biosystems *164*, 1–10.
144. Barbieri, M. (1998). The organic codes. The basic mechanism of macroevolution. Riv. Biol. *91*, 481–513.
145. Beccari, L., Marco-Ferreres, R., and Bovolenta, P. (2013). The logic of gene regulatory networks in early vertebrate forebrain patterning. Mech. Dev. *130*, 95–111.
146. Iwafuchi-Doi, M., Matsuda, K., Murakami, K., Niwa, H., Tesar, P.J., Aruga, J., Matsuo, I., and Kondoh, H. (2012). Transcriptional regulatory networks in epiblast cells and during anterior neural plate development as modeled in epiblast stem cells. Development *139*, 3926–3937.
147. Umulis, D.M., and Othmer, H.G. (2013). Mechanisms of scaling in pattern formation. Development *140*, 4830–4843.
148. Simsek, M.F., and Özbudak, E.M. (2022). Patterning principles of morphogen gradients. Open Biol. *12*, 220224.
149. Stanley, K.O., and Miikkulainen, R. (2003). A taxonomy for artificial embryogeny. Artif. Life *9*, 93–130.
150. Matos, A., Suzuki, R., and Arita, T. (2009). Heterochrony and artificial embryogeny: a method for analyzing artificial embryogenies based on developmental dynamics. Artif. Life *15*, 131–160.
151. McMillen, P., Oudin, M.J., Levin, M., and Payne, S.L. (2021). Beyond Neurons: Long Distance Communication in Development and Cancer. Front. Cell Dev. Biol. *9*, 739024.
152. Waddington, C.H. (1942). Canalization of development and the inheritance of acquired characters. Nature *150*, 563–565.
153. Willadsen, K., and Wiles, J. (2007). Robustness and state-space structure of Boolean gene regulatory models. J. Theor. Biol. *249*, 749–765.
154. Gates, A.J., Brattig Correia, R., Wang, X., and Rocha, L.M. (2021). The effective graph reveals redundancy, canalization, and control pathways in biochemical regulation and signaling. Proc. Natl. Acad. Sci. USA *118*, e2022598118.
155. Busse, S.M., McMillen, P.T., and Levin, M. (2018). Cross-limb communication during Xenopus hindlimb regenerative response: non-local bioelectric injury signals. Development *145*, dev164210.
156. Chernet, B.T., Fields, C., and Levin, M. (2014). Long-range gap junctional signaling controls oncogene-mediated tumorigenesis in Xenopus laevis embryos. Front. Physiol. *5*, 519.
157. Chernet, B.T., and Levin, M. (2014). Transmembrane voltage potential of somatic cells controls oncogene-mediated tumorigenesis at long-range. Oncotarget *5*, 3287–3306.
158. Oviedo, N.J., Morokuma, J., Walentek, P., Kema, I.P., Gu, M.B., Ahn, J.M., Hwang, J.S., Gojobori, T., and Levin, M. (2010). Long-range neural and gap junction protein-mediated cues control polarity during planarian regeneration. Dev. Biol. *339*, 188–199.
159. Fields, C., Bischof, J., and Levin, M. (2020). Morphological Coordination: A Common Ancestral Function Unifying Neural and Non-Neural Signaling. Physiology *35*, 16–30.
160. Grossberg, S. (1978). Communication, Memory, and Development. In Progress in Theoretical Biology, R. Rosen and F. Snell, eds.
161. Moore, D.G., Walker, S.I., and Levin, M. (2018). Pattern Regeneration in Coupled Networks. In Proceedings of The 2018 Conference on Artificial Life held in Tokyo, Japan, T. Ikegami, N. Virgo, O. Witkowski, M. Oka, R. Suzuki, and H. Iizuka, eds. (MIT Press), pp. 204–205.
162. Moore, D.G., Valentini, G., Walker, S.I., and Levin, M. (2018). Inform: Efficient Information-Theoretic Analysis of Collective Behaviors. Front. Robot. AI *5*, 60.
163. Moore, D., Walker, S.I., and Levin, M. (2017). Cancer as a disorder of patterning information: computational and biophysical perspectives on the cancer problem. Converg. Sci. Phys. Oncol. *3*, 043001.
164. Levin, M. (2019). The Computational Boundary of a "Self": Developmental Bioelectricity Drives Multicellularity and Scale-Free Cognition. Front. Psychol. *10*, 2688.
165. Levin, M. (2021). Bioelectrical approaches to cancer as a problem of the scaling of the cellular self. Prog. Biophys. Mol. Biol. *165*, 102–113.
166. Lai, G., Leymarie, F.F., Latham, W., Arita, T., and Suzuki, R. (2021). Virtual Creature Morphology - A Review. Comput. Graph. Forum *40*, 659–681.
167. Duthen, Y., Luga, H., Lassabe, N., Cussat-Blanc, S., Breton, T., and Pascalie, J. (2010). An Introduction to the Bio-logic of Artificial Creatures. Stud Comput Intell *321*, 1–23.
168. Andersen, T., Newman, R., and Otter, T. (2009). Shape homeostasis in virtual embryos. Artif. Life *15*, 161–183.
169. Andersen, T., Newman, R., and Otter, T. (2006). Development of Virtual Embryos with Emergent Self-Repair Held in Arlington, VA (IOS Press), pp. 16–23.
170. Andersen, T., Otter, T., Petschulat, C., Eoff, U., Menten, T., Davis, R., and Crowley, B. (2005). A biologically-derived approach to tissue modeling. In Medicine Meets Virtual Reality 13, J.D. Westwood, R.S. Haluck, H.M. Hoffman, G.T. Mogel, R. Phillips, R.A. Robb, and K.G. Vosburgh, eds., pp. 15–21.
171. Kline, M.J., and Moody, S.A. (1987). Fates of the Blastomeres of the 4-Cell and 8-Cell Stage Frog (Xenopus-Laevis) Embryos. Anat. Rec. *218*, A73–A74.
172. Gilbert, S.F. (2000). Paracrine Factors. In Developmental Biology, 6th edition (Sinauer Associates).

173. Armingol, E., Officer, A., Harismendy, O., and Lewis, N.E. (2021). Deciphering cell–cell interactions and communication from gene expression. Nat. Rev. Genet. *22*, 71–88.
174. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.
175. Sive, H., Grainger, R.M., and Harland, R. (2000). Early Development of Xenopus laevis (New York: Cold Spring Harbor Laboratory Press).
176. Nieuwkoop, P.D., and Faber, J. (1994). Normal Table of Xenopus laevis (Daudin) : A Systematical and Chronological Survey of the Development from the Fertilized Egg till the End of Metamorphosis (Garland Pub).
177. Herrera-Rincon, C., Pai, V.P., Moran, K.M., Lemire, J.M., and Levin, M. (2017). The brain is required for normal muscle and nerve patterning during early Xenopus development. Nat. Commun. *8*, 587.
178. Harvey, I. (2009). In The Microbial Genetic Algorithm, G. Kampis, I. Karsai, and E. Szathmáry, eds. (Springer), pp. 126–133.
179. Villaverde, A.F., and Banga, J.R. (2014). Reverse engineering and identification in systems biology: strategies, perspectives and challenges. J. R. Soc. Interface *11*, 20130505.
180. Prinz, A.A., Bucher, D., and Marder, E. (2004). Similar network activity from disparate circuit parameters. Nat. Neurosci. *7*, 1345–1352.
181. Riedmiller, M., and Braun, H. (1992). Rprop - A Fast Adaptive Learning Algorithm (Proceedings of the International Symposium on Computer and Information Science VII).
182. Manicka, S., and Levin, M. (2022). Minimal Developmental Computation: A Causal Network Approach to Understand Morphogenetic Pattern Formation. Entropy *24*, 107.
183. Adams, D.S., and Levin, M. (2012). Measuring resting membrane potential using the fluorescent voltage reporters DiBAC4(3) and CC2-DMPE. Cold Spring Harb. Protoc. *2012*, 459–464.
184. Adams, D.S., and Levin, M. (2012). General principles for measuring resting membrane potential and ion concentration using fluorescent bioelectricity reporters. Cold Spring Harb. Protoc. *2012*, 385–397.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
$DiBAC_4(3)$ (Bis-(1.3-Dibutylbarbituric Acid) Trimethine Oxonol	Invitrogen (ThermoFisher)	Cat# B438
CC2-DMPE	Invitrogen (ThermoFisher)	Cat# K1070
X-Gal	ThermoFisher	Cat# R0401
Critical commercial assays		
mMessage mMachine	Invitrogen (ThermoFisher)	Cat# AM1340
Experimental models: Organisms/strains		
Xenopus Laevis Wild-type embryos	This lab	N.A.
Recombinant DNA		
Kv1.5	Strutz-Seebohm et al.[138]	N.A.
Dominant-negative Kir6.1 pore mutant – DNKir6.1p	Aw et al.[174]	N.A.
Software and algorithms		
neuralPlatePatterning: Python source code for reproducing the results of this paper	This paper	https://gitlab.com/smanicka/neuralPlatePatterning
Metamorph Software	Molecular Devices	https://www.moleculardevices.com/
GraphPad Prism	GraphPad by Dotmatics	https://www.graphpad.com/
NIH Fiji	Schindelin et al.[33]	https://ImageJ.net/software/fiji/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Michael Levin (Michael.levin@tufts.edu)

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Microscopy data reported in this paper will be shared by the lead contact upon request.
- All original code has been deposited at https://gitlab.com/smanicka/neuralPlatePatterning and is publicly available as of the date of publication.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Xenopus laevis embryos were fertilized *in vitro* according to standard protocols in 0.1X Marc's Modified Ringer's solution (MMR: 10mM Na^+. 0.2 mM K^+, 10.5 mM Cl^-, 0.2 mM Ca^{2+}, pH 7.8).[175] Embryos were housed at 14-18°C and staged according to Nieuwkoop and Faber.[176] Due to technical limitations embryos were not segregated by sex and all experiments include random mixture of sexes. All experiments were approved by the Tufts University Animal Research Committee (M2020-35) following the guide for the care and use of laboratory animals.

METHOD DETAILS

Model

To leverage a clear example of the morphological consequences of bioelectric pre-pattern modulation and one for which there is the most data available concerning gene expression, we focused our model on the particulars of neural development in *Xenopus* embryos at stage ~15.[119,148,149,180] We set the forward features of our phenomenological model to characteristics that are biologically plausible for stage 15 neural plate tissue (Table 1). We adopted the bioelectric feature of our model (Table 1, row (d)) from Refs[88,91] that modeled the formation

and long-range repair of spatiotemporal voltage patterns. However, these papers modeled basic relationships between bioelectric signaling and expression of single genes. Here, we designed a multicellular network model where the cells are connected in a 2-dimensional lattice fashion via symmetric intercellular communication channels termed "intercellular networks" (IN) and the individual cells consist of a bidirectional coupling between voltage and the gene regulatory network (GRN) (details in Table 1). The full mathematical description of the model is provided in Figure 2A.

The key "reverse" features of the model – those designed to be determined by machine learning – are: a) the number of generic genes (≥ 6) in the GRN (the number six is derived from the fact that there are at least six genes known to be sensitive to the endogenous bioelectric pattern); b) the network wiring of the generic GRN, the IN, and of the connections from voltage to genes and from genes to voltage; and, c) the time constants associated with the gene activities and the connections from voltage to genes, genes to voltage, and within the IN, representing the corresponding timescales of communication. Note that we set the target number of genes to six, since our goal was a model that could achieve correct regulation of a set of six genes, reflecting the biology we are modeling. However, we allowed the number of genes in the models to vary above six to accommodate the possibility that a larger set of genes may be required as a scaffold to achieve regulation of the smaller subset of six.

Model simulation

The initial conditions of the variables in each simulation were set as follows: $v = -9.2$mV (the unstable equilibrium point in the bistable cell) for all cells; $G_p = 1.0\ G_0$ for depolarized cells or $1.8\ G_0$ for hyperpolarized cells; the constant G_d was set to a fixed value of $1.5\ G_0$ for all cells under all input conditions. The values of all other variables including G_{ij} and g were set to 0. The model equations were integrated using the standard Euler method with a fixed step size of 0.01 for about 1000 steps. "Input voltage patterns" refers to the patterns set at initiation of the simulation by supplying the appropriate values (hyperpolarized or depolarized) to G_p for each cell to generate the desired voltage pattern when the tissue is decoupled from the GRNs, that is, in a pure bioelectric setting. This is analogous to the approach we use for *in vivo* experiments, controlling V_{mem} by manipulating appropriate ion channels[80] rather than voltage *per se*.

Pattern discrimination problem

While this approach is general enough to help understand spatial computation in tissues more broadly, we designed this model specifically to understand how *Xenopus* embryonic bioelectrical prepatterning leads to correct brain morphogenesis. To that end, we set up a minimal version of this problem for our model to solve, which we will henceforth refer to as the *pattern discrimination problem*. Can the model recognize the 'endogenous' (mimicking the normal *in vivo* pattern depicted in Figure 1A) input voltage pattern across cells? In our models, the ability to discriminate between the endogenous and aberrant patterns is manifested by activating a set of 6 genes in response to the endogenous pattern and repressing them in response to patterns that are known to result in abnormal brain morphologies. This recapitulates the results of *in vivo* experiments on the role of bioelectric contrast prepatterns in *Xenopus* brain development[78,80,87,88,97,105,177]). The input to the model is a voltage prepattern representing either the endogenous bioelectric contrast pattern (henceforth referred to as the "correct" pattern – Figures 1A and 1B) or its abnormal versions, namely fully depolarized or fully hyperpolarized (referred to as the "incorrect" patterns – Figures 1C and 1D). The output of the model is a gene expression pattern. As just described, the desired relationship between inputs and outputs is: 1) the correct voltage input pattern should activate the genes (positive activity); and 2) incorrect voltage input patterns should repress them (negative activity). Here positive activity is taken as a proxy for normal brain morphology and negative activity for abnormal morphology.

Machine learning search and model selection

Classical methods have not led to the discovery of generative models that provide insights, hypotheses, and predictions about what is necessary for the observed tissue-level processing of bioelectrical information. Given the complex recurrent nature of this phenomenon and the necessity to begin to develop Artificial Intelligence tools that assist human scientists in discovering models of morphogenetic regulation, we used machine learning methods to search for insight hypotheses on connectivity among model components and the associated parameters that enable a small model containing 24 cells (4x6 lattice) to solve the pattern discrimination problem (Figure 1F). The chosen model size both facilitates ease of training and analysis and also serves as a coarse-graining for biological tissues that are typically larger.[131,132] A rough estimation suggests that the Xenopus neural plate is ~43 cells X ~43 cells, and hence the bioelectric pattern which includes neural plate cells and surrounding ectoderm would be ~83 cells wide. Thus, the biological cell numbers lie within in the range tested in our model. We used a combination of genetic algorithm (GA) and gradient descent (GD) to train the models (Figure 2B). GA is analogous to biological evolution, whereas GD is analogous to learning during the lifetime of an animal. Our motivation for using this particular combination of search algorithms is not to emulate biological *learning*. Rather, using this mixture of evolutionary and ontogenetic timescales in our searches is a purely pragmatic means to discover a combination of the connectivity and parameters of a model that explains the physiological and functional data on bioelectric control of brain patterning. The overall search began with the GA search for an appropriate network structure (including the numbers of genes and regulatory connections), followed by GD to further fine-tune the parameters of the best network discovered by the GA, with the structure fixed. We used the standard microbial GA,[178] which is based on horizontal gene transfer and a simple tournament selection process: two randomly chosen individuals from the population are pitted against each other and the 'winner' transmits randomly chosen portions of its genetic material to the 'loser' based on a specified "infection rate". This process was repeated for thousands of iterations resulting in the gradual improvement of performance (Figure S1A). Owing to the inherent degeneracy of the parameter space of models of biological

systems,[110,179,180] we randomly chose one of the two best-performing models and further refined it using GD. We used a "resilient backpropagation"[181] flavor of GD for this purpose; this method relies only on the sign (not the magnitude) of the gradients for updating the parameters at every iteration. The resulting performance score was found to be significantly higher than expected for a randomly parameterized model (Figure S1B).

The particular optimization procedure we used for GD is known as "resilient backpropagation" that relies only on the sign (not the magnitude) of the gradients for updating the parameters at every iteration, with a learning rate of 0.01.

For the microbial GA search, we used a population of 40 individual "genotypes". Each genome codes for the parameters and the meta parameters of a single instance of the model. The parameters are the literals listed in red in Figure 2A coding the parameters of a model with a given size of the GRN and the IN, whereas the meta parameters, coding for the size of the GRN and the IN, are the following: number of genes (n_g), number of GRN edges (n_e), number of iGRN edges (n_i), number of GRN to Vmem edges (n_{gv}), and the number of Vmem to GRN edges (n_{vg}). Every gene in a genotype has a value in the range [0.01,1] that is linearly mapped to the corresponding model features according to their respective ranges, as follows: weights have a range of [-16,16], timeconstants [0.1,30], bias of the GRN nodes [-16,16], bias of the Vmem node [-100 mV,0 mV], Vmem gain [0,7], n_g [6,20], n_e has a range of $[n_g - 1, n_g(n_g - 1)]$, n_i has a range of $[1,2\, n_g^2]$, and both n_{gv} and n_{vg} have a range of $[1, n_g]$. We used a "geographical selection" method where the individuals (genotypes) are placed on a 1D ring, and only geographically close individuals are picked to compete in the tournaments. The size of selection-neighborhood, known as the "deme size", was set to 20%. The genome of the winner of a tournament was transmitted to the loser at a "crossover rate" set to 10%. The genome of the loser was then mutated at a "mutation rate" of 5% and reinserted into the population. All genotypes were randomly initialized (in the range [0.01,1]) during the first generation. This process was repeated for about 1200 generations.

Each genotype was evaluated by simulating the corresponding model and computing its performance at the end of the simulation. The initial conditions of the variables in each simulation were set as follows: $v = -9.2$mV (the unstable equilibrium point in the bistable cell) for all cells; $G_p = 1.0\ G_0$ for the cells that are meant to be depolarized and $G_p = 1.8\ G_0$ for the cells that are meant to be hyperpolarized, while constant G_d was set to a fixed value of $1.5\ G_0$ for all cells under all input conditions. Note that in the main text, the different input conditions, namely "depolarized", "endogenous" and "hyperpolarized" refer to voltage patterns, which in practice is achieved by setting appropriate values of G_p that would result in those patterns when the tissue is decoupled from the GRNs, that is, in a pure bioelectric setting. The values of all other variables including G_{ij} and g were set to 0. The model equations were integrated using the standard Euler method with a fixed step size of 0.01 for about 1000 steps.

Performance measures

The quality of a model and its components was evaluated using the following performance measures. The "discrimination error" of a gene is defined as the weighted mean distance between the observed spatial expression patterns and the target patterns corresponding to the endogenous, depolarized, and hyperpolarized input voltage patterns during the final stages of development. The "discrimination score" of a gene is defined as the change in the observed discrimination error relative to the maximum possible error. The "performance score" of a model is defined as the inverse of the ratio between the mean observed and the maximum possible gene discrimination errors. Both the gene discrimination and the model performance scores range between 0 and 1, with the extrema corresponding to the negative and positive discriminations and the resulting model performance and a value of 0.5 indicating the discrimination and performance expected from a randomly parameterized model. The above measures are mathematically defined as follows.

The discrimination score, $\widehat{R}(g_j)$, of a gene g_j is defined as the change in the observed discrimination error, $R_{\text{obs}}(g_j)$, relative to the maximum possible error, R_{max}, where the discrimination *error* is computed as the weighted mean distance between the observed gene expression pattern and the target patterns corresponding to the endogenous (E), depolarized (D), and hyperpolarized (H) input Vmems, $v(0) = (v_1(0), \ldots, v_{n_c}(0))$:

$$\widehat{R}\left(g_j\right) = \frac{R_{\text{max}} - R_{\text{obs}}\left(g_j\right)}{R_{\text{max}}}$$

$$R_{\text{obs}}\left(g_j\right) = 0.5 \times R_{\text{obs}}\left(\widehat{g}_j \middle| v(0) = E\right) + 0.25 \times \left[R_{\text{obs}}\left(\widehat{g}_j \middle| v(0) = D\right) + R_{\text{obs}}\left(\widehat{g}_j \middle| v(0) = H\right)\right];$$

$$R_{\text{obs}}\left(\widehat{g}_j \middle| v(0) = E\right) = \sqrt{\frac{\sum_{t=1}^{T}\sum_{i=1}^{n_c}\left(\widehat{g_{i,j}} - 1\right)^2}{T \times n_c}},$$

$$R_{\text{obs}}\left(\widehat{g}_j \middle| v(0) = H \text{or} D\right) = \sqrt{\frac{\sum_{t=1}^{T}\sum_{i=1}^{n_c}\left(\widehat{g_{i,j}} + 1\right)^2}{T \times n_c}}$$

$$R_{\text{max}} = 0.5 \times R_{\text{max}}(\widehat{g}|v(0) = E) + 0.25 \times [R_{\text{max}}(\widehat{g}|v(0) = D) + R_{\text{max}}(\widehat{g}|v(0) = H)];$$

$$R_{\max}(\hat{g}|v(0) = E) = \sqrt{\frac{\sum_{t=1}^{T}\sum_{i=1}^{n_c}(-1-1)^2}{T \times n_c}},$$

$$R_{\max}(\hat{g}|v(0) = H \text{or} D) = \sqrt{\frac{\sum_{t=1}^{T}\sum_{i=1}^{n_c}(1+1)^2}{T \times n_c}},$$

$$\Rightarrow R_{max} = 2$$

$$\hat{g}_j = g_j \mapsto [-1, 1] \text{ is the normalized gene expression}$$

The desired target gene expression patterns are the following: all-activated $(1, \ldots, 1)$ for E; and, all-repressed $(-1, \ldots, -1)$ for D and H. Thus, the value of $\widehat{R}(g_j)$ lies between 0 and 1, corresponding to the minimum and maximum possible discriminatory scores, while a value of 0.5 indicating a randomly expected score (no discrimination).

The performance score, P, of the model is defined as the complement of the ratio between the observed mean discrimination error, averaged over all genes, and the maximum possible discrimination error:

$$P = 1 - \sqrt{\frac{\sum_{j=1}^{n_g} R_{obs}^2\left(g_j\right)}{\frac{n_g}{R_{\max}}}}$$

Thus, the value of P lies between 0 and 1, corresponding to the minimum and maximum possible performance, while a value of 0.5 indicating the expected performance of a randomly parametrized model.

Causal integration analysis

To investigate the causal relationship between voltage pattern and gene expression in our model we employed the framework of "multi-time-scale causal influence" introduced in ref. 182. Specifically, the amount of first order causal influence (CI) exerted by the voltage pattern on gene expression is given by the following "causal derivatives".

The "Jacobian" tensor $J_{v\to g}(\tau)$ captures the sensitivity (positive or negative) of every gene in every cell, $g_{i,j}$, at time τ with respect to the voltage of every cell, i, as determined by the conductance of the polarizing ion channel, G_{p_k}, at time $t = 0$ for any $\tau \geq 1$. The larger the absolute value of an element of $J_{v\to g}(\tau)$, the more causally sensitive gene $g_{i,j}$ is at time τ to the initial voltage of cell k. It is defined as follows:

$$J_{v\to g}(\tau) = \left[\frac{\partial g_{i,j}(\tau)}{\partial G_{p_k}(0)}\right]_{n_c \times n_g \times n_c}; i, k = 1, \ldots, n_c; j = 1, \ldots, n_g \quad \text{(Equation 1)}$$

The "Hessian" tensor $H_{v\to g}(\tau)$, on the other hand, captures the second-order sensitivity (i.e., sensitivity of sensitivity, which can also be positive or negative) of every gene, $g_{i,j}$, in every cell, i, at time τ with respect to the voltages of every pair of cells, as determined by the conductance of the corresponding polarizing ion channels, (G_{p_l}, G_{p_k}), at time $t = 0$ for any $\tau \geq 1$. It is defined as follows:

$$H_{v\to g}(\tau) = \left[\frac{\partial^2 g_{i,j}(\tau)}{\partial G_{p_l}\partial G_{p_k}(0)}\right]_{n_c \times n_g \times n_c \times n_c}; i, k, l = 1, \ldots, n_c; j = 1, \ldots, n_g \quad \text{(Equation 2)}$$

The Hessian derivative can also be conceived as a measure of context-dependency – the extent to which the influence of a cell's voltage on gene expression depends on the context provided by the voltage of another cell. Alternatively, it can be interpreted as the amount of collective influence exerted by the voltages of pairs of cells on the gene expression. The larger the absolute value of an element of $H_{v\to g}(\tau)$, the more causally sensitive gene $g_{i,j}$ is at time τ to the initial voltages of the pair of cells l and k. Here we used the conductance of the polarizing ion channel, G_p, of a cell as a surrogate for its voltage since the former is fixed throughout the simulation and fully determines the voltage profile of tissue[88] (Figure 5A).

A Hessian *network* is a graphical representation of the Hessians defined for a particular gene $g_{i,j}$ in all cells $j \in \{1, \ldots, n_c\}$ at any given time τ. In other words, it is a network of the second-order influence of voltage on gene expression at any given time τ; this network is hidden in the sense that it can only be inferred via analysis such as causal integration.

Reconstructing gene activity timeseries using causal derivatives

The following equations were fit to reconstruct the normalized gene expression timeseries of the model.

$$\widehat{g_{i,j}(t+1)} = \widehat{g_{i,j}(t)} + \sum_{k=1}^{n_c} c_J[j] \frac{\partial g_{i,j}(t)}{\partial G_{p_k}(0)} \quad \text{(Equation 3)}$$

$$\widehat{g_{i,j}(t+1)} = \widehat{g_{i,j}(t)} + \sum_{l=1}^{n_c} \sum_{k=1}^{n_c} c_H[j] \frac{\partial^2 g_{i,j}(t)}{\partial G_{p_l} \partial G_{p_k}(0)} \quad \text{(Equation 4)}$$

Here, $\widehat{g_{i,j}(0)} = 0$, and[.] refers to the gene index. The above equations were optimized by either fitting the Jacobian coefficient c_J or the Hessian coefficient c_H, both vectors of length $1 \times n_g$. Though the above equations look similar to the conventional Euler integration they differ from the latter in two ways: one, in the kind of differential derivatives they involve, and two, the number of terms. This is explained in more detail below.

Conventional Euler integration involves derivatives of the integrated variable taken with respect to time. Here we present an alternative method using derivatives taken with respect to the parameter of the modeled system (e.g., external inputs, variables that change relatively slowly, etc.). This method can be understood with the help of a toy example as follows. Consider the following ODE consisting of a single recurrent variable x controlled by parameter a:

$$\frac{dx}{dt} = a^2 x$$

Even though this equation can be analytically solved as $x_t = c_1 e^{a^2 t}$, where c_1 is the integration constant, we shall consider the approximated timeseries $x_1, x_2, \ldots, x_t, \ldots$ generated by Euler method:

$$x_{t+1} = x_t + \frac{dx_t}{dt}\delta t = x_t + a^2 x_t \delta t \quad \text{(Equation 5)}$$

The above can also be expressed in equivalent parameter-based Euler (PE) integration consisting of the terms $\frac{dx_t}{da}$, representing the sensitivity of the variable to the controlling parameter. This shall allow us to understand the role of the parameter in reconstructing the timeseries of the variable x. For example,

$$x_1 = x_0 + a^2 x_0 \delta t$$

$$x_2 = x_0 + 2a^2 x_0 \delta t + a^4 x_0 \delta t \quad \text{(Equation 6)}$$

$$= x_1 + \frac{1}{2}\frac{dx_1}{da}(a) + \frac{1}{2}\frac{d^2 x_1}{da^2}\left(a^4 \delta t\right)$$

Where:

$$\frac{dx_1}{da} = 2a x_0 \delta t; \frac{d^2 x_1}{da^2} = 2 x_0 \delta t$$

Likewise,

$$x_3 = x_2 + \frac{1}{4}\frac{dx_2}{da}(a) + \frac{1}{4}\frac{d^2 x_2}{da^2}\left(a^4 \delta t\right) - \frac{1}{6}\frac{d^3 x_2}{da^3}\left(a^5 \delta t\right) + \frac{1}{12}\frac{d^4 x_2}{da^4}\left(a^6 \delta t\right) \quad \text{(Equation 7)}$$

$$x_4 = x_3 + \frac{1}{6}\frac{dx_3}{da}(a) + \frac{1}{6}\frac{d^2 x_3}{da^2}\left(a^4 \delta t\right) - \frac{1}{12}\frac{d^3 x_3}{da^3}\left(a^5 \delta t\right) + \frac{1}{36}\frac{d^4 x_3}{da^4}\left(a^6 \delta t\right) - \frac{1}{72}\frac{d^5 x_3}{da^5}\left(a^7 \delta t\right) + \frac{1}{120}\frac{d^6 x_3}{da^6}\left(a^8 \delta t\right) \quad \text{(Equation 8)}$$

$$\vdots$$

Given that Equation 5 can be rewritten as $x_t = x_0(1+a^2\delta t)^t$, the derivatives of x_t with respect to a can be expressed in generic terms as follows:

$$\frac{dx_t}{da} = 2x_0\left((a\delta t)t\left(1+a^2\delta t\right)^{t-1}\right)$$

$$\frac{d^2 x_t}{da^2} = 2x_0\left(2(a\delta t)^2 t(t-1)\left(1+a^2\delta t\right)^{t-2} + t\left(1+a^2\delta t\right)^{t-1}\right)$$

$$\frac{d^3x_t}{da^3} = 2x_0\left(4(a\delta t)^3 t(t-1)(t-2)\left(1+a^2\delta t\right)^{t-3} + 6\left(a\delta t^2\right)t(t-1)\left(1+a^2\delta t\right)^{t-2}\right)$$

$$\vdots$$

As can be seen, the number of terms appearing in the PE expressions increases as the time index of x_t increases. Moreover, the coefficient of the leading term, given by, $2^n\prod_{i=0}^{n}(t-i)$, where n refers to the order of differentiation, grows nonlinearly with t and n. This is the reason why the coefficients of the derivatives in the PT expansions shrink nonlinearly (to compensate for the nonlinear growth of the derivatives) with time for the higher order derivatives, as can be seen in the expressions above.

Even though it may seem that time-dependent PT coefficients are required to reconstruct the timeseries of the variable, static coefficients, even with just the 1st and 2nd order terms, are sufficient for a qualitative reconstruction. This can be achieved by optimizing what we call "pruned timeless parameter-based Euler" (PTPE) equation:

$$\widehat{x_{t+1}} = \widehat{x_t} + c_1\frac{dx_t}{da} + c_2\frac{d^2x_t}{da^2} \quad \text{(Equation 9)}$$

Here, the coefficients c_1 and c_2 are fit to the equation, given the observations $\{x_t\}$. This is the motivation behind Equations 3 and 4 above.

Microinjections

Capped synthetic mRNAs generated using the mMessage mMachine kit (Ambion) were dissolved in nuclease-free water and injected into embryos immersed in 3% Ficoll using standard protocols.[175] Each injection delivered ~ 1–2 ng of mRNA (per blastomere) into the middle of a cell in the animal hemisphere of embryos at the four-cell stage. Constructs used were: *Kv1.5*[138], *dominant-negative Kir6.1 pore mutant – DN-Kir6.1p*,[33] and *β-galactosidase*.

Membrane voltage imaging with $DiBAC_4(3)$:CC2-DMPE

$DiBAC_4(3)$ and CC2-DMPE ratiometric reporter dyes (Invitrogen) were used as per the standard protocol.[183] Briefly, CC2-DMPE stock (5 mM) was dissolved 1:1000 in 0.1X MMR and the embryos were incubated in the dark for 1 h followed by five washes with 0.1X MMR. $DiBAC_4(3)$ stock (1.9 mM) was dissolved 1:1000 in 0.1X MMR and the CC2-DMPE stained embryos were then incubated in the dark in this solution for 30 min followed by five washes with 0.1X MMR, and then visualized under an Olympus BX-61 microscope equipped with a Hamamatsu ORCA AG CCD camera controlled by MetaMorph software (Molecular Devices). NIH Fiji software was used to quantify the fluorescence intensities of CC2-DMPE:$DiBAC_4$ images. As done previously,[22,88,112,184] static whole-cell electrophysiological recordings of membrane voltage from neural plate cells and flanking ectodermal cells were then used as calibration points for the voltage reporter dye images, and fluorescent intensities were analyzed against these calibration points to approximate membrane voltages at different points within developing embryos.

β-Galactosidase Enzymatic detection

Tadpoles were fixed at stage 45 for 30 min in aldehyde fixative MEMFA at room temperature, washed twice in PBS with 2 mM $MgCl_2$, and stained with X-gal (ThermoFisher) staining solution at 37°C for at least 3 h. Tadpoles were then rinsed three times in PBS followed by dehydration through sequential incubation in 25%, 50%, 75% and 100% methanol. Tadpoles were then incubated in 30% H_2O_2 in methanol overnight for bleaching, washed in 100% methanol, and sequentially rehydrated (100%, 75%, 50%, 25%) to PBS and imaged using a standard Nikon dissection microscope with Lumenera Infinity3 camera and associated Infinity imaging software.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were performed using GraphPad Prism. At least three independent experiments were conducted with N > 50 embryos for each treatment group, using embryos collected from multiple animals across independent clutches. Data were analyzed by ANOVA (with Tukey's multiple comparison post-test). All statistical details of experiments are reported in the figure legends of figures of respective experiments.