Science Signaling
AAAS

www.sciencesignaling.org/cgi/content/full/8/397/ra99/DC1

Supplementary Materials for

Serotonergic regulation of melanocyte conversion: A bioelectrically regulated network for stochastic all-or-none hyperpigmentation

Maria Lobikin, Daniel Lobo, Douglas J. Blackiston, Christopher J. Martyniuk, Elizabeth Tkachenko, Michael Levin*

*Corresponding author. E-mail: michael.levin@tufts.edu

Published 6 October 2015, *Sci. Signal.* **8**, ra99 (2015)
DOI: 10.1126/scisignal.aac6609

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencesignaling.org/cgi/content/full/8/397/ra99/DC1)

Data S2 (Microsoft Excel format). Results of the training set and validation set experiments and the performance of the model for each.

Text S1. Model implementation, simulation, and evaluation

We employed a modeling approach using stochastic ordinary differential equations (ODEs) based on Hill kinetics to describe the signaling mechanism of normal and hyperpigmented tadpoles under experimental perturbations. Our goal was to identify a signaling network, which produced the stochastic phenotypic outcomes (normal pigmented or hyperpigmented) from given experimental initial starting conditions and under different pharmacological interventions. Each ODE in the model represents the production rate of a product in the network. Products in the network can represent a signaling molecule, a pharmacological compound (drug) used in the experiment treatments, or the pigmentation level in the tadpole (the phenotype). The concentrations of pharmacological compounds are assumed to be constant during an experiment and are set to 1 in the model if present in the experimental condition being simulated and 0 if absent. The production rate of the products was modeled as a linear relation among three terms: production, decay, and noise. As described in the Materials and Methods, the production term contained a combination of Hill functions (one for each regulatory relationship), the decay term contained a decay constant (modeling an exponential decay), and the noise term was modeled as a Gaussian random noise. The model also contained the initial concentrations (initial conditions for the ODEs) of every product, except the drugs, which were set according to the experiment treatments, and the pigmentation product, which was set to zero.

We implemented an Euler finite difference method (*153*) to simulate the pharmacological experiments in the dataset. The simulator takes as input a model and an experiment. The model defines the system of ODEs and initial conditions of the signaling products and the experiment defines the initial conditions of the drugs. With this information, the simulator integrates the ODE system for a fixed amount of time. The final level of the pigmentation product indicates if the experiment results in a non-hyperpigmented (pigmentation level below 0.1), intermediate (pigmentation level between 0.1 and 0.9), or hyperpigmented (pigmentation level above 0.9) tadpole. Different runs with the same initial conditions (same experiment) can result in different pigmentation levels due to the stochastic terms in the ODE system.

To evaluate the error of an ODE-based network model, we simulated each experiment 100 times and scored the resultant level of hyperpigmentation. To penalize networks that did not converge to a steady state, we

included a concentration change penalty that was applied when the maximum concentration change rate in the last time step of a simulation was higher than a parameter threshold (μ).

We defined the pigmentation distance between a pigmentation level p and a set of target pigmentation levels P as the minimum logarithmic distance between each of them within a tolerance ε:

$$\mathrm{d}(p, P) = \min_{p' \in \mathrm{P}} (\log(1 + (\mathrm{abs}(p - p') - \varepsilon)^{+})$$

We then defined the phenotypic error of a signaling network model M for an experiment e as the average pigmentation distance and final concentration change over 100 simulations:

$$\mathrm{error}_{\mathrm{phenot}}(M, e) = \frac{1}{100} \sum_{i=1}^{100} \left(\mathrm{d}(p_{i,e}^{M}, \mathrm{P_e}) + \left(\Delta_{i,e}^{M} - \mu \right)^{+} \right)$$

where $p_{i,e}^{M}$ is the resultant pigmentation level of simulation i with model M for the experiment e, $\mathrm{P_e}$ is the set of pigmentation levels resulting from experiment e in the in vivo assays, $\Delta_{\mathrm{i,e}}^{M}$ is the average concentration change rate in the last time step of the simulation i of experiment e with the model M, and μ is the penalty concentration change threshold.

Next, we defined the penetrance error of a signaling model M for an experiment e as the square error between the in-silico and in-vivo outcomes for each of the experiments:

$$\mathrm{error}_{\mathrm{penet}}(M, e) = (H_e^M - H_e)^2 + (N_e^M - N_e)^2$$

where H_e^M and N_e^M are the resultant frequency of hyperpigmented and nonhyperpigmented phenotypes in the simulation of experiment e with the model M, and H_e and N_e are the resultant frequency of hyperpigmented and nonhyperpigmented phenotypes in the in-vivo assay of experiment e.

Finally, we defined the error of a signaling network model M for a set E of n experiments as the average sum of phenotypic and penetrance errors for each of the experiments:

$$\mathrm{error}(M, E) = \frac{1}{n} \sum_{i=1}^{n} \left(\mathrm{error}_{\mathrm{phenot}}(M, e_i) + \mathrm{error}_{\mathrm{penet}}(M, e_i) \right)$$

Text S2. Method for reverse-engineering a stochastic model of hyperpigmentation

We created an automated method based on (*48*) to reverse engineer a stochastic signaling network model of hyperpigmentation directly from the experimental data. The method takes as input a dataset of experiments (each of them specifying a set of applied pharmacological drugs and the resultant phenotypic frequencies) and an optional set of known regulatory interactions required to exist in any candidate model during the search. The method infers any necessary unknown products and links and all the kinetic parameters of the network.

The method uses an evolutionary algorithm (*154*) approach, maintaining a population of candidate models (signaling networks based on a system of ODEs) that evolve in parallel for a fixed number of generations. The algorithm starts by creating an initial population of candidate networks, each of them including all the products and regulations defined in the input dataset of experiments (drugs and their affected signaling products), as well as the input set of known regulatory interactions. All the numerical parameters in the networks are randomly chosen; regulatory links can be either activators or inhibitors with equal probability (except the drug links, which are predefined as activator or inhibitor), and they can be aggregated as necessary or sufficient with equal probability. Then, the networks in the initial population are simulated for all experiments (each experiment defining an initial condition) and their resulting error scored and stored.

The algorithm proceeds iteratively by (i) creating new networks from existing ones in the population through crossover and mutation operators, (ii) simulating and scoring the new networks, and (iii) selecting the best networks from the population for the next generation. The crossover operator mixes randomly two parent networks to produce two new offspring networks. Common products between the two networks are copied randomly (one for each offspring network) and unique products are distributed randomly between the two offspring networks. The regulatory links of a product are copied together with the product; if the other regulatory product of the copied link does not exist in the new offspring network, it is randomly replaced by another regulatory product. The mutation operator changes the signaling network randomly. Product and link parameters can be replaced by a new random value; products and links can be duplicated. The regulated and regulator products of the new duplicated link are chosen randomly. Products and links can also be randomly

deleted, except the drugs and their links, as well as the required known signaling products and the input set of known links, which cannot be deleted (although their parameters can be randomly mutated).

We implemented the algorithm in a cluster computer with 22 CPUs using the island distribution approach (*155*), which improves performance and maintains a diverse set of models in the population. The algorithm used 32 parallel subpopulations with 64 signaling networks each, and a total of 10,000 generations during a search. To compensate for the trend of the subpopulations to saturate, all subpopulations were randomly paired and their signaling networks shuffled randomly between them for every 250 generations. We used a deterministic crowding selection method (*156*) with 75% crossover, 1% parameter change mutation, 1% duplication mutation, and 1.5% deletion mutation probabilities. All the numerical parameters were randomly chosen with a uniform distribution on the interval [0, 1], except the Hill coefficients, which were chosen with a uniform distribution on the interval [1, 5]. The simulations used a Gaussian random noise of 0.0 mean and 10^{-4} standard deviation. The scoring function used a tolerance ε of 0.1 and a penalty concentration change threshold of 10^{-4}.

The simulation and search method was implemented in C++ using the Standard Library. Visualizations used the Qt libraries (The Qt Company Ltd.) and the Qwt library (Uwe Rathmann and Josef Wilgen).

Text S3. System of equations and kinetic parameters for the reverse-engineered model

$$\frac{dHP}{dt} = 0.53 \cdot \min\left(\frac{HP^{1.6}}{0.028^{1.6} + HP^{1.6}}, \max\left(\frac{CREB^{3.6}}{0.44^{3.6} + CREB^{3.6}}, \frac{0.28^{4.6}}{0.28^{4.6} + a^{4.6}}\right)\right) - 0.57 \cdot HP + \xi(t)$$

$$\frac{d5HT}{dt} = 0.65 \cdot \min\left(\frac{0.51^3}{0.51^3 + SERT^3}, \frac{VMAT^{1.7}}{0.23^{1.7} + VMAT^{1.7}}, \max\left(\frac{HP^{4.3}}{0.44^{4.3} + HP^{4.3}}, \frac{0.37^{4.5}}{0.37^{4.5} + R5^{4.5}}, \frac{MSH^{2.2}}{0.21^{2.2} + MSH^{2.2}}, \frac{Med5HT^{2.1}}{0.84^{2.1} + Med5HT^{2.1}}, \frac{b^{3.6}}{0.44^{3.6} + b^{3.6}}\right)\right) - 0.78 \cdot 5HT + \xi(t)$$

$$\frac{dSERT}{dt} = 0.86 \cdot \min\left(\frac{IVM^{1.6}}{0.028^{1.6} + IVM^{1.6}}, \frac{0.24^{1.8}}{0.24^{1.8} + MSH^{1.8}}, \frac{0.28^{3.3}}{0.28^{3.3} + Fluoxetine^{3.3}}, \frac{0.24^{2.2}}{0.24^{2.2} + R1^{2.2}}\right) - 0.43 \cdot SERT + \xi(t)$$

$$\frac{dR1}{dt} = 0.34 \cdot \min\left(\frac{0.76^4}{0.76^4 + Cyanopindolol^4}, \frac{0.38^{4.3}}{0.38^{4.3} + MTP^{4.3}}, \max\left(\frac{5HT^{1.9}}{0.27^{1.9} + 5HT^{1.9}}, \frac{cAMP^{1.4}}{0.21^{1.4} + cAMP^{1.4}}\right)\right) - 0.19 \cdot R1 + \xi(t)$$

$$\frac{dR2}{dt} = 0.56 \cdot \min\left(\frac{0.29^{1.2}}{0.29^{1.2} + R1^{1.2}}, \frac{0.17^{4.9}}{0.17^{4.9} + MTP^{4.9}}, \max\left(\frac{5HT^{3.7}}{0.0095^{3.7} + 5HT^{3.7}}, \frac{0.64^{3.6}}{0.64^{3.6} + Altanserin^{3.6}}\right)\right) - 0.25 \cdot R2 + \xi(t)$$

$$\frac{dR5}{dt} = 0.43 \cdot \min\left(\frac{0.76^{2.7}}{0.76^{2.7} + SB699551^{2.7}}, \frac{0.27^{3.4}}{0.27^{3.4} + MTP^{3.4}}, \max\left(\frac{5HT^{2.4}}{0.54^{2.4} + 5HT^{2.4}}, \frac{0.67^{4.1}}{0.67^{4.1} + SERT^{4.1}}, \frac{c^{3.7}}{0.0095^{3.7} + c^{3.7}}\right)\right) - 0.23 \cdot R5 + \xi(t)$$

$$\frac{dMSH}{dt} = 0.077 \cdot \min\left(\frac{SHU9119^{4.8}}{0.061^{4.8} + SHU9119^{4.8}}, \max\left(\frac{SERT^{3.4}}{0.88^{3.4} + SERT^{3.4}}, \frac{R2^{2.6}}{0.69^{2.6} + R2^{2.6}}, \frac{R5^{1.7}}{0.43^{1.7} + R5^{1.7}}, \frac{0.28^{2.9}}{0.28^{2.9} + MRIF^{2.9}}, \frac{0.65^{2.9}}{0.65^{2.9} + cAMP^{2.9}}\right)\right) - 0.68 \cdot MSH + \xi(t)$$

$$\frac{dVMAT}{\partial t} = 0.74 \cdot \frac{0.23^{3.5}}{0.23^{3.5} + Reserpine^{3.5}} - 0.83 \cdot VMAT + \xi(t)$$

$$\frac{dcAMP}{dt} = 0.48 \cdot \min\left(\frac{MSH^{4.4}}{0.00028^{4.4} + MSH^{4.4}}, \frac{Forskolin^{1.2}}{0.79^{1.2} + Forskolin^{1.2}}, \frac{0.53^{2.5}}{0.53^{2.5} + 2'5' - DDA^{2.5}}\right) - 0.5 \cdot cAMP + \xi(t)$$

$$\frac{dCREB}{dt} = 0.81 \cdot \min\left(\frac{0.65^5}{0.65^5 + R1^5}, \frac{0.24^{2.7}}{0.24^{2.7} + MSH^{2.7}}, \frac{0.027^{1.4}}{0.027^{1.4} + ACREB^{1.4}}, \frac{0.72^{4.3}}{0.72^{4.3} + a^{4.3}}, \max\left(\frac{CREB^{2.7}}{0.49^{2.7} + CREB^{2.7}}, \frac{PKA^{3.3}}{0.41^{3.3} + PKA^{3.3}}, \frac{VP16^{2.7}}{0.026^{2.7} + VP16^{2.7}}\right)\right) - 0.82 \cdot CREB + \xi(t)$$

$$\frac{dPKA}{dt} = 0.31 \cdot \min\left(\frac{SERT^{4.1}}{0.31^{4.1} + SERT^{4.1}}, \frac{cAMP^{2.4}}{0.21^{2.4} + cAMP^{2.4}}, \frac{0.12^{1.1}}{0.12^{1.1} + H89^{1.1}}\right) - 0.66 \cdot PKA + \xi(t)$$

$$\frac{da}{dt} = 0.96 \cdot \min\left(\frac{0.13^{2.6}}{0.13^{2.6} + SERT^{2.6}}, \frac{0.15^{1.8}}{0.15^{1.8} + R1^{1.8}}, \frac{R2^{1.4}}{0.092^{1.4} + R2^{1.4}}\right) - 1 \cdot a + \xi(t)$$

$$\frac{db}{dt} = 0.041 \cdot \max\left(\frac{0.27^{2.8}}{0.27^{2.8} + R2^{2.8}}, \frac{0.14^{4.4}}{0.14^{4.4} + b^{4.4}}\right) - 0.31 \cdot b + \xi(t)$$

$$\frac{dc}{dt} = 0.25 - 0.48 \cdot c + \xi(t)$$

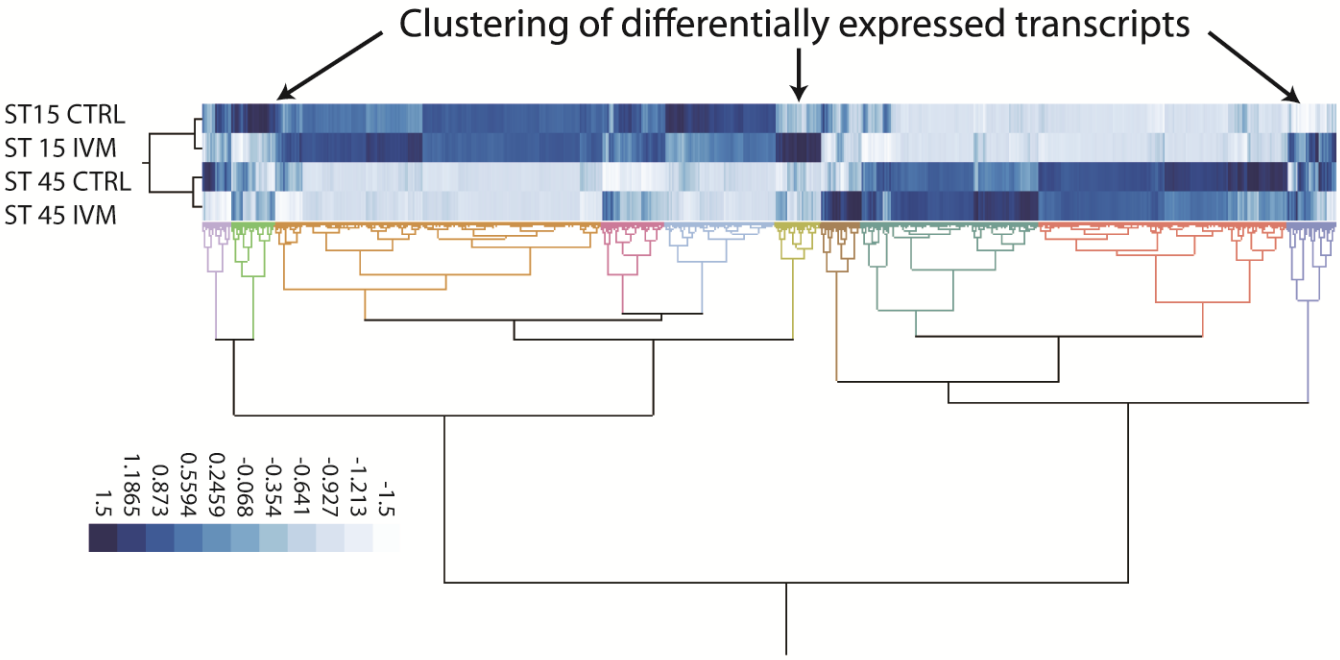

Fig. S1. Hierarchical clustering of the differentially regulated transcripts. Samples were subjected to hierarchical clustering (JMP Genomics v6) using the Fast Ward algorithm. Two Way Clustering was conducted (gene and treatment) and rows were first centered to a mean of zero (0) prior to analysis. Rows were also scaled to a variance of 1 prior to analysis to ensure that each row was scaled the same. Column standardization was also performed and columns centered to generate Z-scores using the standard deviation. This method subtracts the mean for each column and then divides by the column standard deviation. ST15 CTRL, stage 15 control embryos; ST 15 IVM, stage 15 ivermectin-exposed embryos; ST 45 CTRL, stage 45 control embryos; ST 45 IVM, stage 45 ivermectin-exposed embryos.

Table S1. A list of the 45 transcripts differentially expressed by stage 15 in response to ivermectin. The data are provided as an Excel file. The *Xenopus* gene name or EST and fold change in ivermectin-exposed embryos compared to untreated embryos are indicated. The data can also be downloaded from NCBI Gene Omnibus as accession number GSE70834.

Table S2. A list of the 517 transcripts differentially expressed by stage 45 in response to ivermectin. The data are provided as an Excel file. The *Xenopus* gene name or EST and fold change in ivermectin-exposed embryos compared to untreated embryos are indicated. The data can also be downloaded from NCBI Gene Omnibus as accession number GSE70834.

Table S3. **Enriched GO affected in stage 45 embryos.** FDR q is the multiple testing correction using the Benjamini and Hochberg (*157*) method. Enrichment score (ES) is calculated using the formula ES = (b/n) / (B/N), where N is the total number of genes; B is the total number of genes associated with a specific GO term; n is the number of genes in the top of the user's input list or in the target set when appropriate; b is the number of genes in the intersection.

GO term	GO accession	P-value	FDR q-value	Enrichment Score (N, B, n, b)
regulation of acute inflammatory response	GO:0002673	7.35E-10	8.17E-06	12.06 (8327,36,211,11)
regulation of complement activation	GO:0030449	8.23E-10	4.57E-06	21.05 (8327,15,211,8)
regulation of humoral immune response	GO:0002920	1.36E-09	5.03E-06	16.14 (8327,22,211,9)
regulation of protein activation cascade	GO:2000257	2.98E-09	8.27E-06	18.57 (8327,17,211,8)
complement activation	GO:0006956	7.81E-09	1.74E-05	13.66 (8327,26,211,9)
carboxylic acid metabolic process	GO:0019752	1.10E-08	2.03E-05	2.77 (8327,528,211,37)
oxoacid metabolic process	GO:0043436	2.09E-08	3.33E-05	2.61 (8327,590,211,39)
organic acid metabolic process	GO:0006082	2.90E-08	4.02E-05	2.58 (8327,597,211,39)
regulation of inflammatory response	GO:0050727	4.40E-08	5.43E-05	5.31 (8327,119,211,16)
regulation of defense response	GO:0031347	8.53E-08	9.48E-05	3.60 (8327,252,211,23)
small molecule metabolic	GO:0044281	3.89E-07	3.93E-04	1.75 (8327,1603,211,71)

process				
single-organism process	GO:0044699	4.33E-07	4.01E-04	1.22 (8327,5747,211,177)
monocarboxylic acid metabolic process	GO:0032787	4.44E-07	3.80E-04	3.29 (8327,276,211,23)
complement activation, alternative pathway	GO:0006957	5.25E-07	4.17E-04	24.67 (8327,8,211,5)
protein activation cascade	GO:0072376	6.29E-07	4.66E-04	8.66 (8327,41,211,9)
defense response	GO:0006952	8.08E-07	5.61E-04	2.55 (8327,496,211,32)
regulation of response to external stimulus	GO:0032101	1.61E-06	1.05E-03	3.06 (8327,297,211,23)
humoral immune response	GO:0006959	2.58E-06	1.60E-03	7.40 (8327,48,211,9)
single-organism metabolic process	GO:0044710	3.46E-06	2.03E-03	1.48 (8327,2558,211,96)
carbohydrate biosynthetic process	GO:0016051	3.51E-06	1.95E-03	5.09 (8327,93,211,12)
regulation of response to wounding	GO:1903034	5.43E-06	2.87E-03	3.74 (8327,169,211,16)
cytolysis	GO:0019835	6.83E-06	3.45E-03	16.44 (8327,12,211,5)
hexose biosynthetic process	GO:0019319	8.37E-06	4.04E-03	7.52 (8327,42,211,8)
positive regulation of immune response	GO:0050778	8.49E-06	3.93E-03	3.29 (8327,216,211,18)
monosaccharide metabolic process	GO:0005996	8.85E-06	3.94E-03	4.03 (8327,137,211,14)
ion transport	GO:0006811	1.10E-05	4.69E-03	2.29 (8327,534,211,31)
alpha-amino acid metabolic process	GO:1901605	1.34E-05	5.52E-03	3.89 (8327,142,211,14)
activation of immune response	GO:0002253	1.40E-05	5.56E-03	3.47 (8327,182,211,16)
glutamine family amino acid metabolic process	GO:0009064	2.00E-05	7.66E-03	6.72 (8327,47,211,8)
monosaccharide biosynthetic process	GO:0046364	2.35E-05	8.70E-03	6.58 (8327,48,211,8)
glutamine metabolic process	GO:0006541	2.43E-05	8.72E-03	13.15 (8327,15,211,5)
iron ion homeostasis	GO:0055072	3.20E-05	1.11E-02	6.31 (8327,50,211,8)
innate immune response	GO:0045087	3.28E-05	1.11E-02	2.62 (8327,332,211,22)
single-organism biosynthetic process	GO:0044711	4.38E-05	1.43E-02	2.02 (8327,685,211,35)
gluconeogenesis	GO:0006094	4.69E-05	1.49E-02	7.08 (8327,39,211,7)
complement activation,	GO:0006958	4.81E-05	1.48E-02	11.61 (8327,17,211,5)

classical pathway				
immune effector process	GO:0002252	5.22E-05	1.57E-02	3.27 (8327,181,211,15)
immune response	GO:0006955	6.83E-05	2.00E-02	2.32 (8327,425,211,25)
positive regulation of immune system process	GO:0002684	7.68E-05	2.19E-02	2.61 (8327,302,211,20)
proton transport	GO:0015992	7.74E-05	2.15E-02	4.93 (8327,72,211,9)
interaction with host	GO:0051701	8.45E-05	2.29E-02	5.54 (8327,57,211,8)
hydrogen transport	GO:0006818	8.64E-05	2.29E-02	4.87 (8327,73,211,9)
cellular response to interferon-gamma	GO:0071346	9.01E-05	2.33E-02	6.42 (8327,43,211,7)
monovalent inorganic cation homeostasis	GO:0055067	9.01E-05	2.28E-02	6.42 (8327,43,211,7)
single-organism carbohydrate metabolic process	GO:0044723	9.78E-05	2.42E-02	2.50 (8327,332,211,21)
immune system process	GO:0002376	1.03E-04	2.49E-02	1.89 (8327,773,211,37)
cation homeostasis	GO:0055080	1.07E-04	2.53E-02	2.94 (8327,215,211,16)
regulation of immune effector process	GO:0002697	1.19E-04	2.75E-02	3.88 (8327,112,211,11)
regulation of pH	GO:0006885	1.29E-04	2.92E-02	7.40 (8327,32,211,6)
cellular amino acid metabolic process	GO:0006520	1.29E-04	2.86E-02	2.68 (8327,265,211,18)
positive regulation of humoral immune response	GO:0002922	1.54E-04	3.37E-02	23.68 (8327,5,211,3)
transition metal ion homeostasis	GO:0055076	1.55E-04	3.32E-02	5.09 (8327,62,211,8)
energy coupled proton transmembrane transport, against electrochemical gradient	GO:0015988	1.85E-04	3.87E-02	8.97 (8327,22,211,5)
ATP hydrolysis coupled proton transport	GO:0015991	1.85E-04	3.80E-02	8.97 (8327,22,211,5)
regulation of immune response	GO:0050776	1.89E-04	3.82E-02	2.38 (8327,348,211,21)
response to interferon-gamma	GO:0034341	2.11E-04	4.19E-02	5.64 (8327,49,211,7)
transmembrane transport	GO:0055085	2.21E-04	4.31E-02	2.12 (8327,485,211,26)
glucose metabolic process	GO:0006006	2.27E-04	4.35E-02	3.91 (8327,101,211,10)
ferric iron transport	GO:0015682	2.31E-04	4.35E-02	8.58 (8327,23,211,5)
transferrin transport	GO:0033572	2.31E-04	4.28E-02	8.58 (8327,23,211,5)

trivalent inorganic cation transport	GO:0072512	2.31E-04	4.21E-02	8.58 (8327,23,211,5)
hexose metabolic process	GO:0019318	2.37E-04	4.25E-02	3.59 (8327,121,211,11)
organic anion transport	GO:0015711	2.43E-04	4.29E-02	3.15 (8327,163,211,13)

Table S4. Cancer-related genes in humans of the homologs of the genes differentially expressed in stage 45 *Xenopus* tadpole embryos after depolarizing ivermectin treatment. *Xenopus* genes were matched to the human homologs and then Oncomine was used to identify associated cancers. Fold change represents the difference in transcript abundance compared to control untreated embryos.

Name	Description	Related cancer(s) (according to Oncomine)	Fold Change
AGR2	Anterior gradient homolog 2 (Xenopus laevis)	Breast Cancer	-2.79
ALB	Albumin	Breast Cancer	-11.91
ATF5	Activating transcription factor 5	Lymphoma, Breast Cancer	-2.27
ATP6AP1	ATPase, H+ transporting, lysosomal accessory protein 1	Leukemia	-4.07
B2M	Beta-2-microglobulin	Brain & CNS Cancer	2.43
BCL2L10	BCL2-like 10 (apoptosis facilitator)	Breast Cancer	2.15
BIRC2	Baculoviral IAP repeat containing 2	Brain & CNS Cancer, Myeloma	2.6
C2orf40	Chromosome 2 open reading frame 40	Breast Cancer	2.56
CARTPT	CART prepropeptide	Breast Cancer	-2.11
CAT	Catalase	Breast Cancer, Lung Cancer, Leukemia	-3.31
CEBPB	CCAAT/enhancer binding protein (C/EBP), beta	Colorectal, Lymphoma	2.11
CTSL1	Cathepsin L1	Lymphoma	-2.73
CTSS	Cathepsin S	Kidney	2.71
CYP3A4	Cytochrome P450, family 3, subfamily A, polypeptide 4	Breast Cancer	2.11
DPYS	Dihydropyrimidinase	Breast Cancer, Leukemia	-2.53
EGLN3	egl nine homolog 3 (*C. elegans*)	Kidney	5.1
EIF2S1	Eukaryotic translation initiation factor 2, subunit 1 alpha, 35kDa	Colorectal, Sarcoma	2.03
EXT1	Exostosin 1	Head and Neck	-2.39
GPNMB	Glycoprotein (transmembrane) nmb	Lymphoma, Melanoma	-3.18
GSTP1	Glutathione S-transferase pi 1	Prostate Cancer	3.61
IFI30	Interferon, gamma-inducible protein 30	Breast Cancer	4.24
IKBKE	Inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase epsilon	Breast Cancer	2.76
KIF20A	Kinesin family member 20A	Breast Cancer, Sarcoma	2.32
MKI67	Antigen identified by monoclonal antibody Ki-67	Lung Cancer	3.95
MMP1	Matrix metallopeptidase 1 (interstitial collagenase)	Colorectal	3.75
MYD88	Myeloid differentiation primary response gene (88)	Lung Cancer	-4.05
NPY	Neuropeptide Y	Leukemia	-4.05

NR1H4	Nuclear receptor subfamily 1, group H, member 4	Colorectal	-2.39
P4HB	Prolyl 4-hydroxylase, beta polypeptide	Brain & CNS Cancer, Lymphoma	2.24
PDE5A	Phosphodiesterase 5A, cGMP-specific	Colorectal	-2.16
PFKFB3	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3	Breast Cancer, Leukemia	2
PMEL	Premelanosome protein	Melanoma	2.91
POMC	Pro-opiomelanocortin	Brain & CNS Cancer, Lung Cancer	3.59
PPARG	Peroxisome proliferator-activated receptor gamma	Breast Cancer, Sarcoma	2.2
PRKCD	Protein kinase C, delta	Leukemia	2.3
PRSS2	Protease, serine, 2 (trypsin 2)	Cervical Cancer	-3.67
SIRT2	Sirtuin 2	Lymphoma	2.48
STEAP4	STEAP family member 4	Lung Cancer	16.96
TAC1	Tachykinin, precursor 1	Breast Cancer	-3.2
TFPI2	Tissue factor pathway inhibitor 2	Kidney, Melanoma	-2.1
TMPRSS2	Transmembrane protease, serine 2	Kidney, Colorectal	30.12
TXN	Thioredoxin	Lymphoma, Melanoma	6.26
TXNIP	Thioredoxin interacting protein	Breast Cancer	-2.18
URGCP	Upregulator of cell proliferation	Myeloma	3.65

Table S5. Reference genes and primers for RT-qPCR.

Gene	GenBank Accession number	Function	Primer Forward	Primer Reverse
Sox10	AY149116	Regulation of embryonic development, determination of cell fate	TCAGCATTTCCCTCCATCTC	GGCCGAATGGCTGTAATAAG
Slug	AF368040	Neural crest formation	ACACTGCAACAGAGCATTCG	AGCAACCAGATTCCTCATGC
ODC	NM_001086698.1	Housekeeping gene	TCCATTGAGAGCGTAGGACTTG	GAGGCTCGCCGGTGAAATA

Data S1. Differentially expressed transcripts in early and late embryos and their association with disease or cellular process. An Excel file with the details of the pathway analysis is provided. The file contains three worksheets:
"Cell Process 15 stage" lists the pathways associated with the indicated transcripts identified in the early embryos according to Pathway Studio. Total # of neighbors in a given pathway; Overlap represents the number of genes in the 45 identified in the early embryos exposed to ivermectin that were associated with the pathway; Gene Set Seed refers to the entity in the pathway to which the genes are related; Overlapping Entities indicates the names of the identified differentially regulated genes associated with the pathway.

"Cell Process 45 stage" lists the cellular pathways associated with the human homologs of the 517 differentially regulated transcripts according to Pathway Studio. Columns are defined as in the Cell Process 15 stage description.

"Disease 45 stage" lists the diseases associated with the human homologs of the 517 differentially regulated transcripts according to Pathway Studio. Columns are defined as in the Cell Process 15 stage description, with the exception that the Name column refers to a disease, not a pathway.

Data S2. Results of the training set and validation set experiments and the performance of the model for each. An Excel file includes the results for each of the 20 experiments used for training the model and the 7 experiments used for validation. Experimental results are indicated in the "in vivo" columns and model results in the "Model" columns. All drug experiments were performed similarly with the drugs applied at stage 10 and animals scored at stage 45. For experiments with injected RNAs, the experiments were performed as described in the Materials and Methods. RNA injection experiments included the VP16 XlCreb1, ACREB, and IVM + ACREB data. IVM, ivermectin; Med 5HT (serotonin in the culture medium); MTP, methiothepin; 2'5'-DDA, 2'5'-Dideoxyadenosine; MRIF, MSH-RIF.