Supplementary Material

Contents

I. ADIABATIC APPROXIMATION, THE SLAVING PRINCIPLE AND SELF-ORGANIZATION

This section presents the motivation for our continuous time recurrent neural network approach based on the general concepts of self-organization in complex systems [Haken, 2004].

We denote genes and proteins as the elementary functional elements of a cell. The activity of a gene is controlled, among other mechanisms, by protein transcription factors binding to the promoter sites. A readout of this gene activity are the concentrations of the transcribed mRNA, $r(t)$, and the subsequently translated protein $p(t)$. The steady state concentrations of proteins are determined by their turnover rate, i.e. the ratio of production versus their decay rates. The latter is comprised of the protein decay itself and possibly by reactions into other protein variants, e.g. due to phosphorylation. The biological time scales for mRNA and protein concentration changes are in the order of $\approx 10^4 s$. Protein activation/deactivation, however, happen on a time scale of $\mathcal{O}(10^2 s)$. Those processes comprise for example post-translational events such as phosphorylation or active translocation.

Adiabatic Approximation In general, the reaction of a system to a time-dependent perturbation depends on the time scale of the perturbation itself, i.e. it makes a difference whether changes to a dynamic system are applied gradually or all of a sudden. A gradual change in the external conditions characterizes an adiabatic process.

Consider the functional activation of a protein p to the state p^*. The activation decays over time with constant γ_p^*. Based on a Law-of-mass-action this can be formulated mathematically using ordinary differential equations

$$\dot{p^*}(t) = f[p(t)] - \gamma_p^* p^*, \tag{S.1}$$

where p^* denotes the activated, e.g. phosphorylated, protein concentration and $f[p(t)]$ is the activation process of protein p. It is biologically reasonable to assume $\gamma_p^* > 0$ so that the system is stable and damped for $f = 0$.

Formally integrating Eq. (S.1) one obtains the long-term solution

$$p^*(t) = \int_0^t \exp[-\gamma_p^*(t-\tau)] f[p(\tau)] d\tau. \tag{S.2}$$

Making the function $f[p(t)]$ explicit, e.g. by considering the decay of the basal protein

concentration p_0 with a decay constant γ_p

$$f[p(t)] = p_0 \exp(-\gamma_p t), \tag{S.3}$$

one can readily integrate Eq. (S.2).

$$p^*(t) = \frac{p_0}{\gamma_p^* - \gamma_p}(\exp(-\gamma_p t) - \exp(-\gamma_p^* t)). \tag{S.4}$$

In general, changes in activated protein concentration follow its basal state concentration closely over time. This requires that the system acts instantaneously, i.e. $\gamma_p^* \gg \gamma_p$. Consequently, Eq. (S.4) can be approximated by

$$p^*(t) \approx \frac{p_0}{\gamma_p^*} \exp(-\gamma_p t) \equiv \frac{1}{\gamma_p^*} f[p(t)]. \tag{S.5}$$

In other words, for the above requirement of instantaneous reaction to be fulfilled, the time constant $t^* = 1/\gamma_p^*$ of the protein activation must be much shorter than the time constant $t' = 1/\gamma_p$ of the basal protein concentration change.

Note that the same result (Eq. S.5) can be found by setting $\dot{p^*} = 0$ in Eq. (S.1)

$$0 = f[p(t)] - \gamma_p^* p^*. \tag{S.6}$$

Hence, one can approximate the long-term dynamics of the fast system (activated protein) by solving $p^*(t)$ in terms of the slow forcing (protein decay), ignoring the information within the time scale of $\mathcal{O}(1/\gamma_p^*)$. This assumption is called 'adiabatic approximation' and is of fundamental importance for the self-organization of complex systems [Haken, 2004].

The above requirements are generally reflected in cell biology. The time-scales of protein activation involved in signaling are usually in the order of minutes, e.g. for phosphorylation and de-phosphorylation events, while protein translation and decay follow mRNA concentration changes on time-scales of hours. Notable exceptions exist, when the protein decay is actively regulated and actively interfering with the processes under consideration (see below).

Assuming the above proportionality between protein and mRNA concentrations, one can eliminate the protein dependence p in Eqs. (S.1)–(S.8) and rewrite them as a direct function

of the mRNA concentration r by substituting

$$f[p(t)] = f[p(r(t))]. \tag{S.7}$$

In other words, long-term cellular processes lasting hours to days can be described as a function of gene expression.

The Slaving Principle As an illustrative example, we consider a simple gene-protein interaction system. A gene, whose activity is reflected by the mRNA concentration r, produces protein p, which can be activated to p^*, while p^* can inhibit r. Together with the respective decay rates, such a system could read

$$\dot{r} = -k_p^*(p^*)^2 - \gamma_r r \tag{S.8}$$

$$\dot{p} = k_r r - \gamma_p p - k_p p \tag{S.9}$$

$$\dot{p^*} = k_p p - \gamma_p^* p^*, \tag{S.10}$$

where we used linear and quadratic interaction functions for the sake of simplicity. We assume $\gamma_p^* > 0$, so that protein activation is stable and bound. Furthermore, we assume $\gamma_p^* \gg \gamma_p$ to allow for the adiabatic approximation of $p^*(t)$ and we set $\gamma_p \approx \gamma_r$, so that the protein and mRNA concentrations evolve on the same time scale. Setting $\dot{p^*} = 0$ in Eq. (S.10) and expressing the steady state of p in terms the mRNA $r(t)$ in Eq. (S.9) one finds for the temporal behavior of the protein concentrations

$$p(t) \approx \frac{k_r}{\gamma_p + k_p} r(t) \tag{S.11}$$

$$p^*(t) \approx \frac{k_p}{\gamma_p^*} \frac{k_r}{\gamma_p + k_p} r(t). \tag{S.12}$$

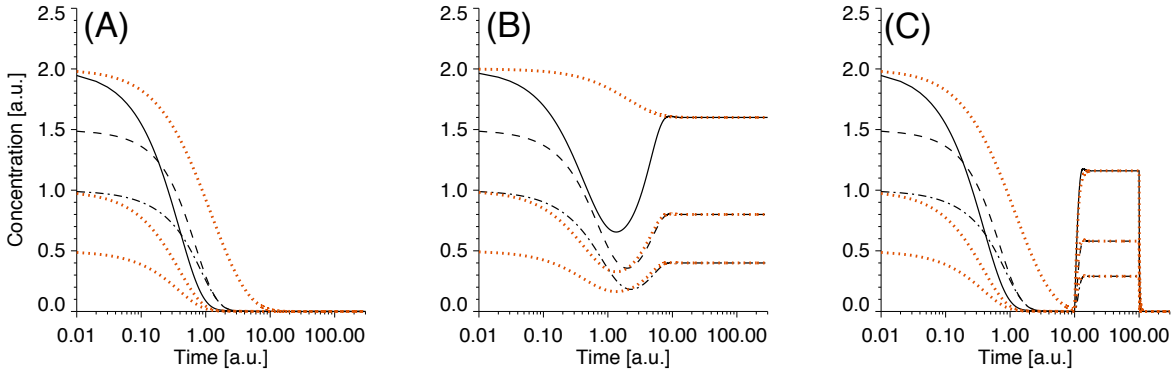
p^* thus follows the mRNA activity over time, hence it is said to be slaved by the slow subsystem. Via the proportionality between r and p and the adiabatic approximation of p^* the protein kinetics react to the gene kinetics of $r(t)$, which determines whether an action occurs or not. Substituting Eqs. (S.11), (S.12) into Eq. (S.8) one finds

$$\dot{r} = -\gamma_r r - k_p^* \left(\frac{k_p}{\gamma_p^*} \frac{k_r}{\gamma_p + k_p} \right)^2 r^2 = -r(k' r + \gamma_r) \tag{S.13}$$

with $k' = k_p^* \left(\frac{k_p}{\gamma_p^*} \frac{k_r}{\gamma_p + k_p}\right)^2$.

Depending on the sign of γ_r in Eq. (S.13), two different solutions can occur. If $\gamma_r > 0$, the only possible solution for $t \to \infty$ is $r = 0$. Hence, gene expression decays over time and likewise the protein concentrations decreases to zero (Fig. 1A). If $\gamma_r < 0$, the steady state solution is $r = k'/\gamma_r$ and consequently $p, p^* \neq 0$ (Fig. 1B). Note, how the approximate long-term solutions (red dotted lines) for the mRNA and the proteins (Eqs. S.11–S.13) approach the true solution after a transient time $t > 10$. This is also true, if the gene encounters a transient up-regulation, e.g. via a transcription factor (Fig. 1C). Both proteins still follow the gene closely and the adiabatic approximation (red lines) retraces these dynamics well.

Since it is the mRNA r, which determines whether a non-zero steady state for the proteins is reached or not, r is referred to as order parameter. Order parameters are those parameters that slave subsystems. Hence, for long-term cellular processes lasting on the order of hours or longer, gene expression kinetics should play a decisive role with respect to the phenotypic outcome of a cellular decision.



Supplementary Figure 1: Simulation of protein-gene interaction system from Eqs. (S.8)–(S.10). The slow gene dynamics slave the fast protein dynamics. (A) $\gamma_r = 0.5$. The gene expression (—) decays and protein concentrations $p(t)$(- - -) and $p^*(t)$ $(- \cdot -)$ follow likewise to zero. (B) $\gamma_g = -0.5$: finite, non-zero gene expression. Protein concentration follow to a non-zero steady state concentration. The red dotted lines show the approximate solutions for the gene (Eq. S.13) and protein kinetics (Eqs. S.11,S.12. (C) $\gamma_g = 0.5$: addition of an input term $I = 1$ for $10 < t < 100$ in Eq. (S.8). Other parameters: $[\gamma_p, \gamma_p^*, k_p, k_p^*, k_r)] = [0.5, 5, 2.50, 5, 1.5]$. Start values: $[r(0), p(0), p^*(0)] = [2, 1.5, 1]$.

Self-Organization In general, a protein signaling network, consisting of the basal proteins and their activated forms can be described by the following system of coupled ordinary differential equations

Proteins:

$$\begin{aligned}
\dot{p}_1 &= -\gamma_p^1 p_1 + f_p^1(r_1; p_1, \ldots, p_n; p_1^*, \ldots, p_n^*), \\
\dot{p}_2 &= -\gamma_p^2 p_2 + f_p^2(r_2; p_1, \ldots, p_n; p_1^*, \ldots, p_n^*), \\
&\vdots \\
\dot{p}_n &= -\gamma_p^n p_n + f_p^n(r_n; p_1, \ldots, p_n; p_1^*, \ldots, p_n^*).
\end{aligned} \tag{S.14}$$

activated Proteins:

$$\begin{aligned}
\dot{p}_1^* &= -\gamma_p^{*1} p_1^* + f_p^{*1}(p_1, \ldots, p_n; p_1^*, \ldots, p_n^*), \\
\dot{p}_2^* &= -\gamma_p^{*2} p_2^* + f_p^{*2}(p_1, \ldots, p_n; p_1^*, \ldots, p_n^*), \\
&\vdots \\
\dot{p}_n^* &= -\gamma_p^{*n} p_n^* + f_p^{*n}(p_1, \ldots, p_n; p_1^*, \ldots, p_n^*),
\end{aligned} \tag{S.15}$$

where the variables p_i and p_i^* , $1 \leq i \leq n$ denote the basal and activated proteins, respectively. All variables are coupled via the non-linear functions f_p^i, which depend on all other proteins p, p^* and additionally on the mRNA concentrations r_i for the basal proteins p_i. We require the protein concentrations to remain stable, i.e. $\gamma_p^{*i}, \gamma_p^i > 0 | \forall i$ so that the system has a stable state, which would correspond to the cell homeostasis. Next, we apply the adiabatic approximation by setting $\dot{p}_i^* = 0$, justified by the fact that protein activation is a much faster process than the overall protein decay of p. By virtue of $\gamma_p^* \gg \gamma_p$ it is assumed that the $|p^*|$'s are much smaller than the $|p|$, i.e. that events such as (de-)phosphorylation constitute a fast modulation of the basal protein concentration. However, this requirement needs to be checked for any real biological system. Hence, setting $p_i^* = 0$ for all f_p^{*i} allows

to solve for the p_i^*'s in terms of the basal protein concentrations p_i in Eq. (S.15)

$$\begin{aligned}
\gamma_p^{*1} p_1^* &= f_p^{*1}(p_1, \ldots, p_n), \\
\gamma_p^{*2} p_2^* &= f_p^{*2}(p_1, \ldots, p_n), \\
&\vdots \\
\gamma_p^{*n} p_1^* &= f_p^{*n}(p_1, \ldots, p_n).
\end{aligned} \tag{S.16}$$

Putting Eq. (S.16) into Eq. (S.14) leads to equations that depend on the basal protein concentrations p_i and the mRNA r_i alone

$$\dot{p}_i = -\gamma_p^i p_i + f_p^i[r_i; p_1, \ldots, p_n; p_1^*(p_1, \ldots, p_n), \ldots, p_n^*(p_1, \ldots, p_n)]. \tag{S.17}$$

Provided that protein concentration follows mRNA regulation closely over time, i.e. $p_i(t) \propto r_i(t)$ (no considering the time delay between mRNA transcription and translation), we can substitute all protein variables in Eq. (S.17) by the respective mRNA variables.

$$\dot{r}_i = -\gamma_r^i r_i + \frac{1}{c_i} f_r^i[p_1(r_i), \ldots, p_n(r_i); p_1^*(r_1, \ldots, r_n), \ldots, p_n^*(r_1, \ldots, r_n)] + I_i(t), \tag{S.18}$$

where we used $\gamma_p \approx \gamma_r$. The c_i denote the respective constants of proportionality between the mRNA and protein concentrations based on the ratio of effective mRNA translation/transcription and protein decay. The dependence of f_p^i on r_i in Eqs. (S.17) has been translated into a transcription activity of the respective gene, which is accounted for by the input function term $I_i(t)$

This leads to an equation describing the long-term behavior of the cellular protein signaling network, which is expressed solely in terms of the mRNA concentrations. The solutions of Eq. (S.18) determine, depending on $\gamma_r^i <> 0$, whether the slaved protein dynamics show a non-zero action or not. Hence, the mRNAs take up the role of ordering parameters in the system.

In many practical examples only very few variables (or modes) become unstable. If all other variables remain damped and adiabatically follow the unstable variables, then the behavior of the whole system is determined by the few order parameters. This has two consequences: first, complex systems may show a very regulated transition behavior between

different states and furthermore, depending on the particular variables, which become unstable and cause a transition, different macroscopic output states can be achieved.

Returning to biology, this means that a cell can undergo reversible transitions to different phenotypes, i.e. migration, proliferation, differentiation [Huang and Ingber, 2007], and these transitions will be most likely regulated by a few genes, only.

The limitations of this analysis mainly depend on two points. Firstly, the time window in which a cellular decision is produced has to be sufficiently broad (over a time span of hours). For example, the heat-shock stress response escapes our analysis, since the phenotypic reaction is present after a few minutes already. Secondly, cellular reactions that break the mRNA-protein correlation cannot be analyzed on the gene expression level. This holds true for the ubiqitin-proteasome signalosome, which is able to massively rebalance the protein concentrations, uncoupled from transcriptional regulation, or for post-translational modifications of histones, which can be epigenetically inherited.

Connection with the CTRNN model In our paper we use the following recurrent neural network to model gene regulation

$$\dot{r}_i = \frac{1}{\tau_i}\left[-r_i + \sum_{j=1}^{N} W_{ij}\sigma(r_j(t-\Delta t_j)+\theta_j) + I_i^0 \exp(-\lambda t)\right] \tag{S.19}$$

Comparing Eq. (S.19) with Eq. (S.18), we find a similar structure by identifying τ_i^{-1} with γ_r^i and $\sum W_{ij}\sigma(r_j)$ with $f_p^i(\dots)$. The external input function $I_i^0 \exp(-\lambda t)$ considered for each gene corresponds to the function for the transcriptional gene activity $I_i(t)$ in Eq. (S.18). Hence, by unraveling the mRNA interaction network and using the proportionality between the mRNA and the protein concentration, we arrive at a model structure that can be realized by a CTRNN and which can serve as one explanation why it has been possible to both measure and model on the gene regulatory level, yet verify model predictions on the protein signaling level.

[Haken, 2004] Haken H (2004). Synergetics: Introduction and Advanced Topics. Springer, Berlin.

[Huang and Ingber, 2007] Huang S, Ingber D (2007). A Non-Genetic Basis for Cancer Progression and Metastasis: Self-Organizing Attractors in Cell Regulatory Networks. Breast Disease **26:** 27-54.

II. MASIGPRO ANALYSIS

The list of probes across the DNA microarrays from all experiments has been filtered using interquartile filtering [vonHeydebreck *et al*, 2004]. The aim of this nonspecific filtering is to remove the genes that are unlikely to carry information about the phenotypes under investigation. This filtering removes genes having small changes within the experimental time points. The resulting list of around 6000 probes is used for the subsequent maSigPro analysis.

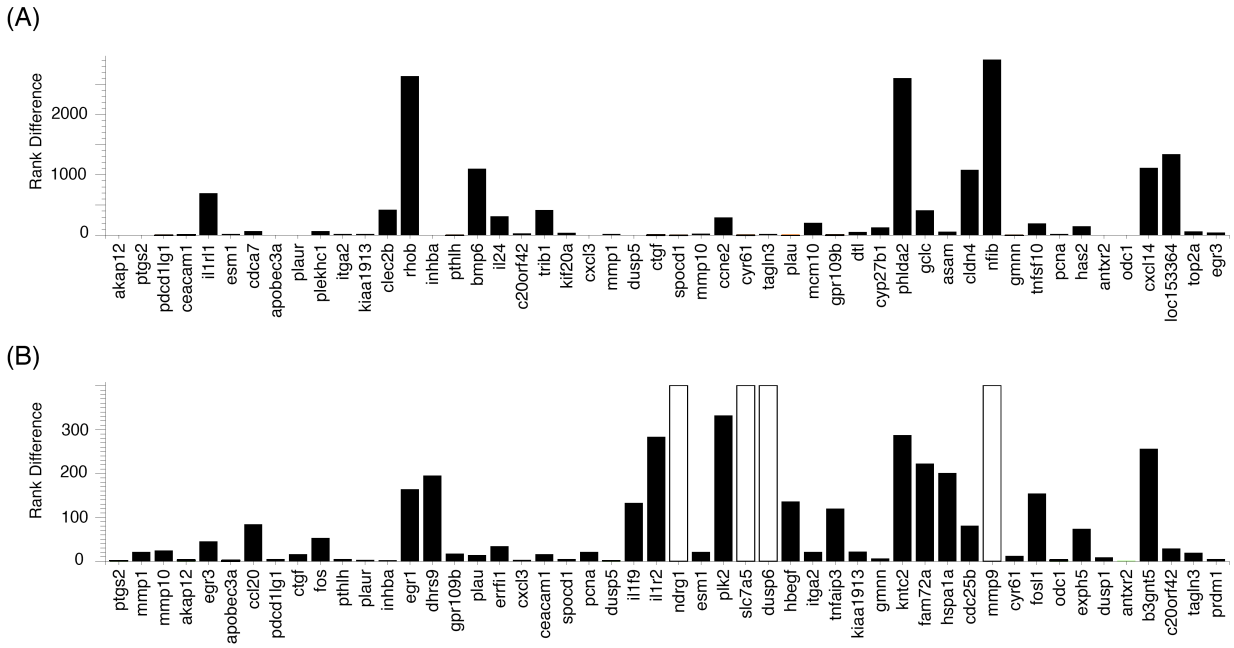
maSigPro [Conesa *et al*, 2006] is an R package for the analysis of single and multi-series time course microarray experiments. It follows a two-step regression strategy to find genes with significant temporal expression changes and significant differences between experimental groups. The method defines a general regression model for the data, where the experimental groups are identified by dummy variables. The procedure first adjusts this global model by a least squares technique to identify differentially expressed genes and then selects significant genes applying false discovery rate control procedures. Next, a stepwise regression is applied as a variable selection strategy to study differences between experimental groups and to find statistically significant temporal gene expression profiles. The coefficients obtained in this second regression model can be used to cluster together significant genes with similar expression patterns and to visualize the results. Here, a quadratic regression fit was performed using significance level of 0.05, and the Benjamini-Hochberg correction method for multiple testing problems. A stepwise regression fit for time series gene expression experiments was performed using a T-fit with a two-ways-backward method. Finally, a list of significantly regulated genes was obtained for a set of variables, whose significance values have been computed with the T-fit function.

In the case of HGF-induced migration versus the 0h control experiment, maSigPro identified 536 significantly regulated genes from 607 probesets having a p-value $p < 0.05$. Comparing the 50 highest ranked genes according to the mean and peak fold expression Eq. (1) and the maSigPro methods, one finds a good agreement with respect to identified genes and their respective rank, in particular for those genes, which have been found to be decisive for

the cell migration decision, such as *akap12*, *ptgs2*, *plau* or *egr3* (Supplementary Fig. 2).

[vonHeydebreck *et al*, 2004] von Heydebreck A, Huber W, Gentleman R (2004) Differential expression with the bioconductor project. In: Bioconductor projects working papers.

[Conesa *et al*, 2006] Conesa A, Nueda MJ, Ferrer A, Talon M (2006) maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments. Bioinformatics **22:** 1096-1102.



Supplementary Figure 2: Comparison of gene ranking methods: Gene rank difference of the 50 most significantly regulated genes according to (A) maSigPro and (B) gene ranking according to mean and peak fold expression. The hollow bars in (B) denote genes that are not identified by maSigPro as being significant.

III. GENETIC ALGORITHMS

The task of model fitting is to determine the unknown parameters of the gene regulatory network such that it behaves as reflected in the experimental gene expression kinetics. Due to the internal structure of the recurrent neural network model many local optima exist, wherein deterministic gradient-based search algorithms are likely to get stuck. Hence, model training of neural methods usually use Back Propagation through Time [Werbos, 1990] or evolutionary algorithms such as Particle Swarm Optimization [Kennedy *et al*, 2001] or Genetic algorithms [Wahde and Hertz, 2001].

Here, we have chosen to use a genetic algorithm (GA) to determine the unknown parameters of the network. The GA works with a fixed-size population of possible solutions. Each solution, called an individual, evolves step-wise from generation to generation using three principal genetic operators: selection, crossover, and mutation. All individuals are evaluated in each generation according to a fitness function (cf. Eq. 7 in Materials and Methods) and then selected to produce two offspring, comprising the next generation. Having selected two parents, their chromosomes, i.e. their set of parameters, is recombined in the offspring, which can furthermore undergo mutation. The parent individuals are selected according to their respective fitness with no restriction on the number of offspring an individual can produce. The purpose of the crossover and mutation is to combine useful parental information to form new and hopefully better performing offspring. After many generations, only those individuals having the highest fitness survive with the best individual providing a (near) optimal solution to the problem, i.e. best fitting the experimental data.

In our study there are several advantages of the GA in comparison to other optimization approaches. Given the scarcity of data points in relation to the number of model parameters, the reverse engineering problem is highly over-parameterized resulting in a very large set of candidate solutions with many local optima. Genetic algorithms work probabilistically and consider several solutions at the same time, which allows them to explore large areas of the parameter space rather efficiently [Willett, 1995]. Moreover, GAs work with the least assumption on parameter values: only parameter ranges and the fitness function need to be specified. This gives this optimization technique a high flexibility, if changing the number

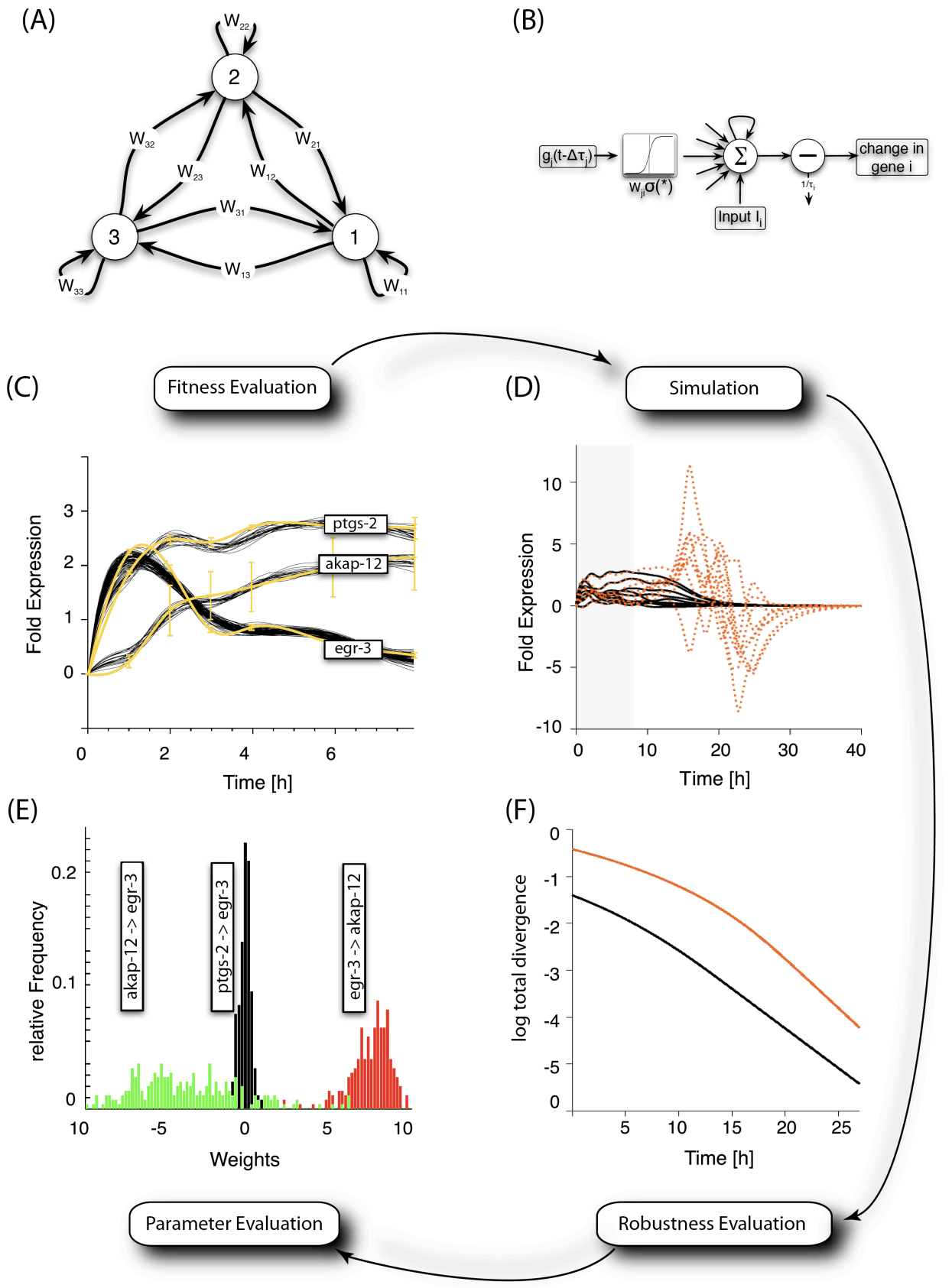
of network nodes or the network structure itself.

[Kennedy *et al*, 2001] Kennedy J, Eberhart R, Shi Y (2001). Swarm Intelligence, Morgan Kaufmann Publishers.

[Wahde and Hertz, 2001] Wahde M, Hertz J (2001) Modeling Genetic Regulatory Dynamics in Neural Development. J. Comp. Biol. **8:** 429-442.

[Werbos, 1990] Werbos PJ (1990) Backpropagation Through Time: What It Does And How to Do It. Proceedings of IEEE **78:** 1550-1560.

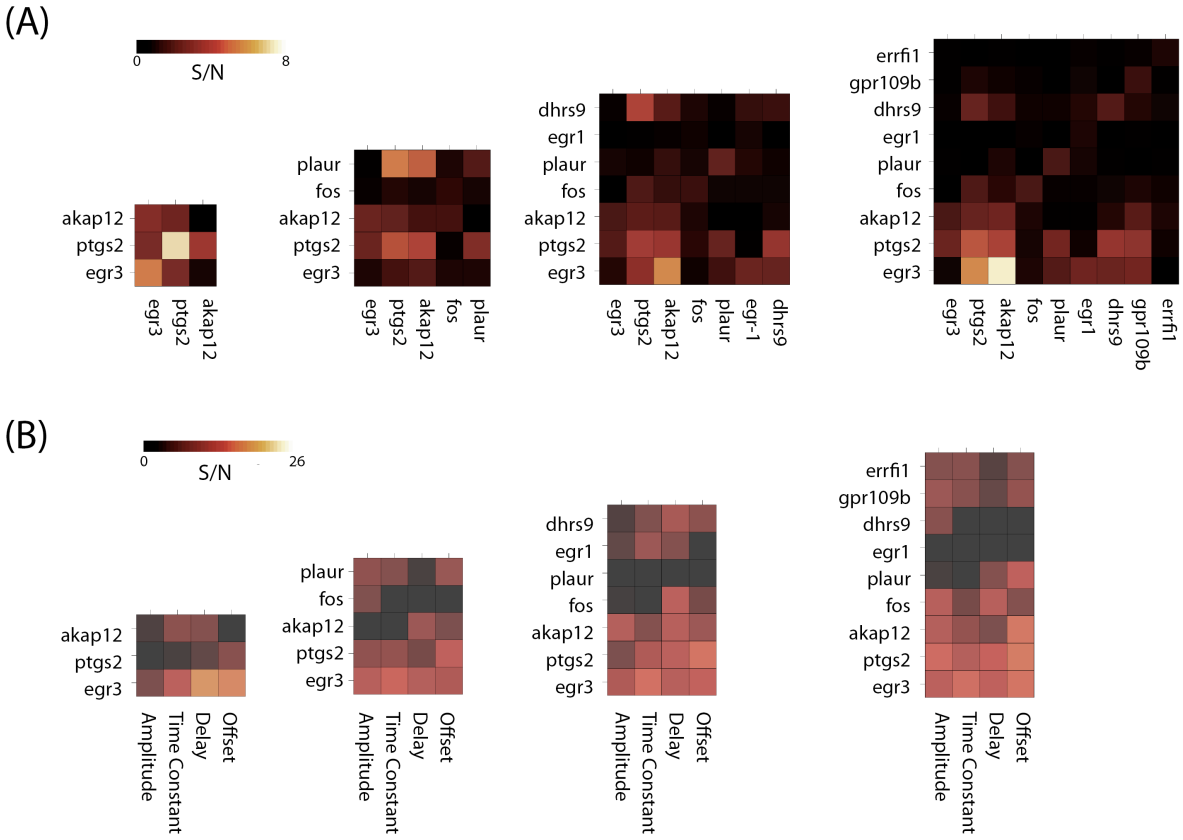
[Willett, 1995] Willett P (1995) Genetic algorithms in molecular recognition and design. Trends Biotechnol. **13:** 516-521.



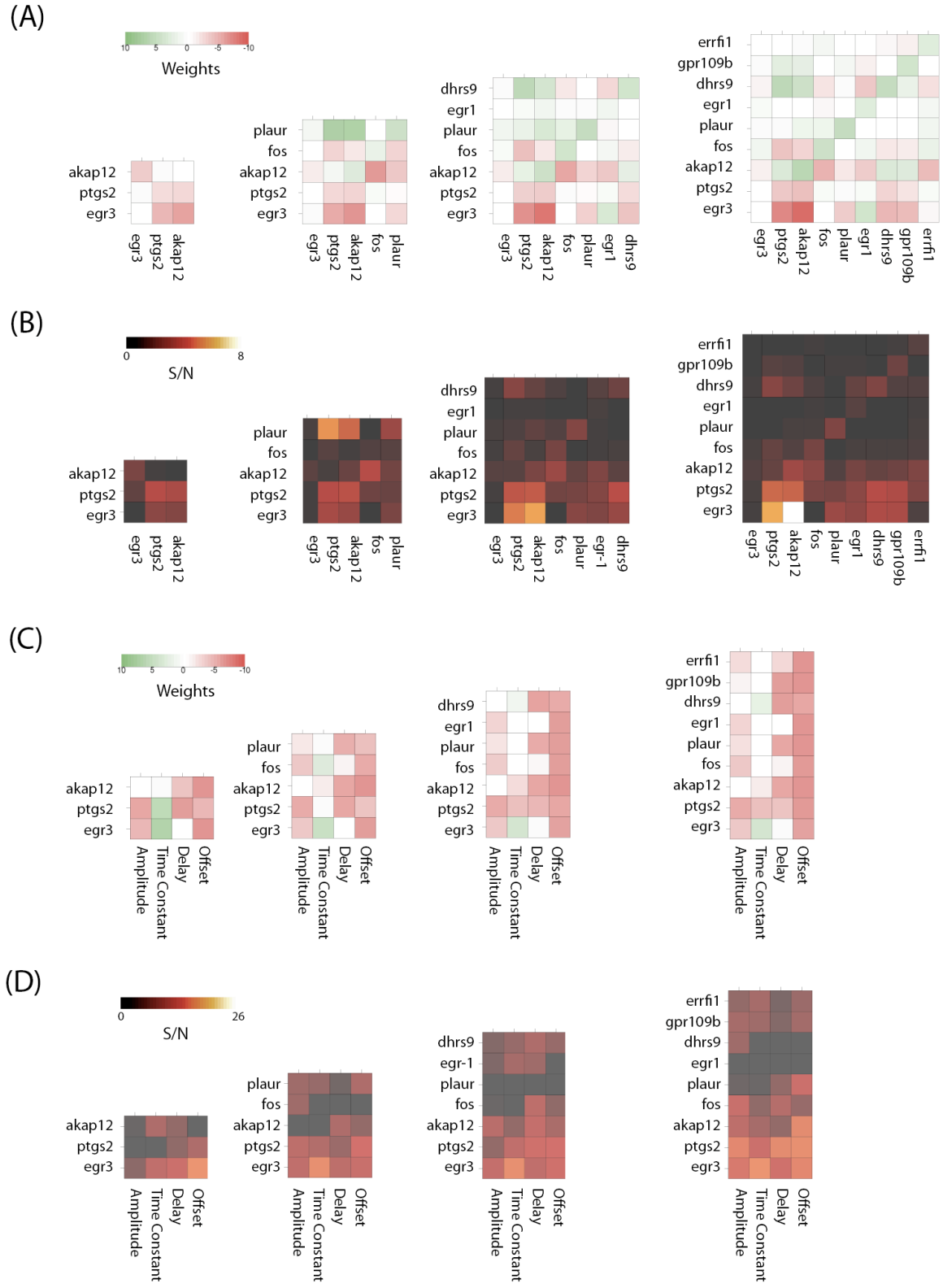
Supplementary Figure 3: (A) Schematic view of a fully connected three-gene network. Weights indicate strength of directed interactions. (B) Information processing flow of a node (gene) in the CTRNN. The input to a gene i consists of an external stimulus I_i as well as the summed interaction with all genes, including self- interaction. The latter is composed of the delayed gene fold expression of each gene at time $t - \Delta\tau_j$, which is modified by the activation function σ and multiplied by the interaction weight W_{ij} before finally being summed into gene i.

Supplementary Figure 3 (continued)

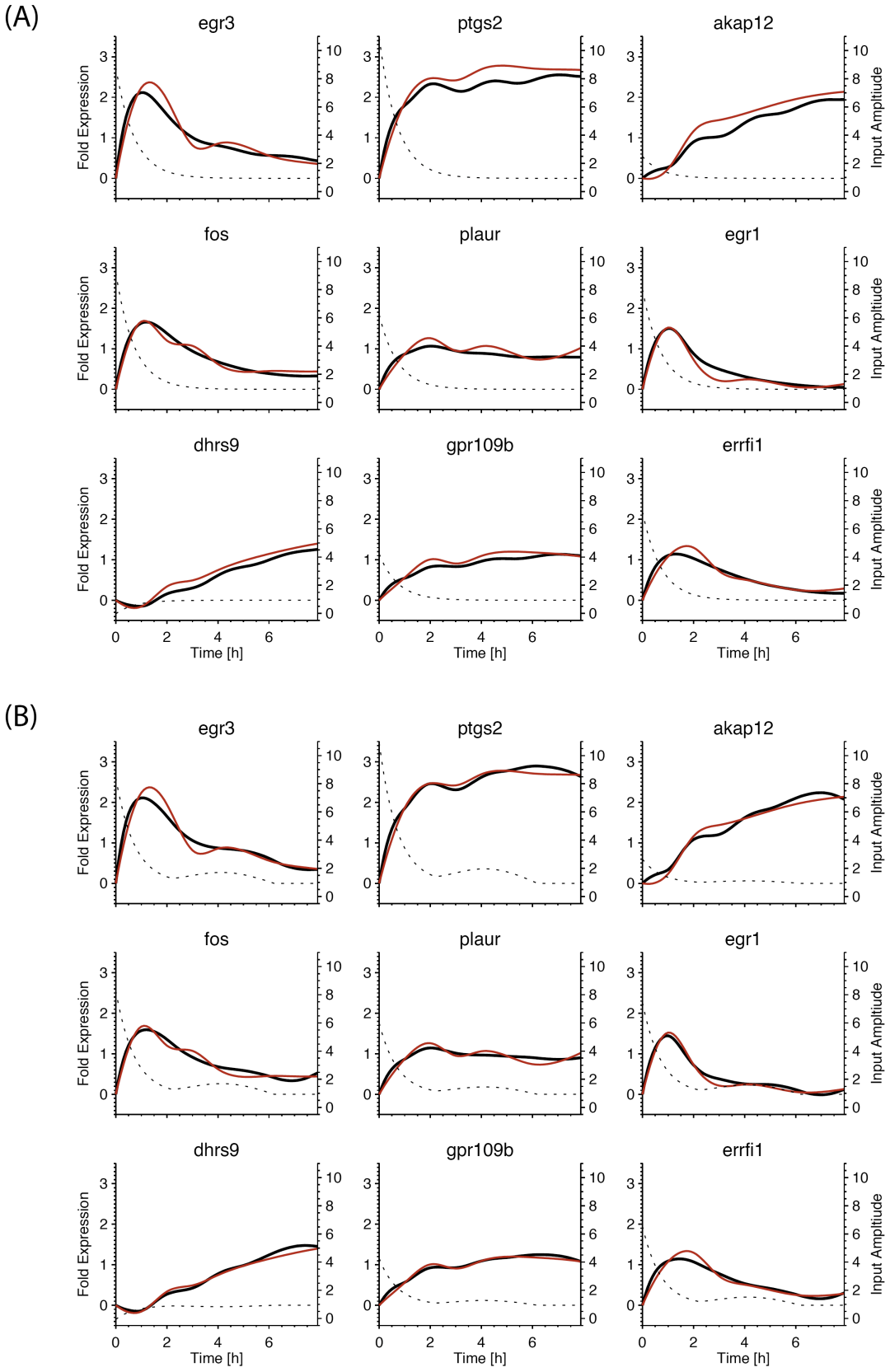
$1/\tau_i, \tau_i > 0$ determines the time constant and correspondingly the stability of the fixed point of each gene. The function $g_i(t)$ for each gene is learned such that the time course maximizes the fitness function (Eq. 7) with respect to the experimental spline-interpolated fold expression data. (C-F) Model evaluation consists of three different steps (cf. workflow in Fig. 1A). First, model fitness is assessed by the fit to the experimental data. Then, the model robustness is evaluated by its divergence (D, F). (C) Sample trajectories from a nine-gene network of different, independent model fits to the experimental data for genes $egr3$, $ptgs2$ and $akap12$. Each of the black lines corresponds to one independent model fit. Spline interpolated experimental data is indicated by the yellow lines. The error of the experimental data has been calculated as described in the Materials and Methods. (D) Example of system robustness: simulated gene expression time series of all genes in a nine-gene network using 2 different parameter sets (red and black). Despite similar fitness to the experimental data (first 8 hours, grey area), they show different dynamic behavior at later time points. (E) Histograms of sample interaction weights describing functional connectivity between genes. Histograms are drawn from 250 out of 5000 best ranked model fits. The histogram shows a consistent directed interaction between $ptgs2$ and $egr3$ (black) and between $egr3$ and $akap12$ (red), whereas the influence of $akap12$ on $egr3$ (green) is less consistent in different models due to its obviously low signal-to-noise ratio, as calculated from the mean to standard deviation ratio of the histogram. (F) Logarithm of the total divergence of the $ptgs2$ trajectories from (C) with corresponding colors. The total divergence is a measure of separation of nearby trajectories as a function of time. The more negative the logarithm of the divergence, i.e. the smaller the divergence, the faster the network settles back into a stable steady state upon stimulation.



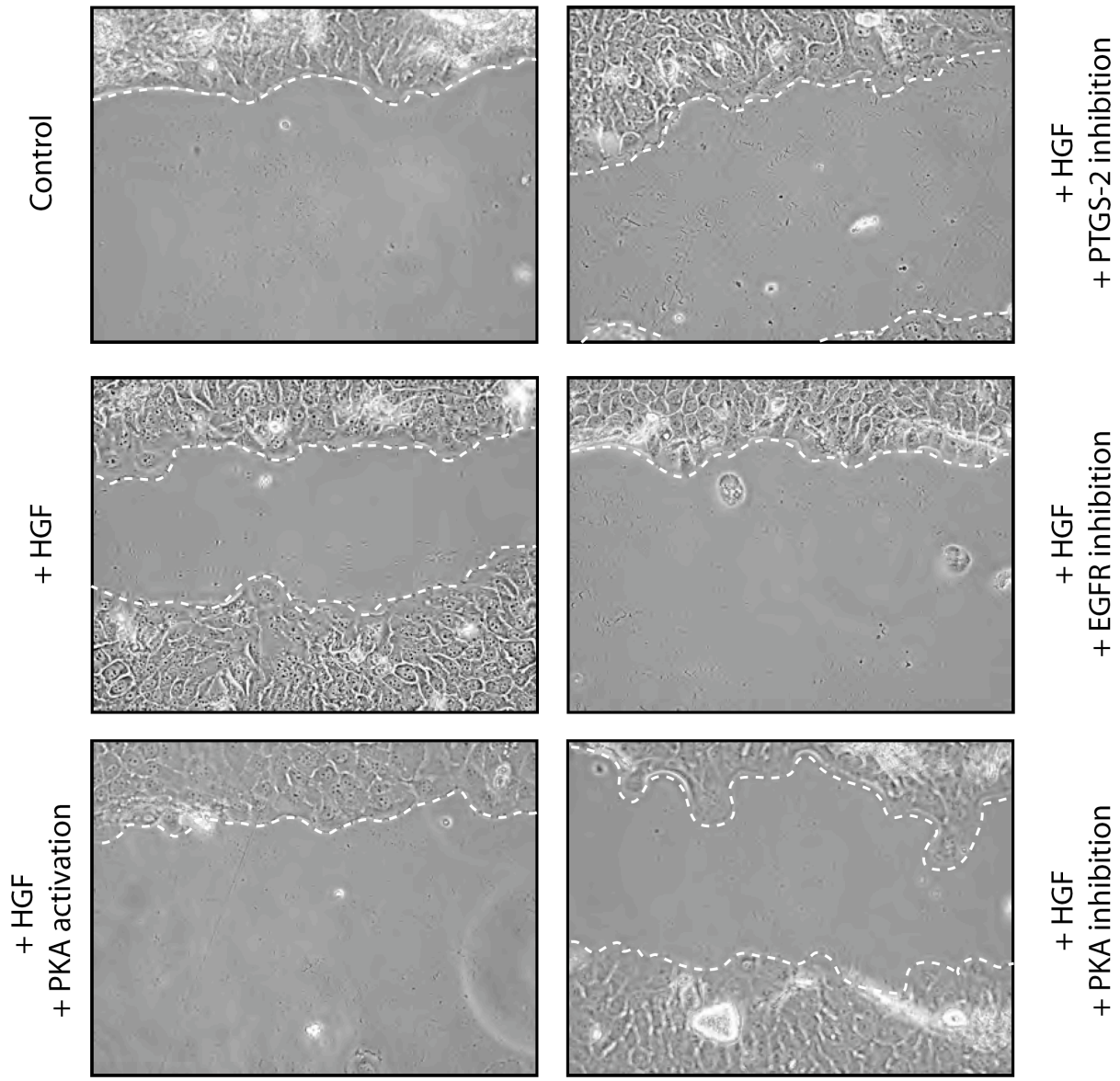
Supplementary Figure 4: (A) Signal-To-Noise (S/N) ratio for the inferred interaction weights for networks of size 3, 5, 7, 9 genes (from left to right). The S/N ratio is calculated from the ratio of mean to standard deviation of the weighted sum of the highest ranked parameter fits (cf. Methods Section). (B) Signal-To-Noise ratio of the amplitude, time constant and offset parameters of the inferred gene regulatory network for sizes 3, 5, 7, 9.



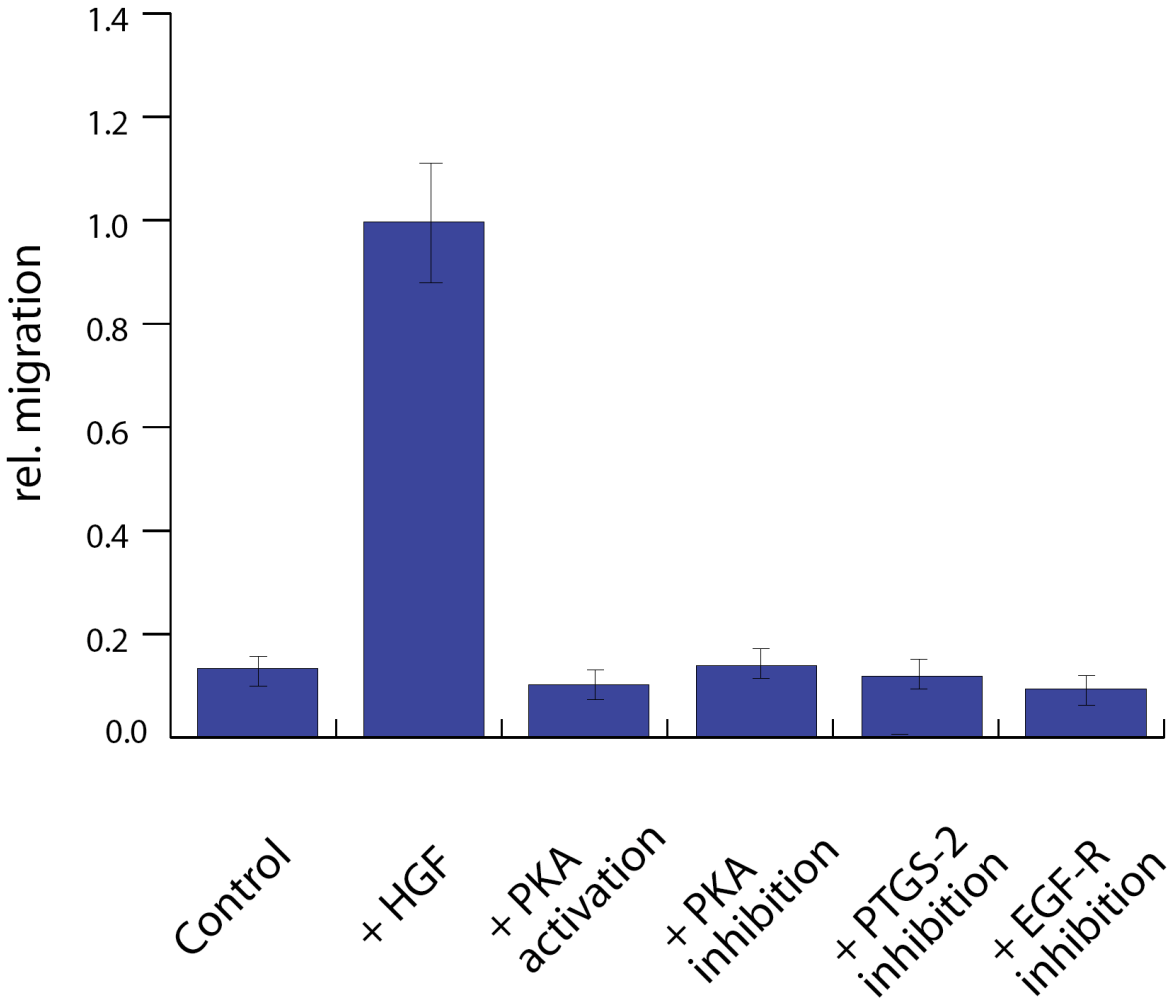
Supplementary Figure 5: (A) Interaction weights W_{ij} for the CTRNNs obtained by inverse modeling for networks of size 3,5,7,9 genes (from left to right) with the genes first receiving an exponentially decaying and then a second, time-delayed, pulsed input (cf. Methods section). (B) Signal-To-Noise values for the interaction weights W_{ij} depicted in (A). (C) Other CTRNN parameters for the 2nd stimulus model with network sizes 3, 5, 7 and 9 genes (from left to right). (D) Signal-To-Noise values for the other parameters depicted in (B).



Supplementary Figure 6: Model simulations of nine gene networks. (A) Network simulation using parameter values obtained from fitting the experimental data with a single, exponentially decaying input (cf. Eq. 4 and Supplementary Table III A). (B) Simulation using parameter values obtained from fitting the experimental data with two time-consecutive inputs (cf. Eq. 5 and Supplementary Table III B). Red line: spline-interpolated gene expression kinetics from microarray experiment, black line: model simulation of gene kinetics, black dotted line: learned input function to each gene.



Supplementary Figure 7: Modulation of HGF induced migration by PTGS-2, PKA- and EGF-R activity, respectively. To induce a migratory response, HaCaT cells were incubated (30 h) with or without $10ng/ml$ recombinant human HGF. PTGS-2 activity was modulated using $50\mu M$ Meloxicam (4-Hydroxy-2-methyl-N-(5-methyl-2-thiazolyl)-2H-1,2-benzothiazine-2-carbox-amide-1,1-dioxide1, PKA activity was modulated by addition of $1\mu M$ H-89 (Calbiochem) and $200\mu M$ 8-(4-Chlorophenylthio)adenosine 3',5'-cyclic monophosphate sodium salt (Sigma) to inhibit and enhance PKA-activity, respectively. Inhibition of PKA by addition of $10\mu M$ Myristoylated PKI (14-22) amide, cell-permeable PKA inhibitor (Biomol) gave rise to similar results (data not shown). EGF-R activity was blocked by addition of $1\mu M$ GW 2974 (Sigma). Phase contrast microscopy photographs were taken 20h after initial HGF single pulse stimulation with 40x fold magnification. Migration front is indicated by a dotted line.



Supplementary Figure 8: Modulation of keratinocyte migration by PKA-pathway, EGF-R signaling and PTGS-2 activity depend on the initiation of migration by HGF. To induce a migratory response, HaCaT cells were incubated (30h) with or without $10ng/ml$ recombinant human HGF. PKA activity was modulated by addition of $1\mu M$ H-89 (Calbiochem) or 200 μM 8-(4-Chlorophenylthio)adenosine 3',5'-cyclic monophosphate sodium salt (Sigma). Inhibition of PKA by addition of $10\mu M$ Myristoylated PKI (14-22) amide, cell-permeable PKA inhibitor (Biomol) gave rise to similar results (data not shown). EGF-R activity was blocked by addition of $1\mu M$ GW 2974 (Sigma). Error bars denote $\pm$ standard deviation of mean migration distance from three independent experiments.