Different cell fates from cell-cell interactions : core architectures of two-cell bistable networks

Supporting Material

Hervé Rouault[†‡] and Vincent Hakim[†*]

†Laboratoire de Physique Statistique, CNRS, Université P et M Curie, Université Paris Diderot, Ecole Normale Supérieure, 75231 Paris Cedex 05, France.
‡Institut Pasteur, CNRS URA2578,75015 Paris, France
* Corresponding author

Contents

1 Description of the evolutionary algorithm **2**
- 1.1 The evolutionary procedure . . . 2
- 1.2 Description of transcriptional regulation . . . 3
- 1.3 Post-transcriptional interactions . . . 4
- 1.4 Cell-cell interactions . . . 5
- 1.5 Evaluation procedure and score function . . . 5
- 1.6 Initial differences between the two cells during the test phase. . . . 6
- 1.7 Network pruning . . . 7
- 1.8 Implementation of the algorithm . . . 7

2 Mathematical analysis of different networks **9**
- 2.1 The auto-activation complexation network: a bistable switch . . . 9
- 2.2 The mixed-feedback loop in the bistable regime . . . 11
- 2.3 A bistable two-gene network based on purely transcriptional non-cooperative interactions . . . 12
- 2.4 Simple bistable two-cell networks . . . 14
  - 2.4.1 Interaction between homologous proteins in the signal-sending and signal-receiving cells . . . 14
  - 2.4.2 Interaction between heterologous proteins in the signal-sending and signal-receiving cells: a network purely based on post-transcriptional interactions . . . 16

3 Description of the different networks **20**

4 Supplementary figures **27**

Chapter 1

Description of the evolutionary algorithm

1.1 The evolutionary procedure

The aim of the algorithm is to obtain genetic networks that implement a given task encoded in a score function. A starting collection of genetic networks is modified iteratively using repeated rounds of "growth" and "selection".

The starting collection of genetic networks is chosen to consist of 30 networks containing a fixed number of genes. In the initial networks, there are no interactions between the different genes : each of them produces initially its mRNAs and proteins at different rates but without any further interactions.

In the growth phase, the number of genetic networks is doubled by adding to the original collection a mutated version of each genetic network. Mutations can simply consist in changes in kinetic constants, such as mRNA or protein production or degradation rates. or, in a chemical constant describing gene regulation by a protein or an interaction between two proteins. They can also be real evolutionary novelties that can change the network topology or the number of its components, like the creation of a new transcriptional interaction between a protein and a gene promoter or the introduction of a new post-transcriptional interaction.

Selection follows the growth phase. All the genetic networks are evaluated and ranked according to a predetermined score function. The networks with scores in the top 50% are kept and the other ones are deleted. The network collection thus regains its original size and is ready for another round of growth and selection. The procedure is iterated for until a maximal allowed number of 1200 generations. To obtain bistable networks without being over-specific, we chose the score function to simply require that the protein concentrations remained sufficiently different, as time evolved, between cells starting from initial conditions with different imposed values of a given protein A, as detailed below.

Mutation rates are chosen in the present algorithm according to the mutation types. Simple or activated phosphorylations are chosen for instance comparably more frequently than dimerizations (see Table 1.1 for the precise probability of each type of modification). Moreover, initial conditions and kinetic constants are kept under reasonable range by ignoring a mutation of a kinetic constant or an initial condition outside the range $[0.1, 20]$. In this case, another mutation is chosen randomly for

the network. And in the same vein, we restrict the number of transcriptional regulations, to 5 for evolutionary simulations with the whole set of possible interactions and to 8 when only transcriptional interactions were included, and the number of dynamical variables to 12. To assess, the influence of these bounds, we performed simulations runs where these maximal numbers were doubled as reported in the main text and in Fig. S4.3.

Fig. S4.1 shows for illustration two runs of simulations that lead to the creation of cell-autonomous networks. The score of the ancestors of the final networks are also shown. It can be noted that they the score of the ancestors of the final networks are not always the best in the population and can even deteriorate for a few generations, although the best scoring network is always kept. We observed generally, that the number of generation until convergence, when convergence occurred in within the maximum allowed number of generations, decreased more slowly than $1/N$ when N, the number of networks in the collection was increased.

Before providing further details of the algorithm impementation, we note that for computational efficiency, the algorithm departs from biological evolution in several ways. The duplication with mutations of the best scoring networks differs from reproductive fitness and, as a consequence, the best-scoring networks of one generation always appear in the next generation. An important difference also comes from the choice of the score function that in our procedure directly evaluates the network for its ability to produce different cell fates. In reality this most certainly comes about as a by-product of a different selection pressure. These differences should certainly lead to caution about drawing conclusions about biological evolution, beyond the interesting fact that networks that perform somewhat complex tasks can be produced rather easily by trial and error under the guidance of a score function.

1.2 Description of transcriptional regulation

The dynamics of mRNA production (transcription) and protein production (translation) are governed by ordinary differential equations.

The transcription of a gene a producing an mRNA m_a is described as follows

$$\frac{dm_a}{dt} = \rho_a(P_1, \cdots, P_n) - \delta_a m_a, \tag{1.1}$$

where δ_a is m_a degradation rate, ρ_a its production rate and it has been assumed that the transcription of gene a is regulated by proteins (or protein complexes) $P_1, \cdots, P_n$. Explicitly, ρ_a is described with a Hill-like function, as resulting from the competition between the different factors that regulate the transcription of gene a (fast equilibrium dynamics on the promoter is assumed) :

$$\rho_a(P_1, \cdots, P_n) = \frac{\rho_0 + \sum_{j=1,\cdots,n} \rho_j (P_j/P_j^{(0,a)})^h}{1 + \sum_{j=1,\cdots,n} (P_j/P_j^{(0,a)})^h}. \tag{1.2}$$

In Eq. (1.2) ρ_0 is the basal transcription rate of gene a, the concentration $P_j^{(0,a)}$ measures the concentration for half-activation or repression of protein P_j on the promoter of gene a, and ρ_j the rate that results when transcription of gene a is fully dominated by P_j (i.e. for $P_j \gg P_j^{(0,a)}$). Note that if $\rho_0 < \rho_j$, protein P_j activates the transcription of gene a and represses it for $\rho_0 > \rho_j$. The Hill coefficient for binding, h, is taken here equal to 1, for non-cooperative transcriptional interactions and to 2 for cooperative

ones, as specified in the text. The Hill function description is a simplification that can sometimes prevent interesting behavior. For instance, neglecting time in the promoter binding/unbinding dynamics does not allow the "mixed-feedback loop" in its simplest version to oscillate (1) . Since we are primarily interested in stationary states in the present paper, we assume this simplification is of no consequence.

Protein production is more simply taken to be proportional to the mRNA concentration. That is, the production of protein A from its mRNA m_a is given by

$$\frac{dA}{dt} = \beta_A m_A - \delta_A A + Int_A, \tag{1.3}$$

where δ_A is the degradation rate of A. The last term Int_A eventually describes the interaction of A with other proteins such as complex formation or catalytic transformation to a modified form, an example of which is provided in Eq. (1.11) below (for a protein B).

1.3 Post-transcriptional interactions

The possible post-transcriptional interactions in the algorithm can be of two possible types, namely complex formation between two proteins ("dimerization" in Table 1.1), and modification of a protein, which can itself either happen at a fixed rate ("phosphorylation" in Table 1.1) or can depend on the concentration of another described protein (denoted as "activated phosphorylation" in Table 1.1). These different possibilities are mathematically implemented as follows. In each case, we only describe the terms which are added when the interaction is created and denote by dots the other interactions which the proteins can participate in.

Dimerization between protein A and B to produce the complex C, is described by

$$\frac{dA}{dt} = \cdots - \kappa_{AB} AB \tag{1.4}$$

$$\frac{dB}{dt} = \cdots - \kappa_{AB} AB \tag{1.5}$$

$$\frac{dC}{dt} = \cdots + \kappa_{AB} AB - \delta_C C \tag{1.6}$$

where κ_{AB} is the complex formation rate and δ_C the complex degradation rate. These two constants are chosen at random when the interaction is created, , as described above.

"Phosphorylation" of protein A, which produces the modified protein A^*, is simply described by

$$\frac{dA}{dt} = \cdots - \kappa_A A \tag{1.7}$$

$$\frac{dA^*}{dt} = \cdots + \kappa_A A - \delta_{A*} A^* \tag{1.8}$$

where κ_A is the rate at which A is modified and δ_{A*} the rate at which the modified form is degraded.

Finally, "activated phosphorylation" of protein A by protein B corresponds to

$$\frac{dA}{dt} = \cdots - \kappa_{AB} AB \tag{1.9}$$

$$\frac{dA^*}{dt} = \cdots + \kappa_{AB} BA - \delta_{A*} A^* \tag{1.10}$$

The concentration of protein B is not modified in this case, in contrast to the "dimerization" case and the rate of transformation of A into A^* depends on the concentration of B in contrast to the simple "phosphorylation" of A.

We did not investigate regulation at the level of translation or involving mRNA nor regulations coming from other types of RNAs. This subject of much current interest lies beyond the focus of the present study but it is certainly worth further computational investigations.

1.4 Cell-cell interactions

A central difference between the present algorithm and previous ones is the inclusion of cell-cell interaction. In order to implement it in a simple way and to avoid being too specific, interaction between cell 1 and 2 is introduced by taking protein A in cell 1 to induce the transformation of protein B into protein B^* in cell 2. This is mathematically represented by the following terms in B and B^* equations in cell 2

$$\begin{aligned} dB_2/dt &= \cdots - \gamma_{12} A_1 B_2 \\ dB_2^*/dt &= \cdots + \gamma_{12} A_1 B_2 \end{aligned} \tag{1.11}$$

where A_1 denotes the concentration of A in cell 1, B_2 and B_2^* the concentrations of B and B^* in cell 2, and the dots are meant to represent the other interaction which increase or decrease the amount of B_2 or B_2^*. As described in the text, in a few simulations the effects of receptor saturation and cooperativity were investigated by replacing Eq. (1.11) by

$$\begin{aligned} dB_2/dt &= \cdots - \gamma_{12} A_1 \frac{B_2^\sigma}{B_0^\sigma + B_2^\sigma} \\ dB_2^*/dt &= \cdots + \gamma_{12} A_1 \frac{B_2^\sigma}{B_0^\sigma + B_2^\sigma} \end{aligned} \tag{1.12}$$

with B_0 the concentration of B at half-maximum activation and the exponent σ was chosen to be 1 (saturation only) or 2 (saturation and cooperativity).

Note that we assume that A in the signal-sending cell influences the state of B in the signal-receiving cell but that this interaction by itself has no effect in the signal-sending cell (but of course when A is present in the signal-receiving cell and B is present in the signal-sending cell, the reciprocal interaction exists in which the sending/receiving characters of the cells are inverted). In other words, we do not consider here bidirectional signaling although it exists in some biological cases (like for instance for Eph/Ephrin interaction).

As with other reactions, when an interaction was introduced in a simulation run A and B were chosen at random among the proteins of the networks. As specified in the result section, A and B were further constrained to be different in some evolutionary runs.

1.5 Evaluation procedure and score function

The evaluation procedure and score functions should obviously be chosen according to the desired outcome. In order to select for networks producing different fates in different cells, the action of each genetic circuit was here tested for its action in two cells.

This test phase consisted in integrating the differential equations corresponding to the considered genetic circuit and evaluation of a chosen score function. The differential equation integration required the specification of initial conditions. These were prescribed by first imposing the concentration of protein A in each cell to a given value (dependent of the cell) and letting the other concentrations relax to these imposed values, by integrating the other concentrations starting from identical initial values in different cells. The initial values of protein other than A were attached to each network and let free to evolve in the course of the evolutionary simulations as other constants (see Table 1.1 for their respective probabilities of mutations). The different initial imposed values of A concentrations can be thought of as roughly mimicking different initial concentrations arising for instance from interaction with a positional information gradient.

The score function is particularly important for the success of the evolutionary procedure. It should provide a gradually improvable score and should also avoid being over-specific (like by requiring precise values for some protein concentrations) which would impede the success of the simulated evolution. To obtain bistable networks without being over-specific, we chose to simply require that the protein concentrations remained sufficiently different, as time evolved, between cells starting from initial conditions with different imposed values of protein A. This was implemented by first keeping the dynamical variables within a defined range by assigning a drastically deleterious score to networks producing mRNAs or protein concentration outside the $[0.005, 500]$ range (in arbitrary units).

Then a proximity between the expression profiles of a protein P in two different cells i and j was defined by

$$\mathrm{Prox}_{i,j}(P) = \int_0^T \Pi(P_i(t), P_j(t))\, dt \tag{1.13}$$

with T the time over which the network dynamics was integrated ($T = 300$ in dimensionless units unless otherwise specified) and Π the function defined by

$$\Pi(x, y) = \begin{cases} 1 - |x - y|/2 & \text{if } |x - y| < 2, \\ 0 & \text{otherwise.} \end{cases}$$

When two different fates were desired, two cells 1 and 2 were considered to have different fates (as assessed by measuring protein A concentration) when their protein expression profiles were never close namely when $\mathrm{Prox}_{1,2} < 1.0$.

1.6 Initial differences between the two cells during the test phase.

In our two-cell simulations, the initial concentrations of A were drawn at random in each evolutionary run in the same way as other protein concentrations and unchanged during the whole evolutionary run. All cells in one evolutionary run had the same initial concentrations of A. For the simulation runs with a steeper gradient (Fig. S4.3), the same procedure was adopted but after the random draw of the two initial A concentrations, the larger concentration was multiplied by 4 and the smaller divided by 4.

1.7 Network pruning

The raw networks observed once the convergence criterion was reached, were very often complex and decorated with proteins or reactions unessential to the desired function. In order to try and automatically reduce the successful networks to their essential, we added an additional pruning phase. While keeping the evolutionary procedure under the constraint of the score function, we let it run for 500 additional generations. In this optimization mode, the possibility to add reactions or proteins was removed in favor of protein and reaction elimination, and parameter adaptation. This automatic pruning is illustrated in Fig. S3.1, which displays a bistable AAC network before and after automatic pruning.

1.8 Implementation of the algorithm

The developed programs have been written in C++ and are available at *https://github.com/hrouault/Genherite*. They have been executed on an octoprocessor Intel Xeon machine with 32 Go RAM.

The random number generation and ODEs integrations have been performed with algorithms provided by the *GNU Scientific Library*. The random number generator used is the default Mersenne Twister algorithm. The stepping function for ODEs integration makes use of the Runge-Kutta Prince-Dormand method and step size control is achieved with the standard method proposed by the library. Parallelization is brought into play with multithreading. The program outputs the generated networks in the `dot` file format for graph visualization.

Tables

Table 1.1: **Mutation rates**

Mutations	**Relative rate**
Initial condition modification	3.0
Reaction rate modification	2.0
Degradation rate modification	2.0
Promoter constants modification	2.0
Cell communication constants modification	2.0
Dimerization	0.6
Activated phosphorylation	0.6
Phosphorylation	1.2
Promotion	0.3
Reaction withdrawal	0.2
Cell communication insertion	0.25

Frequencies of the different mutations. When a network is mutated, first the mutation type is drawn with a probability proportional to the rates indicated in the right column. Then the specific reaction to be mutated is drawn at random among all the reactions of the chosen type. The chosen rates ensure that mutations affecting the topology of the network are less frequent than changes in kinetic constants. They also allow evolution to keep networks of a reasonable size. When a mutation affects a kinetic constant or the initial condition, the affected constant is chosen at random among the whole set of the network constants. The chosen constant is then multiplied by 2^{2r-1} where r is a random number uniformly distributed between 0 and 1. When reactions are added, the inputs are chosen randomly among the existing dynamical variables and an output species is created if needed. When a reaction withdrawal is drawn, with probability $1/2$ removal of either a transcriptional regulation or a post-transcriptional regulation is chosen. For a transcriptional regulation, one is chosen at random among all the existing ones. For a post-transcriptional reaction, a protein species is chosen at random among all the existing one and it is removed when it does not participate in any reaction besides its own production (i.e. it does not regulate a gene or participate in the production of another protein).

Chapter 2

Mathematical analysis of different networks

2.1 The auto-activation complexation network: a bistable switch

The auto-activation and complexation (AAC) switch is based on a protein A that activates the transcription of its own gene and is able to bind another protein B in a complex. A simple description of its dynamics can be based on the following two equations,

$$\frac{dA}{dt} = \rho_A \frac{A}{A_0 + A} - \kappa AB - \delta_A A \tag{2.1}$$

$$\frac{dB}{dt} = \rho_B - \kappa AB - \delta_B B \tag{2.2}$$

Here, we have not explicitly described mRNA dynamics so that ρ_A and ρ_B are the rates of production of protein A and B, κ is the AB complex formation constant and δ_A and δ_B the protein degradation constants. We have also assumed that protein A activates the transcription of gene a with half-maximum activation at a concentration A_0 and a Hill coefficient of one. Finally, the basal transcription rate of a in absence of A has been assumed to be negligible, for simplicity.

In a stationary state, the time derivatives in Eq. (2.1,2.2) vanish. The concentration of B is therefore a function of the concentration of protein A

$$B = \frac{\rho_B}{\delta_B + \kappa A}. \tag{2.3}$$

The possible stationary concentrations of A are thus given by

$$\rho_{eff}(A) - \delta_A A = 0, \tag{2.4}$$

in which we have found it convenient to introduce the effective production of free A (i.e. not complexed with B) $\rho_{eff}(A)$ with

$$\rho_{eff}(A) = \frac{\rho_A A}{A_0 + A} - \frac{\rho_B A}{A_B + A} \tag{2.5}$$

We have also defined the concentration $A_B = \delta_B/\kappa$.

We consider first the limit $A_B \ll A_0$ which is easily understood intuitively. In this limit, which can be achieved with realistic parameter values (1), the production of B $\rho_{eff}(A)$ is equal to zero when $\rho_A < \rho_B$, that is when there are enough B produced to complex all produced A. On the contrary when $\rho_A > \rho_B$, the effective production of A is simply diminished by the maximal complexation rate ρ_B when the concentration of A is high enough to render $\rho_{eff}(A) \geq 0$,

$$\begin{aligned} \rho_{eff}(A) &= 0, \text{ for } A < A_0 \frac{\rho_B}{\rho_A - \rho_B} \\ \rho_{eff}(A) &= \frac{\rho_A A}{A_0 + A} - \rho_B, \text{ for } A \geq A_0 \frac{\rho_B}{\rho_A - \rho_B} \end{aligned} \tag{2.6}$$

As emphasized in ref. (2, 3), this simple subtraction by complexation transforms the non-cooperative activation of gene a into an "ultrasensitive" activation function. Returning to Eq. (2.4), this effective ultrasensitive activation results in the existence of multiple steady states for the AAC network, when the degradation rate of A is low enough. More precisely, this holds when $\delta_A \leq (\sqrt{\rho_A/\rho_B} - 1)^2 \rho_B / A_0$.

The previous approximate analysis simply shows that the AAC network possesses multiple steady states for $\rho_A > \rho_B$ when A_B/A_0 and δ_A are small enough. It is not difficult to refine the argument and obtain the exact boundaries of the bistable domain. Taking A_0 as the concentration scale, it is clear that the shape of $\rho_{eff}(A)$ only depends on the two dimensional parameters $r = \rho_A/\rho_B$ and $\alpha = A_B/A_0$,

$$\rho_{eff}(aA_0) = \rho_B \, a \, R(a), \text{with } R(a) = \frac{r}{1+a} - \frac{1}{\alpha + a} \tag{2.7}$$

With these notations, Eq. (2.4), which determines the possible steady concentrations of A other than 0, reduces to

$$R(a) = \delta_A A_0 / \rho_B \tag{2.8}$$

An elementary calculation shows that the extremum of R for $a > 0$, when it exists, is located at $a^* = (1 - \alpha\sqrt{r})/(\sqrt{r} - 1)$. One concludes that Eq. (2.8) can have two positive solutions only when $r > 1$ and $\alpha < 1/\sqrt{r}$. In this case, when a grows from 0 to a^*, $R(a)$ increases from $R(0)$ to $R(a^*)$ and then it decreases toward 0 when a grows beyond a^*[1]. Two solutions exist when $R(0) < \delta_A A_0/\rho_B < R(a^*)$. Finally, the parameter domain where the AAC network has three steady states, one with $A = 0$ and two with $A > 0$, is given by

$$\rho_A > \rho_B, \ \sqrt{\rho_A/\rho_B} < \kappa A_0/\delta_B, \text{ and } \rho_A/\rho_B - \kappa A_0/\delta_B < \delta_A A_0/\rho_B < \frac{(\sqrt{\rho_A/\rho_B} - 1)^2}{1 - \delta_B/(\kappa A_0)} \tag{2.9}$$

These conditions also imply that the steady state with $A = 0$ and the one with the largest concentration of A are stable as it is easily shown by linearization of Eq. (2.1) and Eq. (2.2). Thus, the conditions of Eq. (2.9) indeed define the parameter regime in which the AAC network is bistable. The different parameter regimes in the (ρ_A, ρ_B) plane are displayed in Fig. 1A.

As a final remark, we note that a bistable regime exists for the AAC network when the degradation of the full AB complex is replaced by a catalytic modification or a catalytic degradation of A by B.

[1]When $a^* < 0$, $R(a)$ is monotonic and Eq. (2.8) can have at most one solution. For $r < 1$ and $\alpha > 1/\sqrt{r}$, the other case when a^* is positive, $R(a)$ decreases until it reaches a negative value $R(a^*)$ and then grows again toward zero. Thus, in this case, Eq. (2.8) cannot have two solutions for $a > 0$ since its r.h.s. is positive.

For simplicity, we consider a non-degraded catalyst B. In this case, Eq. (2.2) is replaced by

$$\frac{dB}{dt} = -\kappa AB + \rho_m C \tag{2.10}$$

$$\frac{dC}{dt} = +\kappa AB - \rho_m C \tag{2.11}$$

where we have denoted by C the AB complex and by ρ_m the rate of production of the modified (or degraded) form of A once the complex is formed. In a stationary state, the concentration of non-complexed B proteins is

$$B = \frac{\rho_m B_T}{\rho_m + \kappa A} \tag{2.12}$$

with B_T the total B concentration. Eq. (2.12) is identical to Eq. (2.3) with the replacements $(\rho_B \leftrightarrow \rho_m B_T, \delta_B \leftrightarrow \rho_m)$. One therefore deduces from the previous analysis that a bistable regime exists when $\rho_m B_T < \rho_A < \kappa^2 A_0^2 B_T/\rho_m$ and the degradation rate of A lies in the range $\rho_A/A_0 - B_T\kappa < \delta_A < (\sqrt{\rho_A} - \sqrt{\rho_m B_T})^2/(A_0 - \rho_m/\kappa)$.

In Fig. 1, the graph has been drawn for a dimensionless value of the complex formation rate equal to 10 ($\kappa A_0/\delta_B = 10$).

2.2 The mixed-feedback loop in the bistable regime

The mixed-feedback loop is a simple network based on two genes a and b with two interactions of different kinds. On the one hand, protein A produced by gene a regulates the transcription of gene b. On the other hand, it can form a transcriptionally inactive complex with protein B. The different possible dynamical regimes of the MFL have been studied previously (1). Here, we limit ourselves to the case depicted in Fig. 1B where protein A is a repressor of gene b transcription. In this regime, the MFL can be studied without explicitly describing mRNA dynamics or the binding dynamics of protein A to the promoter of gene b. This reduces the MFL dynamics description to the following two equations,

$$\frac{dA}{dt} = \rho_A - \kappa AB - \delta_A A \tag{2.13}$$

$$\frac{dB}{dt} = \rho_B \frac{A_0}{A_0 + A} - \kappa AB - \delta_B B \tag{2.14}$$

where ρ_A and ρ_B are the rates of production of protein A and B, κ is the AB complex formation constant and δ_A and δ_B the protein degradation constants. We have also assumed that protein A represses the transcription of gene b with half-maximum repression at a concentration A_0 and a Hill coefficient of one, which is sufficient for bistability. Finally, the transcription rate of b in the fully repressed state has been assumed to be negligible, for simplicity.

In a stationary state, the concentration of B is given as a function of the concentration of A by

$$B = \frac{\rho_B A_0}{(A_0 + A)(\delta_B + \kappa A)} \tag{2.15}$$

The production $\rho_{eff}(A)$ rate of free A is the total production rate of A diminished by the formation rate of the complex between A and B,

$$\rho_{eff}(A) = \rho_A - \frac{\rho_B A}{(1 + A/A_0)(A_B + A)} \tag{2.16}$$

where we have introduced the concentration $A_B = \delta_B/\kappa$ of protein A above which complexation with A dominates over degradation for the dynamics of B. As stressed in ref. (2, 3), complexation gives to $\rho_{eff}(A)$ a shape similar to what strong cooperativity would produce, which renders bistability possible. The stationary concentrations of A are obtained when degradation of A and complexation balance the production of A

$$\rho_A = \delta_A A + \frac{\rho_B A}{(1 + A/A_0)(A_B + A)} \tag{2.17}$$

Eq. (2.17) depends on three dimensionless parameters

$$r_A = \frac{\rho_A}{\delta_A A_0},\ r_B = \frac{\rho_B}{\delta_A A_0},\ \eta = \frac{A_B}{A_0} = \frac{\delta_B}{\kappa A_0}. \tag{2.18}$$

Bistability is possible when the right-hand-side (r.h.s) of Eq. (2.17) is non monotonic that is for

$$r_B \geq (1 + \eta^{1/3} + \eta^{2/3})^3 \tag{2.19}$$

In this regime, the r.h.s of Eq. (2.17) has an N-shape as a function of A. Bistability is achieved when Eq. (2.17) has three roots, that is when ρ_A stands between the local maximum and minimum of the r.h.s. of Eq. (2.17). Simple analytic expressions are easily obtained in the biologically relevant limit $A_B \ll A_0$, that is $\eta \ll 1$. Then, one obtains for the lower limit of ρ_A, in dimensionless form,

$$r_A \geq 2\sqrt{r_B} - 1 \tag{2.20}$$

The upper limit is simply $r_A \leq r_B$ to zeroth order in η. Keeping the first η-correction one obtains

$$r_A \leq r_B - 2\eta\sqrt{r_B(r_B - 1)} \tag{2.21}$$

The two approximate expressions (2.20) and (2.21) have been plotted in Fig. 1B together with the numerically determined exact boundary of the bistability domain for a dimensionless value of the complex formation rate equal to 100 ($\kappa A_0/\delta_B = 100$).

2.3 A bistable two-gene network based on purely transcriptional non-cooperative interactions

When the interactions were restricted to be purely transcriptional and non-cooperative, the algorithm repeatedly created the two-gene network shown in Fig. S4.2C. In this network, A transcriptionally activates its own gene. In parallel with this direct activation loop, A also indirectly activates its own gene by transcriptionally repressing its repressor B. We show, here, that this simple network can indeed be bistable even with non-cooperative transcriptional interactions.

The network dynamics can be described by the two equations

$$\frac{dA}{dt} = \frac{\rho_A A/A_0}{1 + A/A_0 + B/B_0} - \delta_A A \tag{2.22}$$

$$\frac{dB}{dt} = \frac{\rho_B}{1 + A/A_B} - \delta_B B \tag{2.23}$$

where, for simplicity, mRNA dynamics are not explicitly described, and we have taken the basal production rate of A and the fully repressed production rate of B to be zero.

Can the simple system of eqs (2.22, 2.23) support multiple stable steady states? In a steady state, the production of B is equal to its degradation, and the steady concentration of B is related to that of A by

$$B(A) = \frac{\rho_0/\delta_B}{1 + A/A_B} \tag{2.24}$$

Given this relation, the steady concentration of A is either equal to 0 or obey the equation

$$f(A) \equiv \frac{\rho_A/A_0}{1 + A/A_0 + B(A)/B_0} = \delta_A \tag{2.25}$$

Linearization of Eq. (2.22) shows that the $A = 0$ solution is stable when the degradation rate of A is larger than its production in presence of the maximal possible steady concentration of B, $B = \rho_B/\delta_B$, namely when

$$\delta_A > \frac{\rho_A/A_0}{1 + \rho_B/(\delta_B/B_0)} \tag{2.26}$$

The non-zero steady concentrations of A are given by the solution of Eq. (2.25). The function on the l.h.s of Eq. (2.25) decreases towards 0 when A grows. Given the simple form of $f(A)$, a computation of its derivative shows that two different cases can be distinguished depending on the relative magnitude of the two concentrations, A_0 and A_B, governing transcriptional activation and repression by A. When A_B/A_0 is large enough , f monotonically decreases starting from $A = 0$ while, for small enough A_B/A_0, it begins by increasing at low A concentration before reaching a maximum and eventually decreasing. A computation of the derivative of f straightforwardly gives the exact criterion,

$$\frac{A_B}{A_0} < \frac{\rho_B}{\delta_B B_0} \tag{2.27}$$

When f decreases monotonically (i.e. when the inequality (2.27) does *not* hold), Eq. (2.25) shows that a single non-trivial steady state exists as soon as $\delta_A < f(0)$, a condition which exactly corresponds to the loss of stability for the $A = 0$ state, i. e. to the violation of Eq. (2.26). As explained below, a simple linear stability calculation shows that the non-trivial state is stable. Therefore in the case when f decreases monotonically , which corresponds to a situation where the network indirect loop is hard to activate, the network has always a single stable state, either $A = 0$ when Eq. (2.26), holds or a non-trivial one with $A \neq 0$ that replaces it otherwise.

When the condition (2.27) is obeyed, the network properties are more interesting. As A grows the function f grows until $A = A^*$, reaches its maximum $f(A*)$ and then decreases toward 0, with

$$A^* = A_B \left(\sqrt{\frac{\rho_B}{\delta_B B_0} \frac{A_0}{A_B}} - 1 \right) \tag{2.28}$$

Therefore, there are two supplementary fixed points, in addition to the trivial $A = 0$ one, when

$$f(0) = \frac{\rho_A / A_0}{1 + \rho_B / (\delta_B B_0)} < \delta_A < f(A^*) = \frac{\rho_A / A_0}{1 + 2\sqrt{(\rho_B / \delta_B B_0)(A_B / A_0)} - A_B / A_0} \tag{2.29}$$

Linear stability of a non-trivial fixed point A_s is simple to perform. The trace of the 2×2 stability matrix is always negative, while its determinant has the sign of $-f'(A_s)$. This readily shows that the non-trivial fixed points with the larger value of A is stable (the determinant is positive) while the other one is unstable.

In summary, the simple network of Fig. S4.2C is bistable in the parameter regime where both Eq. (2.27) and (2.29) are obeyed.

Before concluding, we note a feature of this circuit that surprised us, at first. When looking for bistable circuit based on non-cooperative transcriptional interactions, simulated evolution repeatedly produced the network of Fig. S4.2C where an indirect activation loop is made by a chain of two repressions. We expected that bistability could also be produced by having instead an indirect loop made of two activations but this never appeared in the simulations. Two non-cooperative activations in fact cannot produce bistability, when transcriptional regulation is described by Eq. (2) of the main text, where different factors compete for the same binding site. This can be simply seen as follows.

In the general case, the equation for A is

$$\frac{dA}{dt} = \frac{\rho_0 + \rho_A A / A_0 + \rho_B B / B_0}{1 + A / A_0 + B / B_0} - \delta_A A \tag{2.30}$$

The steady states are determined when production equals degradation of A, that is for

$$\frac{\rho_0 / A + \rho_A / A_0 + \rho_B B(A) / (A B_0)}{1 + A / A_0 + B(A) / B_0} = \delta_A \tag{2.31}$$

where $B(A)$ is the stationary concentration of B for a given concentration of A. When A activates the transcription of B, $B(A)$ is an increasing function of A as well as the denominator of the l.h.s. of Eq. (2.31). However, when A activates the transcription of B in a non cooperative fashion, $B(A)/A$ is always a decreasing function of A [2]. Therefore, the numerator of l.h.s. of Eq. (2.31) is also a decreasing function of A (since it is a sum of two decreasing functions) as well as the l.h.s of Eq. (2.31). Therefore, Eq. (2.31) cannot have multiple solutions and two non-cooperative activations cannot produce a bistable circuit in the case of Fig. S4.2C.

2.4 Simple bistable two-cell networks

2.4.1 Interaction between homologous proteins in the signal-sending and signal-receiving cells

In the simplest network of two interacting cells produced by the evolutionary algorithm, a single protein that we call A both "sends" the signal in one cell and "receives" it in the other cell. In this

[2] In the non-cooperative case, the stationary concentration of B reads $B(A) = \frac{\rho_0^b + \rho_A^b A / A_B}{1 + A / A_B}$. The function $B(A)/A = \frac{\rho_0^b / A + \rho_A^b / A_B}{1 + A / A_B}$ is readily seen to be decreasing as a ratio of decreasing functions

case, the two-cell network dynamics is described by the following two equations

$$\frac{dA_1}{dt} = \frac{\rho_0 A_0 + \rho_1 A_1}{A_0 + A_1} - \gamma A_1 A_2 - \delta A_1 \tag{2.32}$$

$$\frac{dA_2}{dt} = \frac{\rho_0 A_0 + \rho_1 A_2}{A_0 + A_2} - \gamma A_1 A_2 - \delta A_2 \tag{2.33}$$

where A_1 and A_2 denote the concentrations of protein A in the two cells. For simplicity, we have not explicitly described mRNA dynamics so that ρ_0 denote the basal rate of production of protein A, ρ_1 its maximum production rate, δ its degradation constant and γ quantify the strength of the interaction between the two cells. We have also assumed that protein A activates the transcription of gene a with half-maximum activation at a concentration A_0 and a Hill coefficient of one.

We begin by assuming that the basal transcription rate of a in absence of A is negligible. That is, we begin by taking $\rho_0 = 0$ and study afterwards the changes brought by a non-zero ρ_0.

We first consider the possible stationary states of Eq. (2.32,2.33). The l.h.s. of Eq. (2.32) vanishes when either $A_1 = 0$ or

$$A_2 = \frac{1}{\gamma}\left[-\delta + \frac{\rho_1}{A_0 + A_1}\right] \tag{2.34}$$

Similarly, the stationarity of Eq. (2.33) gives either $A_2 = 0$ or the equation obtained by exchanging the role of A_1 and A_2 in Eq. (2.34).

$$A_1 = \frac{1}{\gamma}\left[-\delta + \frac{\rho_1}{A_0 + A_2}\right] \tag{2.35}$$

When $\rho_1/A_0 < \delta$, the production rate of A is smaller than its degradation rate for non-interacting cells. This is *a fortiori* true when they are interacting and therefore $A_1 = A_2 = 0$ is the only stationary state.

When $\rho_1/A_0 > \delta$, the stationary state $A_1 = A_2 = 0$ is unstable and there are three other possible states, the two asymmetric states $(A_1 = 0, A_2 = \rho_1/A_0 - \delta)$ and $(A_1 = \rho_1/A_0 - \delta, A_2 = 0)$, and the symmetric state obtained by the intersection of the curves given by Eq. (2.34) and (2.35). The curve given by Eq. (2.34) starts for $A_1 = 0$ at $A_2 = A_i$ with

$$A_i = \frac{1}{\gamma}\left[-\delta + \frac{\rho_1}{A_0}\right]. \tag{2.36}$$

It ends when A_2 vanishes for $A_1 = A_f$ with $A_f = \rho_1/A_0 - \delta$. Two cases can be distinguished depending upon whether $\gamma < \delta/A_0$ for which $A_i < A_f$ or $\gamma > \delta/A_0$ for which $A_i > A_f$. It is not difficult to check that in the former case the symmetric state is stable and the asymmetric states are unstable whereas for $\gamma > \delta/A_0$ the reverse is true. The cells therefore spontaneously assumes two different states as soon as the interaction strength γ is greater than a critical $\gamma_c = \delta/A_0$. At $\gamma = \gamma_c$ itself, the bifurcation described by Eq. (2.32,2.33) (for $\rho_0 = 0$) is non-generic since Eq. (2.34) and (2.35) coincides and there is a line of fixed points.

For $\rho_0 \neq 0$, the dynamics is qualitatively similar in that there is a symmetric stationary state that is stable for small γ and unstable for large γ. In this symmetric state, the concentration A_s of A in the two cells obeys

$$\gamma A_s^2 + \delta A_s = \frac{\rho_0 A_0 + \rho_1 A_s}{A_0 + A_s} \tag{2.37}$$

Comparison of the two sides of Eq. (2.37) shows that A_s always exists and is uniquely determined. When A_s increases from zero to infinity, the l.h.s of Eq.(2.37) increases from zero to infinity with an

increasing slope whereas its r.h.s increases from $\rho_0 > 0$ to ρ_1 with a decreasing slope. It is therefore clear that these two functions cross at a unique point. Moreover, since the l.h.s of Eq.(2.37) is an increasing function of γ, A_s decreases when γ increases and actually approaches 0.

The stability of this symmetric solution is easily studied by linearization of Eqs. (2.32,2.33). The symmetry between the two cells implies that the two eigenmodes are either symmetric or antisymmetric with real eigenvalues. The two modes are stable when $\gamma = 0$ (non-interacting cells) and for γ small enough. When the interaction γ between the cells is increased, it is found that the antisymmetric mode always becomes instable before the symmetric one. This happens when the corresponding eigenvalue crosses zero i.e for

$$\lambda_a = \frac{(\rho_1 - \rho_0)A_0}{(A_0 + A_s)^2} - \delta = 0. \tag{2.38}$$

As stated, for $\gamma = 0$, λ_a is negative and the symmetric state is stable. When γ increases, A_s decreases and λ_a increases. It crosses zero at a finite $\gamma = \gamma_c$, i.e. Eq. (2.38) is satisfied, if and only λ_a is positive for $A_s = 0$. That is, a necessary condition for instability is

$$\rho_1 - \rho_0 > A_0\delta. \tag{2.39}$$

When the condition (2.39) is realized, the symmetry between the two cells is spontaneously broken when $\gamma > \gamma_c$. It is not difficult to see from Eq.(2.37,2.38) that when ρ_1 becomes sufficiently large, the finite basal production rate ρ_0 becomes irrelevant and γ_c tends toward δ/A_0, its $\rho_0 = 0$ value. In the other limit, when ρ_1 tends towards its minimum value $\rho_0 + A_0\delta$, larger and larger interaction strengths are needed to break the symmetry between the two cells with γ_c diverging as

$$\gamma_c \sim \frac{4\delta^2\rho_0}{(\rho_1 - \rho_0 - A_0\delta)^2} \tag{2.40}$$

Close to the bifurcation line, a weakly nonlinear calculation shows that the amplitude F_a of the antisymmetric mode obeys an equation of the form

$$\frac{d}{dt}F_a = \lambda_a F_a + \kappa F_a^3 \tag{2.41}$$

with κ proportional to $\delta/A_0 - \gamma_c$ and of the same sign. Therefore, κ is negative and the transition is continuous (of the so-called "supercritical pithfork" type) since, for $\rho_0 \neq 0$, γ_c is larger than δ/A_0, the critical interaction strength for a zero basal protein production rate. The transition is however unusually sharp for $\rho_1 \gg \rho_0$ since κ vanishes in this limit, a trace of the anomalous bifurcation behavior for $\rho_0 = 0$.

On Fig. 3, the plot has been drawn in dimensionless units by measuring the values of ρ_1 and γ relative to the relevant effective production rate δA_0 and effective interaction δ/γ. The basal protein production rate has been chosen such that $\rho_0/(\delta A_0) = 0.4$. For comparison, the dashed lines mark the boundary of the bistability domain when $\rho_0 = 0$.

2.4.2 Interaction between heterologous proteins in the signal-sending and signal-receiving cells: a network purely based on post-transcriptional interactions

One the most commonly produced network bistable two-cell networks with an interaction between different proteins in the signal-sending and signal-receiving cell only relies on post-transcriptional

interactions. As depicted in Fig. 4A, signal transmission is obtained by the transformation of protein B into the modified form B^* in the signal-receiving cell due to the presence of A in the signal-sending cell. Quite remarkably, evolutionary simulations show that the simple addition of a complexation between A and B^* is sufficient to create a parameter regime in which the network is bistable. In order to show it explicitly and better understand the dynamics of this basic network, we analyze it mathematically in the following. The network dynamics is described by the following three equations in cell 1

$$\frac{dA_1}{dt} = \rho_A - \kappa A_1 B_1^* - \delta_A A_1 \tag{2.42}$$

$$\frac{dB_1}{dt} = \rho_B - \gamma A_2 B_1 - \delta_B B_1 \tag{2.43}$$

$$\frac{dB_1^*}{dt} = \gamma A_2 B_1 - \kappa A_1 B_1^* - \delta_{B^*} B_1^* \tag{2.44}$$

where A_1 and A_2 denote the concentrations of protein A in the two cells and B_1 and B_1^*, the concentrations in cell 1 of the two forms of protein B. Of course, the analogous set of equations with permuted indices 1 and 2, describes the dynamics in cell 2. For simplicity, we have not explicitly described mRNA dynamics so that ρ_A and ρ_B denote the production rate of protein A and B. The protein degradation constants are respectively, δ_A, δ_B and δ_{B^*} and γ and κ respectively quantify the strengths of the interaction between the two cells and the kinetics of complex formation between A and B^*.

In order to analyze the network dynamics described by Eq. (2.42,2.43,2.44), we first determine the steady concentrations for two cells in identical states and then analyze the stability of this symmetric state. In a steady symmetric state, one has

$$B = \frac{\rho_B}{\gamma A + \delta_B} \tag{2.45}$$

$$B^* = \frac{\gamma A B}{\kappa A + \delta_{B^*}} \tag{2.46}$$

with A determined by

$$\rho_A - \delta_A A = \rho_B \frac{\gamma\, \kappa A^2}{(\gamma A + \delta_B)(\kappa A + \delta_{B^*})} \tag{2.47}$$

Eq. (2.47) clearly has a unique positive solution since, as A increases from zero, the l.h.s of Eq. (2.47) decreases linearly from ρ_A while its r.h.s increases monotonically from zero to ρ_B.

To perform the stability analysis of this steady state in a simple but relevant limit, we focus on the case when complexation is the fastest process. That is, we suppose that κ is large enough so that B^* (Eq. (2.44)) quickly reaches the steady value $B_1^* \simeq \gamma A_2 B_1/(\kappa A_1)$, while A and B protein concentrations evolve on a slower time scale. In this fast complexation limit, the dynamics reduce to that of the two-dimensional system

$$\frac{dA_1}{dt} = \rho_A - \gamma A_2 B_1 - \delta_A A_1 \tag{2.48}$$

$$\frac{dB_1}{dt} = \rho_B - \gamma A_2 B_1 - \delta_B B_1 \tag{2.49}$$

The stability of the steady symmetric solution is easily studied by linearization of Eqs. (2.48,2.49). The symmetry between the two cells implies that the two eigenmodes are either symmetric or antisymmetric. The symmetric mode is found to be always stable. For the antisymmetric mode, the linearization of Eqs. (2.48,2.49) is obtained by writing the protein concentrations under the form, $A_1 = A + a, B_1 = B + b, B_1^* = B^* = B^* + b^*, A_2 = A - a, \cdots$. The linear evolution around the symmetric steady values A, B, B^* (Eqs. (2.45-2.47)), then reads,

$$\frac{da}{dt} = (\gamma B - \delta_A)a - \gamma A b \tag{2.50}$$

$$\frac{db}{dt} = \gamma B a - (\gamma A + \delta_B)b \tag{2.51}$$

Stability of the symmetric state requires that the two eigenvalues of the linear dynamics matrix L have negative linear parts, namely that their sum $\mathrm{tr}(L)$ is negative and their product $\det(L)$ positive

$$\mathrm{tr}(L) = \gamma B - \gamma A - \delta_A - \delta_B < 0 \tag{2.52}$$

$$\det(L) = -\delta_B \gamma B + \delta_A \gamma A + \delta_A \delta_B > 0 \tag{2.53}$$

Two cases can be distinguished depending on the relative values of the degradation constants of A and B.

We consider first the case $\delta_B > \delta_A$. In this case, division of Eq. (2.53) by δ_A readily shows that the determinant condition (2.53) is stronger than the condition (2.52) on the trace. The stability boundary is thus given by the vanishing of $\det(L)$ which can be written with the help of Eq. (2.45,2.47)

$$\rho_B = \rho_A + \frac{\delta_B \delta_A}{\gamma} \tag{2.54}$$

Below the line (2.54), the linear system (2.50, 2.51) has two real and negative eigenvalues. When the line (2.54) is traversed from below, one of the two eigenvalues becomes positive rendering the symmetric state unstable. This corresponds to the appearance of bistability and spontaneous symmetry breaking between the two cells.

Before leaving the analysis of the case $\delta_B > \delta_A$, we note a subtlety. The stability boundary (2.54) is obtained in the limit in which κ is large and all the other parameters are fixed. The elimination of B^* is no longer valid when the steady-state value of A (Eq. (2.47)) is small enough so that $\kappa A \sim 1$, that is for $\rho_A \sim 1/\kappa$. An alternative reduction of Eq. (2.42,2.43,2.44) can be performed in the limit of large κ and small $\rho_A \sim 1/\kappa$. By supposing that $A_1 \sim 1/\kappa$ and $B_1^* \sim 1/\kappa$, one obtains that B_1 quickly reaches the steady value $B_1 = \rho_B/\delta_B$ and can be eliminated from the dynamics. The linear stability boundary of the remaining two-dimensional systems for the symmetric and antisymmetric A and B^* modes can be performed as above and it provides the stability condition

$$\rho_B = \frac{\delta_B \delta_A}{\gamma} \left(\frac{\rho_A + \delta_A \delta_B^*/\kappa}{\rho_A - \delta_A \delta_B^*/\kappa} \right)^2 \tag{2.55}$$

Finally, Eq. (2.54) and (2.55) can be combined into a uniform approximation of the stability boundary

$$\rho_B = \rho_A + \frac{\delta_B \delta_A}{\gamma} \left(\frac{\rho_A + \delta_A \delta_B^*/\kappa}{\rho_A - \delta_A \delta_B^*/\kappa} \right)^2 \tag{2.56}$$

Eq. (2.56) is shown in Fig. 4B together with the exact boundary of the bistable domain.

When $\delta_B < \delta_A$, the determinant condition (2.53) is no longer stronger than the trace condition (2.52). It is therefore possible than the symmetric state looses stability with $\mathrm{tr}(L)$ becoming positive. This corresponds to the real part of two complex eigenvalues becoming positive and to the appearance of oscillations. The condition $\mathrm{tr}(L) > 0$ can be translated on the parameter conditions

$$\rho_A < \rho_b + \frac{\delta_A - \delta_B}{2}\sqrt{(\delta_A/\gamma)^2 + 4\rho_B/\gamma} - \frac{\delta_A(\delta_A + 3\delta_B)}{2\gamma} \tag{2.57}$$

For fast complexation between A and B^*, linear analysis thus delimits the oscillatory region as the region bounded by Eq. (2.57) and Eq. (2.56) (i.e such that $\det(L) > 0$). A dynamical trace with cells oscillating in antiphase is shown in Fig. 4C.

On Fig. 4, the kinetic parameters used to draw the graph are $\delta_{B^*} = 1$, $\gamma = 1, \kappa = 10$. For (C) : $\delta_A = 1 > \delta_B = 0.25$ with $\rho_A = 2, \rho_B = 3$, and the other kinetic parameters as in B.

Chapter 3

Description of the different networks

We display here networks as they were created by our evolutionary algorithm. We use a more abstract notational convention to be able to display all the parameters that are required to simulate them. This provides a full description for all the networks that were displayed in the figures of the main text. Please note that we reproduce here, the networks as they were created by the evolutionary algorithm. The labels on the genes and proteins may thus be different from the labels in the figures of the main text (for instance, the two genes in the AAC network are denoted a and b in Fig. 1, but in the example displayed in Fig. S3.2, evolution used gene a and c to create the corresponding network).

Our notational conventions are as follows (an explicit example is provided below):
- upper red boxes display the different genes as well as the parameters (ρ, δ, α) describing basal mRNA production (ρ), mRNA stability (δ), and translation efficiency (α),
- ovals enclose the corresponding proteins and their degradation rates (δ),
- arrows describe regulatory interactions and their strengths. An arrow from an oval to a box describes a transcriptional regulation of the boxed gene by the protein in the oval, with (K, ρ) the dissociation constant and activated (or repressed) transcription rate and (for gene a and protein P_j, $K = (P_j^{(0,a)})^{-h}$ and $\rho = \rho_j$ with the notations of Eq. (2) of the main text). Hill coefficients are equal to one in all figures, except when otherwise specified. Arrows converging from two ovals denote a post-transcriptional interaction between two proteins, that is either a complex formation (solid arrows) or a catalytic modification of one of the protein (indicated by a solid arrow) by another (indicated by a dashed arrow). The kinetic constant for this interaction is shown at the arrow ends and another arrow points to the oval corresponding to the formed entity (the complex, denoted by concatenating the name of the two constitutive proteins, or the modified protein, denoted by the addition of a star to the name of the mother protein).

As an example for the convenience of the reader, we explicitly provide the equation corresponding to the AAC switch displayed in Fig. S3.1B with a_m and e_m the concentrations of the two mRNAs corresponding, to the two genes a and e, A and E the concentrations of the two proteins and AE the

concentration of their complex.

$$\frac{da_m}{dt} = \frac{0.19 + 9 \times 2.1\,A}{1 + 2.1\,A} - 1.8\,a_m \tag{3.1}$$

$$\frac{de_m}{dt} = 2 - 1.8\,e_m \tag{3.2}$$

$$\frac{dA}{dt} = 1.1\,a_m - 0.1A - 1.4\,A \times E \tag{3.3}$$

$$\frac{dE}{dt} = 0.47\,e_m - 0.2\,E - 1.4\,A \times E \tag{3.4}$$

$$\frac{dAE}{dt} = 1.4\,A \times E - 0.1\,AE \tag{3.5}$$

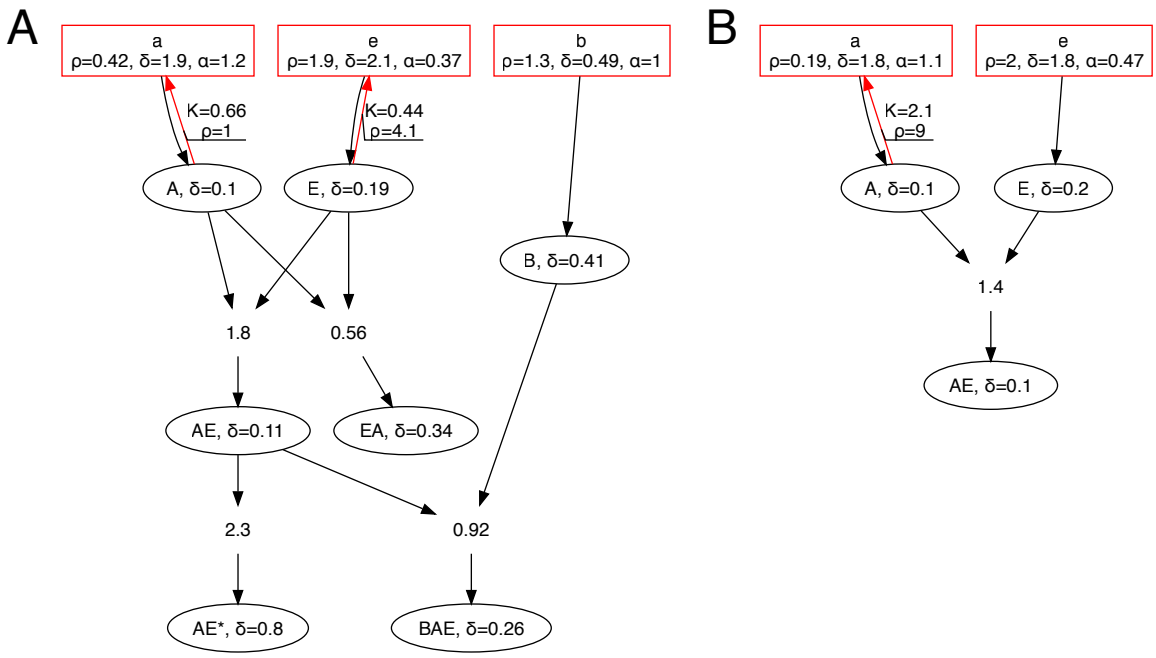

Figure S3.1: Examples of network pruning performed by the algorithm. A) A cell-autonomous bistable network after convergence but before the pruning phase. B) The same network after the pruning phase. Note that since evolution proceeds during the pruning phase, the networks kinetic parameters are modified during the pruning phase and some of them are different after and before pruning.

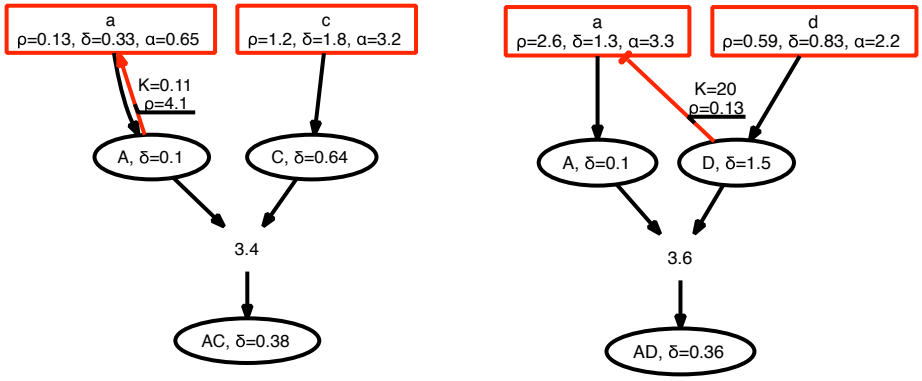

Figure S3.2: Examples of a AAC network (left) and Mixed-Feedback-Loop (right) created by the algorithm.

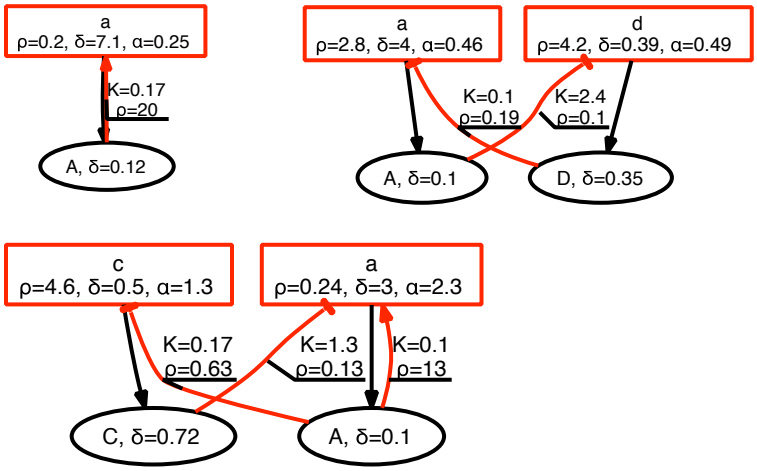

Figure S3.3: Examples of bistable purely transcriptional networks corresponding to the networks of Fig. 1 in the main text. Transcriptional interactions have Hill coefficients equal to 2 for the two top networks.

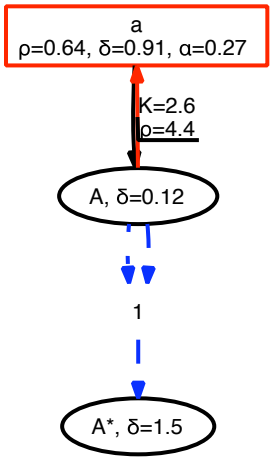

Figure S3.4: An example of the simplest network, pictured in Fig. 3 in the main text, for which interactions allow two identical cells to be in different states.

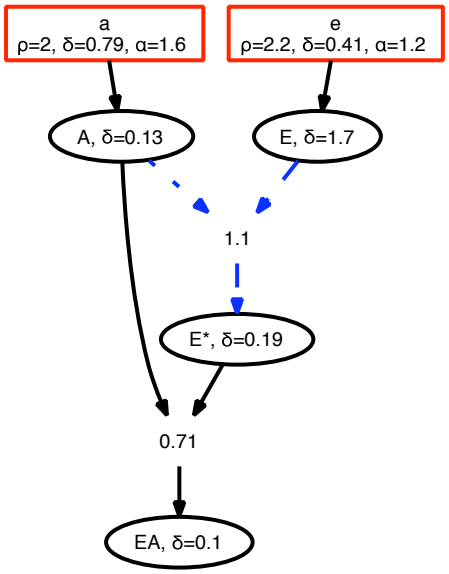

Figure S3.5: An example of the type 1 created bistable network displayed in Fig. 4. In the example shown, the interaction takes place between A, in the signal-sending side, and E, in the signal-receiving side.

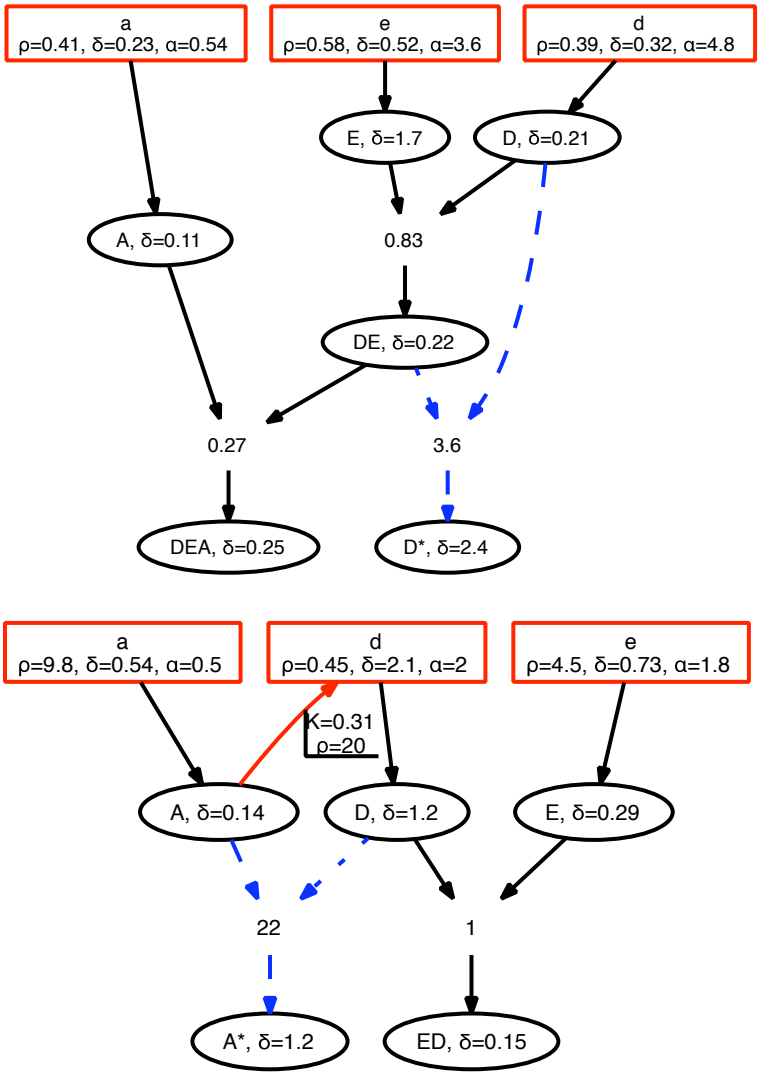

Figure S3.6: Created examples of the type 2 networks displayed in Fig. 5 in the main text.

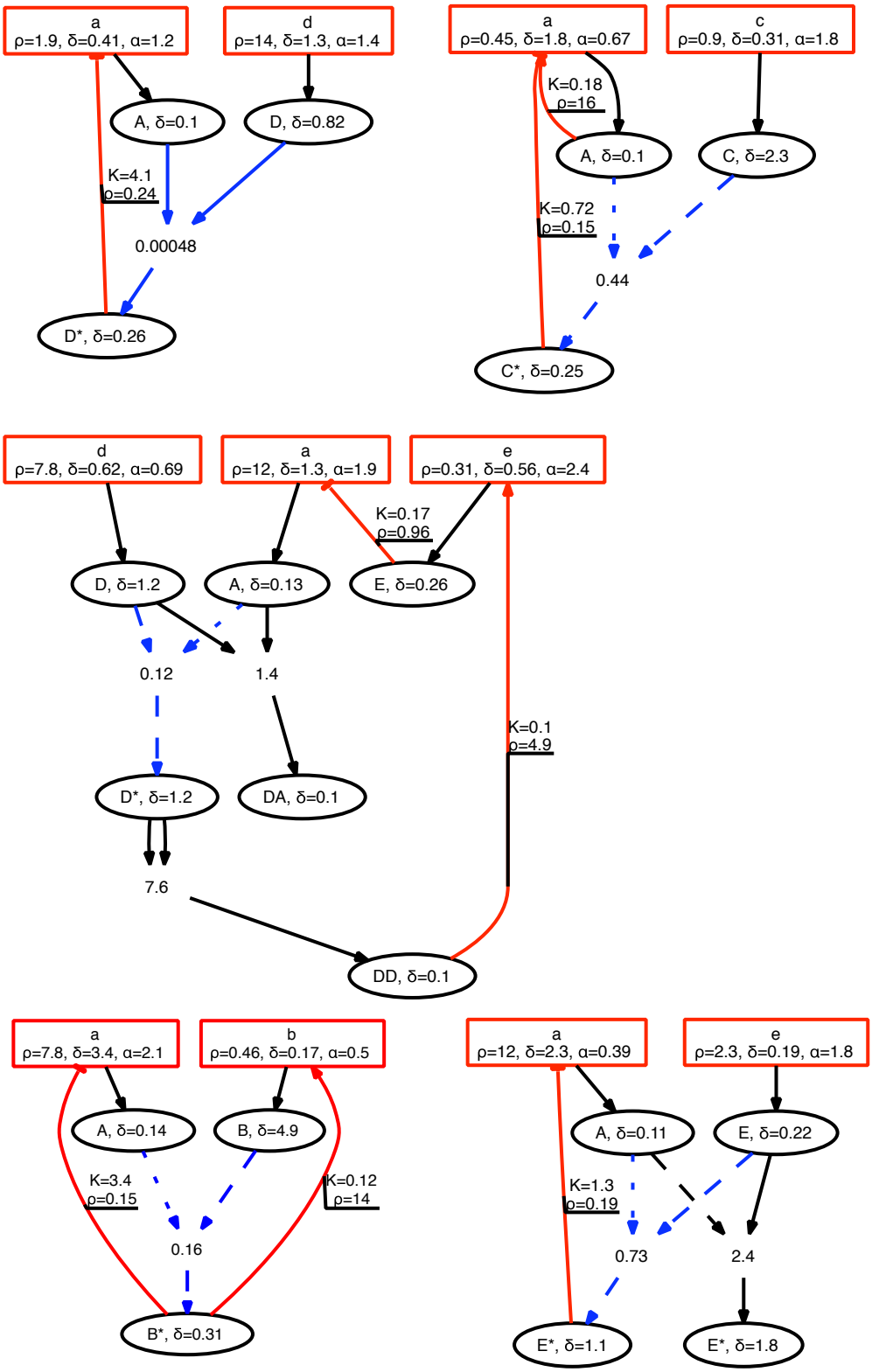

Figure S3.7: Created examples of the type 3 network displayed in Fig. 6 in the main text. Note that the transcriptional inhibition by the modified signal-receiving protein is actually implemented by a dimer of proteins D^* in the shown bottom-left network .

Chapter 4

Supplementary figures

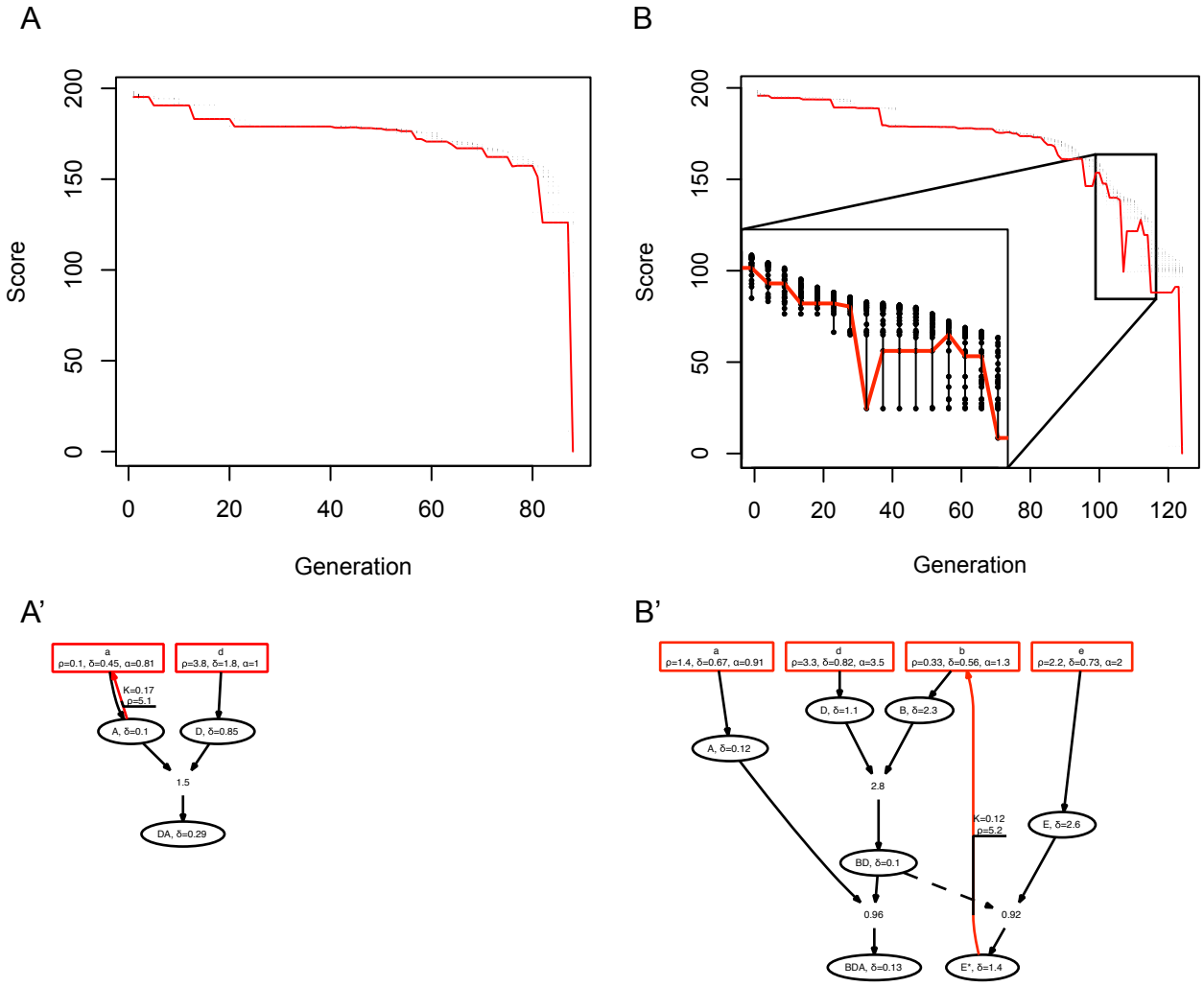

Figure S4.1: **Examples of score evolutions in two simulation runs.** The score evolution is shown in two evolutionary simulations runs leading to the creation of cell-autonomous bistable networks. (A, B) The score of each network in the evolving population is shown as a black dot. The red line shows the scores of the ancestors of the final network. The score evolution during 16 generations is enlarged in panel B insert to show the non-monotonic variation of these ancestor scores. (A′,B′) The final networks in the two runs displayed as explained above with all their kinetic constants.

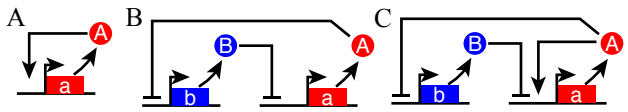

Figure S4.2: **Simple bistable networks purely based on transcriptional interactions**. (A) Simple autoactivation loop. (B) Repression of a repressor. (C) Auto-activation and repression of a repressor. While the designs shown in A and B require Hill coefficients greater than one, the network shown in C is bistable even when all transcriptional interactions are described by Hill coefficients equal to one.

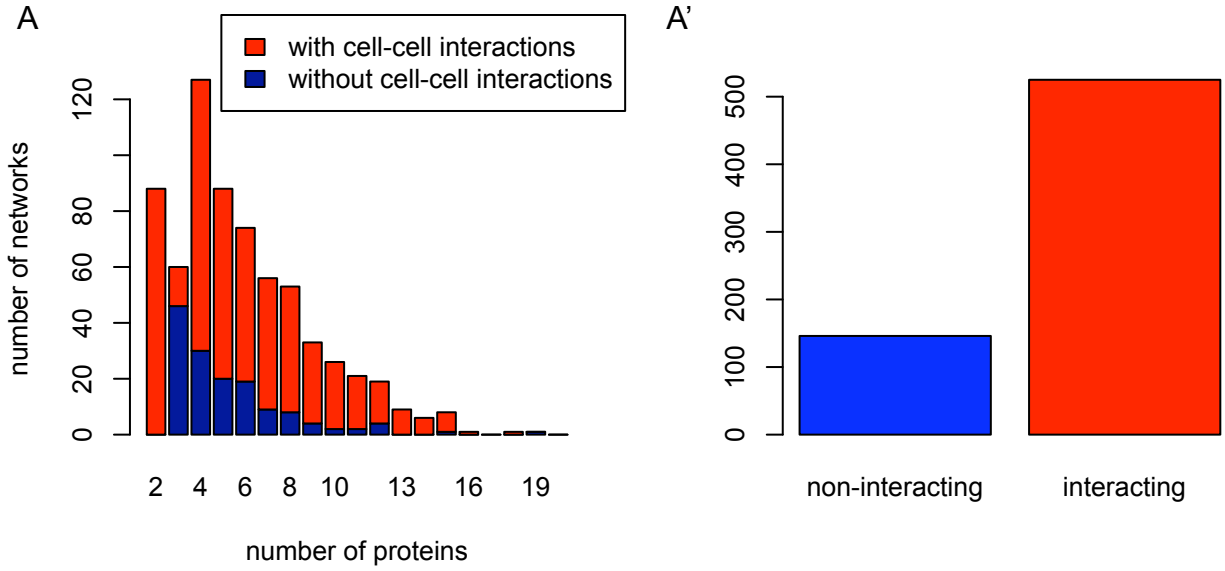

Figure S4.3: **Two cell network statistics for larger maximal numbers of allowed proteins and interactions.** Same as Fig. 2 for simulation runs in which a larger maximal number of proteins and regulations were allowed (see the main text and section 1).

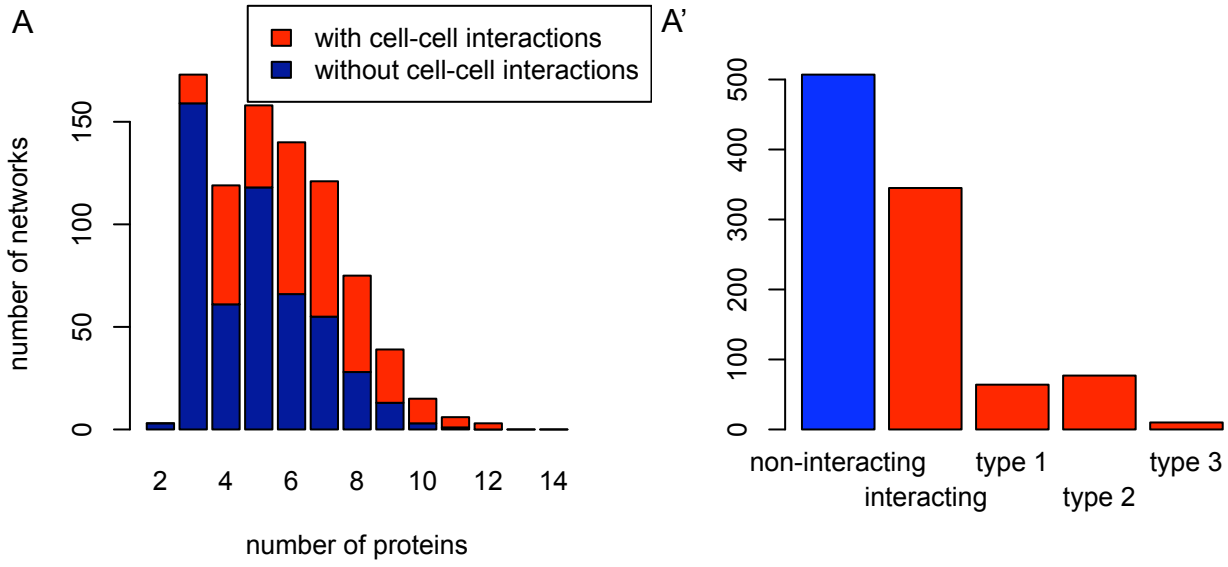

Figure S4.4: **Two cell network statistics for a steep initial gradient.** Same as Fig. 2 but for steep initial gradient of A (see the main text and section 1.6).

Bibliography

1. Francois, P., and V. Hakim, 2005. Core genetic module: the mixed feedback loop. *Phys Rev E Stat Nonlin Soft Matter Phys* 72:031908.

2. Buchler, N. E., and M. Louis, 2008. Molecular titration and ultrasensitivity in regulatory networks. *J. Mol. Biol.* 384:1106–1119.

3. Buchler, N. E., and F. R. Cross, 2009. Protein sequestration generates a flexible ultrasensitive response in a genetic network. *Mol. Syst. Biol.* 5:272.