BMC Systems Biology

Research article

Threshold-dominated regulation hides genetic variation in gene expression networks

Arne B Gjuvsland*[1,3], Erik Plahte[2,3] and Stig W Omholt[1,3]

Address: [1]Department of Animal Science and Aquaculture, Norwegian University of Life Sciences, 1432 Ås, Norway, [2]Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, 1432 Ås, Norway and [3]Centre for Integrative Genetics (CIGENE), Norwegian University of Life Sciences, 1432 Ås, Norway

Email: Arne B Gjuvsland* - arne.gjuvsland@cigene.no; Erik Plahte - erik.plahte@umb.no; Stig W Omholt - stig.omholt@cigene.no

* Corresponding author

Abstract

Background: In dynamical models with feedback and sigmoidal response functions, some or all variables have thresholds around which they regulate themselves or other variables. A mathematical analysis has shown that when the dose-response functions approach binary or on/off responses, any variable with an equilibrium value close to one of its thresholds is very robust to parameter perturbations of a homeostatic state. We denote this threshold robustness. To check the empirical relevance of this phenomenon with response function steepnesses ranging from a near on/off response down to Michaelis-Menten conditions, we have performed a simulation study to investigate the degree of threshold robustness in models for a three-gene system with one downstream gene, using several logical input gates, but excluding models with positive feedback to avoid multistationarity. Varying parameter values representing functional genetic variation, we have analysed the coefficient of variation (*CV*) of the gene product concentrations in the stable state for the regulating genes in absolute terms and compared to the *CV* for the unregulating downstream gene. The sigmoidal or binary dose-response functions in these models can be considered as phenomenological models of the aggregated effects on protein or mRNA expression rates of all cellular reactions involved in gene expression.

Results: For all the models, threshold robustness increases with increasing response steepness. The *CV*s of the regulating genes are significantly smaller than for the unregulating gene, in particular for steep responses. The effect becomes less prominent as steepnesses approach Michaelis-Menten conditions. If the parameter perturbation shifts the equilibrium value too far away from threshold, the gene product is no longer an effective regulator and robustness is lost. Threshold robustness arises when a variable is an active regulator around its threshold, and this function is maintained by the feedback loop that the regulator necessarily takes part in and also is regulated by. In the present study all feedback loops are negative, and our results suggest that threshold robustness is maintained by negative feedback which necessarily exists in the homeostatic state.

Conclusion: Threshold robustness of a variable can be seen as its ability to maintain an active regulation around its threshold in a homeostatic state despite external perturbations. The feedback loop that the system necessarily possesses in this state, ensures that the robust variable is itself regulated and kept close to its threshold. Our results suggest that threshold regulation is a generic phenomenon in feedback-regulated networks with sigmoidal response functions, at least when there is no positive feedback. Threshold robustness in gene regulatory networks illustrates that hidden genetic variation can be explained by systemic properties of the genotype-phenotype map.

Background
Historical perspective

In the early 1970s, Leon Glass, Stuart Kauffman, and René Thomas started their pioneering efforts in exploring the possibility of modelling what was then called "Genetic Control Circuits" (Thomas) and "Biochemical Control Networks" (Glass and Kauffman) by using concepts and ideas from mathematical logic. Combining these ideas with earlier ideas from Monod and others on allostery and cooperativity which suggested sigmoidal rate dependences of key metabolites, Glass and Kauffman [1,2] and Thomas [3] proposed that gene transcription could be modelled by sigmoidal response functions depending on transcription factor concentrations. In the case of several transcription factors acting on a gene, they assumed the effect could be expressed by Boolean combinations of the separate response functions, and proposed a simple framework of ordinary differential equations for modelling of gene regulatory networks based on these principles. Glass and Kauffman observed that the behaviour of these regulatory networks was remarkably insensitive to the steepness of the sigmoids, and suggested to use Heaviside or step function in stead of sigmoids as dose-response functions to simplify the models and their analysis.

From these early attempts, phenomenologic frameworks for the modelling of Gene Regulatory Networks (GRNs) have been developed, based on a few fundamental premises: (i) genes are controlled by transcription factors (TFs) which combine into logical input functions, and these can be described by Boolean logic; (ii) the effect of a transcription factor on the transcription rate of a gene (the response function) can be described by a sigmoidal function of its concentration with a pronounced threshold behaviour (graded response) or by a Heaviside step function (binary response); (iii) this can be modelled in a discrete way in which transcription factors are either absent of present, and proteins are either transcribed or not, or in a continuous way by means of ordinary differential equations; (iv) proteins act as transcription factors, so that networks become closed with feedback loops; (v) posttranscriptional, translational and posttranslational regulation, transport processes, metabolic processes etc. can be phenomenologically encompassed by the sigmoidal or binary response functions.

Recent experimental evidence seem to confirm the validity and usefulness of the basic assumption on which GRNs are resting. A number of studies have shown that gene networks have *cis*-regulatory elements governed by a Boolean-like logic [4-11]. There is also extensive experimental documentation of sigmoidal or binary responses in gene regulation [5,11-16] as well as theoretical justifi-

cations based on classical principles from physical chemistry [17-20].

How common is steep transcription response? Analyses based on classical methods from statistical physical chemistry show that a steep transcription response curve could be due to cooperativity in the transcription factor binding [18-20]. It has been shown that transcriptional and signalling cascades do in fact lead to graded or binary responses [21,22]. There is also extensive evidence that transcription response in single cells is binary (see references in [14]), and that individual cells responds in an on/off way to varying external inputs [13]. Thus, there are good reasons to expect that high gain regulation is quite common in gene regulatory networks.

The method developed in [23-27] to deal with models with steep sigmoidal response functions works for quite general models, also with other nonlinearities in addition to the steep sigmoidal functions. In the course of this work it was discovered that when the responses functions are very steep, equilibrium values for actively regulating variables show a remarkable robustness towards changes in all parameters except the level of the threshold around which the active regulation occurs. We call this phenomenon *threshold robustness*. To be precise, this is a mathematical result valid in the limit when the sigmoid function approaches a step function (Heaviside function), but for continuity reason it is also valid for steep sigmoid functions. But to what extent is it found in models with more empirically sound threshold functions? We have investigated this question by a simulation study of a wide class of 3-dimensional regulatory systems where the regulatory dose-response relationships are varied from a hyperbolic Michaelis-Menten situation to an extremely steep sigmoidal situation.

If conserved when the steepness of the sigmoidal interactions is slackened to realistic values, insensitivity or robustness to functional genetic polymorphism may be a deep generic property of some of the loci in a wide range of regulatory networks. When present, threshold robustness adds significant and characteristic phenomena to the genotype-phenotype map. This implies for example that the functional mutational changes in network which shows threshold robustness will only results in small phenotypic variations in the homeostatic values of the protein products.

Analytical foundation

The above basic assumptions of Gene Regulatory Networks lead to the following generic model for the time course of gene product concentrations:

$$\dot{y}_j = \kappa_j R_j(Z) - \lambda_j Q_j(Z) y_j, \qquad (1)$$

where γ_j is the concentration of gene product number j, $j = 1,...,n$, $Z_j = S(\gamma_j, \theta_j, p_j)$ is a sigmoid or binary function with threshold θ_j and steepness parameter p_j, and γ and Z are the vectors with γ_j and Z_j as components. The functions $R_j \in [0, 1]$ and $Q_j \in (0, 1]$ are regulatory functions, frequently taken to be algebraic equivalents of Boolean functions [25], describing the regulation of production and decay, respectively, while the positive parameters κ_j and λ_j represent the maximal production and decay rates.

Eq. (1) could be justified in at least two ways. It could be considered a model of transcription regulation with the γ_j still representing protein concentrations. This model could be derived from a larger model for protein and mRNA concentrations where transcription of mRNA is regulated by protein concentrations, and the conversion from mRNA to protein is described by linear equations. If all mRNA degradation rates are much larger than all protein degradation rates, we can apply a quasi-stationary hypothesis to the mRNA concentrations, leading to Eq. (1). This procedure can be justified mathematically as well as biologically. A simple example is presented below, and the case $n = 2$ is studied in [28].

Alternatively, taken as a model of gene regulation, Eq. (1) is a generic phenomenological model of protein concentration dynamics, not a mechanistic description of gene regulation. The threshold functions model the aggregated effect of all the processes involved in the real cellular reg-

ulatory networks [29]: transcription, translation, intracellular transport, post-translational modifications, protein-protein interactions, metabolic processes, and signalling cascades. Such drastic simplification is hard to justify theoretically, but models based on the generic Eq. (1) have been applied successfully to many real systems [5,30,31]. Considered in this way, Eq. (1) is a generic, phenomenological framework assumed to catch the essential features of a wide range of regulatory systems, where the regulatory control may be at the level of transcription, mRNA stability, translation, or post-translation, and where the state variables may for example be concentrations of proteins, hormones, mRNA, and intracellular ions [29].

In almost all cases, regulation of the degradation is disregarded, thus we assume $Q_j = 1$. We let S be a Hill function, $Z_j = S(\gamma_j, \theta_j, p) = \gamma_j^p/(\gamma_j^p + \theta_j^p)$, with the property that when $p \to \infty$, then Z_j approaches the Heaviside step function with threshold θ_j, and put all $p_j = p$ (Fig. 1). Of course, in real systems sigmoids most likely do not have the same steepness. Our justification for taking all p_j equal is to investigate systematically whether there is a threshold robustness effect for varying steepnesses. Once this effect is established, one might take one step further to investigate necessary and sufficient conditions on the steepnesses for robustness in more realistic models.

Models encompassed by Eq. (1) have been extensively investigated mathematically, in particular in the special
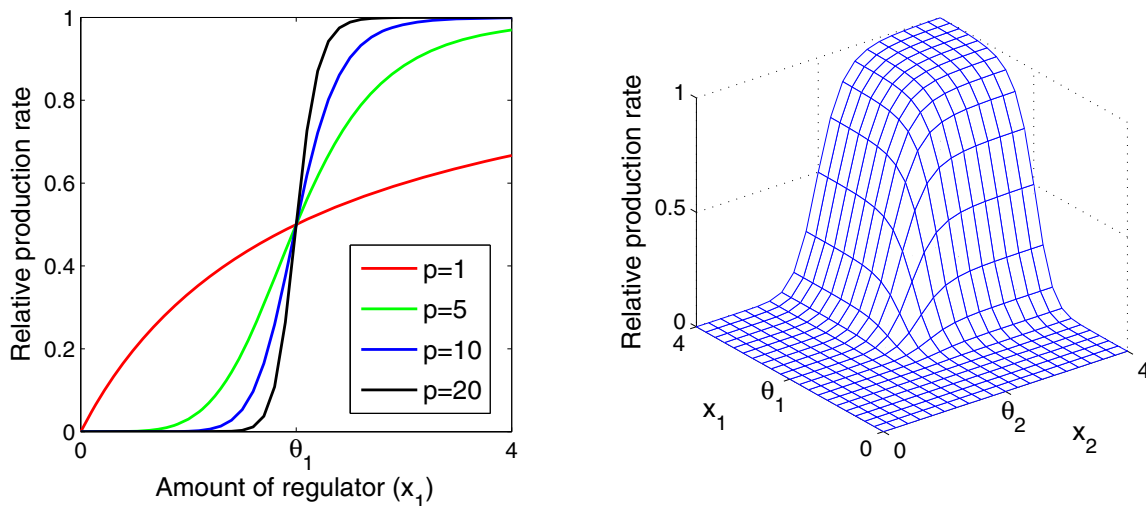
Figure 1
Regulatory functions used in the simulations. (a) The Hill function describes the dose-responce relationship between the amount of a regulator x_1 and the relative production rate of the regulated gene. The threshold parameter $\theta_1 = 2$ is the same for all four curves. (b) Regulatory function for two regulators x_1 and x_2. Two Hill functions are combined by the algebraic equivalent of the Boolean AND function (see Table 1). The parameter values used in this panel are $\theta_1 = \theta_2 = 2$ and $p = 10$.

case of S_j being step functions. (See [32] for an extensive list of references.) An efficient way to analyse models of the generic type Eq. (1) with steep sigmoids is presented in [26,27].

Frequently models encompassed by the generic type Eq. (1) have stationary points lying close to the thresholds of one or several variables when the sigmoids are steep, i.e. when p is large. A stationary point $y^*(p)$ is called a singular stationary point (SSP) for Eq. (1) if at least one of its components approaches its threshold when $p \to \infty$. These components are called singular, the others regular. It has been proved that if a SSP Y^* exists in the step function limit, then there exists a stationary point $y^*(p)$ for sufficiently large p with the property that $y^*(p) \to Y^*$ when $p \to \infty$ [26]. Mathematically, SSPs of Eq. (1) have remarkable generic robustness properties. The key issue is that the singular components of Y^* are locked to their thresholds, while the regular components vary with the other parameters in the model. For $y^*(p)$ this implies that when p is large, the singular components are highly insensitive to changes in all parameters except the thresholds of the singular variables. Biologically this insensitivity means that the expression levels of genes corresponding to singular components will be kept virtually constant despite stochastic or mutational variation in the expression process rates.

An illustrative example is provided by a popular model for a negatively autoregulated gene, which are very common in e.g. *E. coli*

$$
\begin{aligned}
\dot{m} &= r + \alpha(1 - Z) - \gamma m, \\
\dot{y} &= \kappa m - \lambda y,
\end{aligned}
\qquad (2)
$$

where m and y are the concentrations of mRNA and protein, respectively, $Z = S(y, \theta, p)$ is the sigmoidal response function, r is the basal transcription rate, θ is the regulation threshold, p is the steepness parameter, and the remaining four parameters are production and degradation rate constants. Incidentally, asssuming quasistationarity of mRNA concentration as explained above, leads to $\dot{y} = (\kappa/\gamma)[r + \alpha(1 - Z)] - \lambda y$, which is of the generic type Eq. (1).

After elimination of m, the equilibrium equation for y can be solved graphically (Fig. 2). There are three qualitatively different situations to consider: the solution can be close to the threshold and approach the threshold as steepness increases (red line), or the gene can be almost off (green line) or almost constitutively on (magenta line), the differences being accentuated for higher steepness. In the first case the stationary point is a SSP, in the two other cases it is regular. In the singular case we find the solution
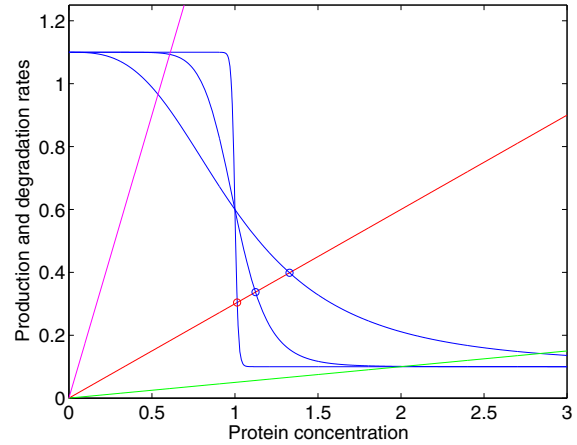
Figure 2
Graphical solution of the equilibrium condition of Eq. (2) for varying steepness of the response function and varying relative degradation rates. The blue curves are graphs of $r + \alpha(1 - Z)$, the straight lines graph the degradation term γy. For intermediate values of γ the equilibrium concentration is close to θ and approaches θ when the response function steepness increases (red line), and there is active regulation. When γ is small, the basal transcription rate is sufficient to balance the degradation, and y^* gets large (green line). When γ is large, degradation is so rapid that the protein concentration never reaches the level where it regulates. Maximal production is necessary to balance the degradation (magenta line).

in the step function limit by putting $y^* = \theta$ and solve for m^*, getting $m^* = \lambda \theta / \kappa$, $Z^* = 1 + r/\alpha - g$, where $g = \gamma \lambda \theta / (\alpha \kappa)$. From the requirement $Z^* \in \langle 0, 1 \rangle$ it follows that this solution exists when $r/\alpha < g < 1 + r/\alpha$. Thus, when g lies in this interval, the protein concentration y^* is very close to the threshold, independent of parameter values, and $y^* = \theta$ in the step function limit.

However, if g is perturbed outside this interval, y^* is no longer locked to the threshold, and robustness is lost. If $p \to \infty$ and $g < r/\alpha$, then $Z^* = 1$ (the magenta case), $m^* = r/\gamma$, $y^* = \kappa r / (\gamma \lambda)$, while if $g > 1 + r/\alpha$, then $Z^* = 0$ (the green case), $m^* = (\alpha + r)/\gamma$, $y^* = \kappa(\alpha + r)/(\gamma \lambda)$. We see that in both these cases, y^* vary with the other parameters.

Model system
For our simulations we chose a particular realisation of Eq. (1) which after a scaling is given by the dimensionless equations

$$
x_j' = \alpha_j R_j(Z_1, Z_2) - \gamma_j x_j, \quad j = 1,...,3,
\qquad (3)
$$

where x_j is the scaled version of γ_j, $\alpha_3 \in \langle 0, 1 \rangle$ and all $\gamma_j \in \langle 0, 1 \rangle$, and $Z_j = x_j^p/(x_j^p + 1)$. The prime denotes differentiation with respect to a scaled time. The Boolean functions R_j are chosen so that the system has a unique stable point in which both x_1 and x_2 are singular (close to threshold at equilibrium). Altogether 14 models satsify this requirement (Fig. 3, see also the Methods section).

The key question is whether the robustness of SSPs is still generic and preserved when the Hill exponents atttain smaller, more realistic values. Considering instances of Eq. (1) as models for gene regulatory networks, we checked this out for a large set of 3-dimensional particular realisations of Eq. (3). We took the set of stable equilibrium values $x_j^*(p)$ as the phenotype and the set $\{ \alpha_j, \gamma_j, \theta_j,$

$p, R_j\}$, $j = 1,...,3$, of parameter values and regulatory functions as the genotype. The equilibrium conditions for Eq. (3) then define the genotype-phenotype map for this system. Our interest is to investigate the robustness properties of the phenotype under mutations, i.e. under perturbations of the genotype. For the 14 models with a unique SSP in the step function limit we investigated the robustness properties of the singular and the regular components of the SSP for a range of parameter perturbations and for varying Hill exponent of the response functions from $p = 1$ (Michaelis-Menten conditions) to $p = 100$, which for all practical purposes is very close to a step function.
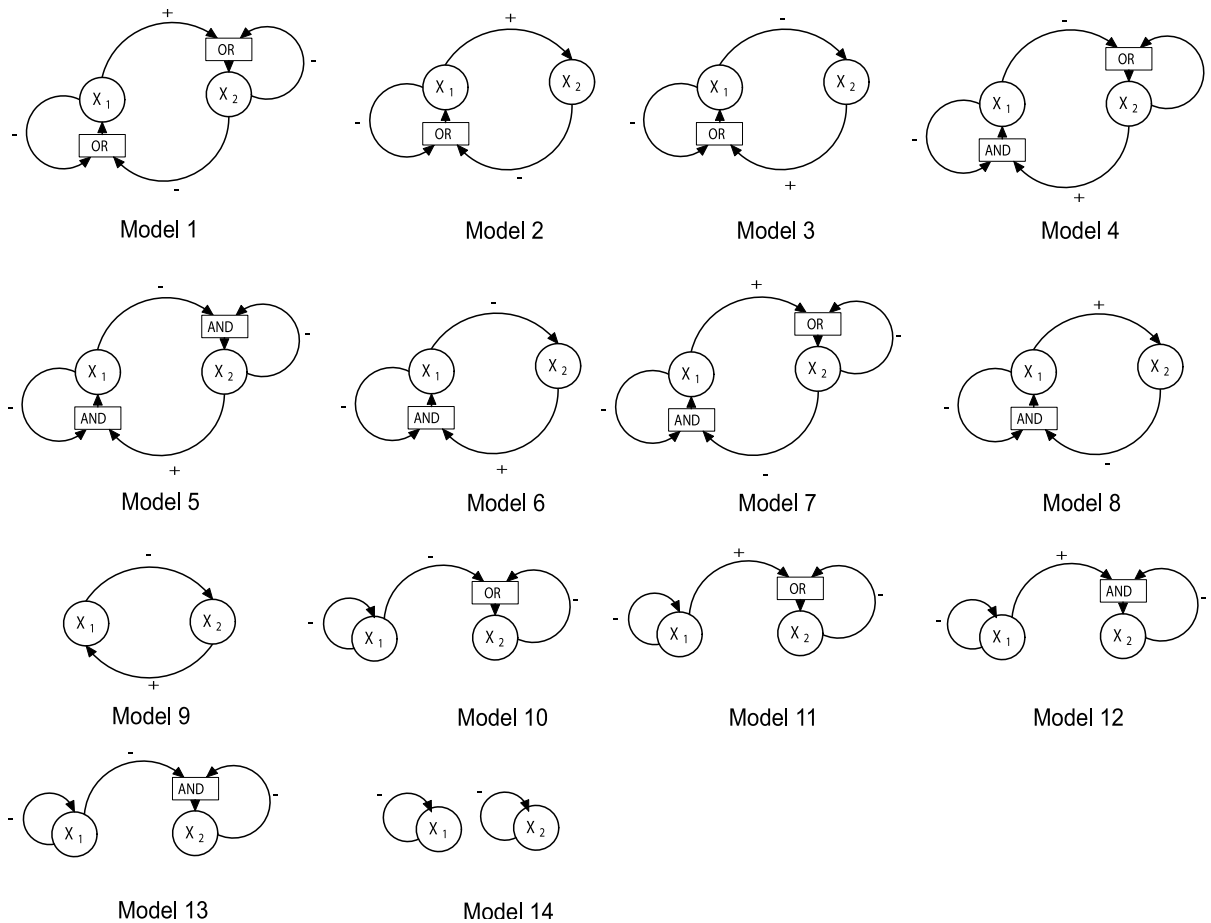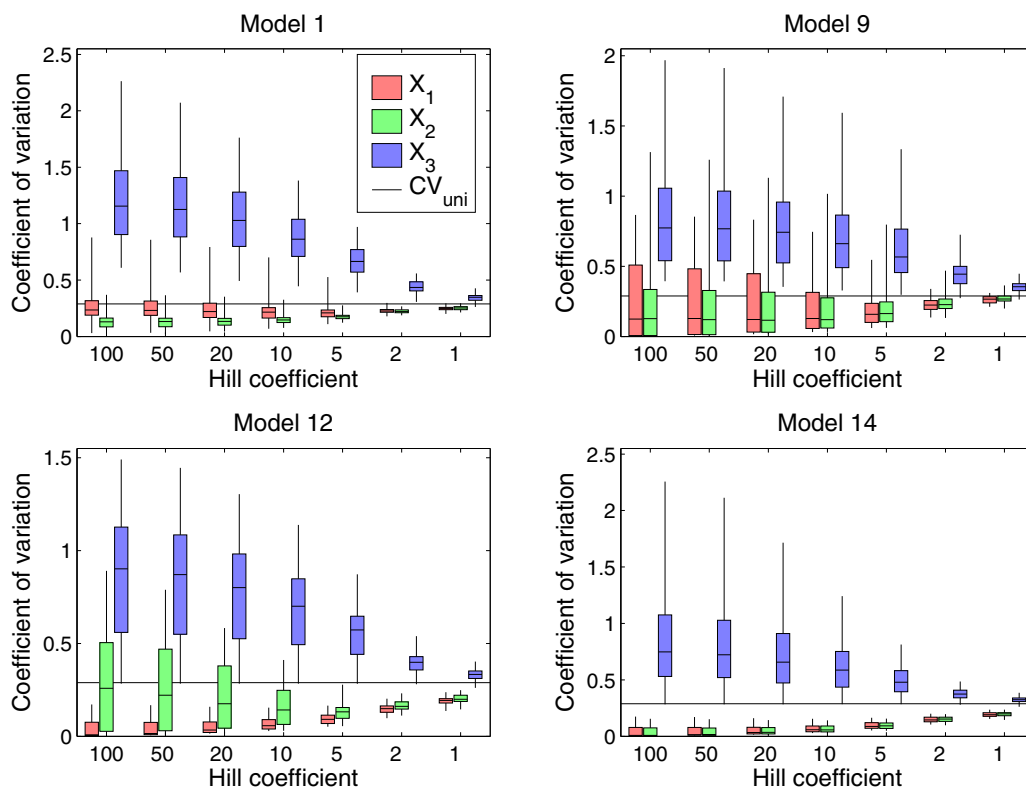

Figure 3
Connectivity diagrams for the 14 network models in the simulation study. Genes 1 and 2 are represented by circles, the downstream gene 3 being omitted for clarity. The sign of an arrow indicates whether the type of regulation is activation (+), in which case the input variable is Z_i, or inhibition (-), in which case the input variable is $1 - Z_i$. When a gene has two regulators, the individual signals are combined with a logic block, represented by a rectangle, merging the two signals into one by the continuous analogue of the Boolean functions AND or OR. (See the Methods section for explanations of the Boolean variables and functions.)

The coefficient of variation *CV* for a distribution is defined as the ratio of the standard deviation to the mean. Being a dimensionless number and scaled by the mean, it is suitable for comparison of the variation of distributions with large differences in mean values. To measure and compare the *CV*s for the equilibrium values x_i^* for each of the 14 models we generated 81 parameter sets, giving a total of 1134 different systems with a unique SSP for which x_1 and x_2 are singular and x_3 regular. For each data set we sampled 50 random pertubations of each production parameter α_j from the uniform distribution $U(\alpha_j/2, 3\alpha_j/2)$ with corresponding coefficient of variation $CV_{uni} \approx 0.288$. We then computed the coefficients of variation CV_j^k, $j = 1, 2, 3$, $k = 1,...,81$ for the steady state levels of all three variables in all the 14 networks separately. Details are described in the Methods section. We use the minimum coefficients of var-

iation as robustness measure in order to be able to compare the robustness of most favourable parameter sets.

Results

Models 1, 9, 12, and 14 represent four different classes among the 14 models: Model 1 represents models with a negative feedback loop between the two genes plus autoregulation, Model 9 has just a pure negative feedback loop and no autoregulation, Model 12 has interaction but no feedback loop between genes 1 and 2, and Model 14 has no interaction at all between genes 1 and 2 (Fig. 3). The comparison (Fig. 4) of the distributions of CV_1^k and CV_2^k over the 81 parameter sets to the distribution of CV_3^k shows that there is a marked difference between the CVs of the singular variables x_1 and x_2 and the downstream, regular variable x_3. This difference is most marked for high Hill coefficient, but is present even under Michaelis-
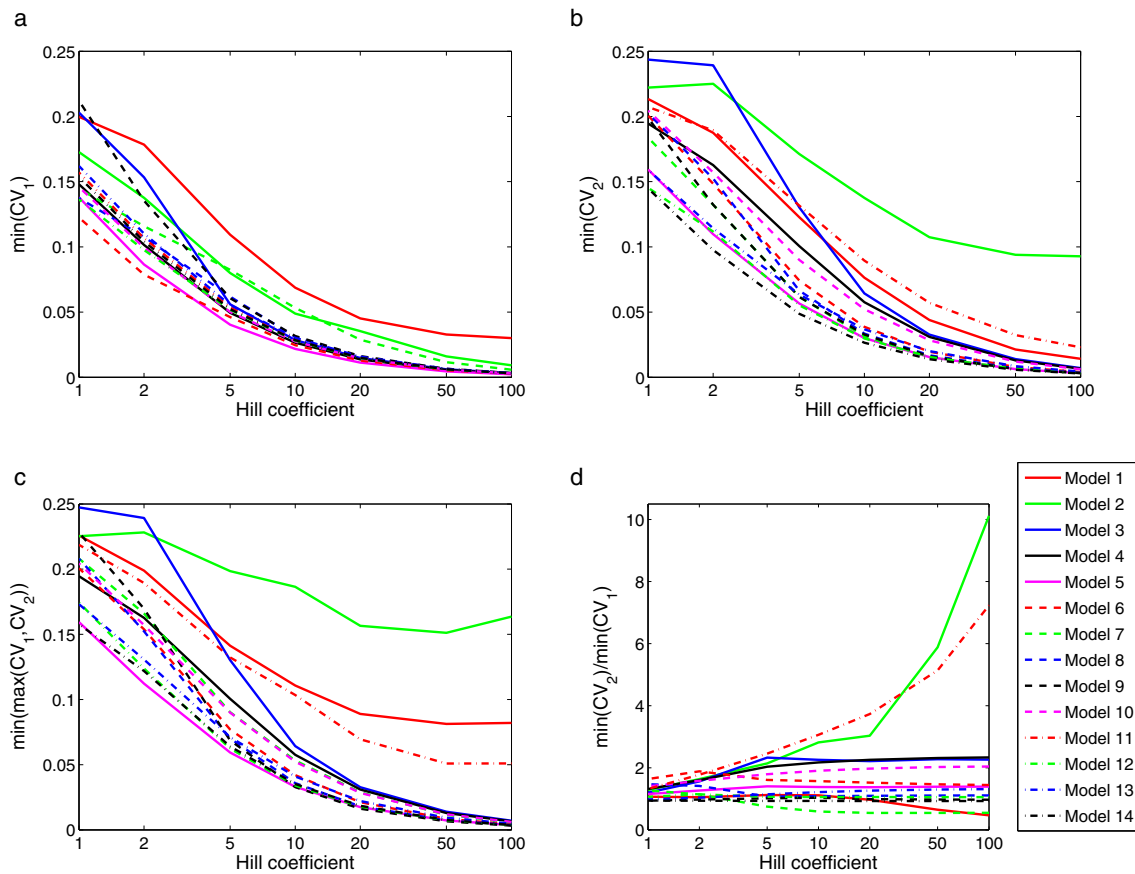

Figure 4
Variation in steady state values for Models 1, 9, 12, and 14. Boxplots show the distributions of the 81 coefficients of variance for all three genes using seven different Hill coefficients ranging from $p = 1$ to $p = 100$. For each Hill coefficients the three plots show from left to right the coefficient of variation for gene 1 (red), 2 (green), and 3 (blue), respectively. The boxes show the quartiles and the median. The black vertical lines extend to the largest observation and the smallest. The long black horisontal line shows the coefficient of variation 0.288 of the perturbed production rates α_j.

Menten conditions. While in almost all cases the variation in x_3^* is larger than CV_{uni}, the variations in x_1^* and x_2^* are considerably smaller for most parameter sets, in Model 14 for all parameter sets. The result for Model 14 is not surprising, as it is well known that negative autoregulation leads to a high degree of robustness [30,33,34].

What is more important is that in almost all cases, all the coefficients of variation for genes 1 and 2 are significantly smaller than those for gene 3. The reduced sensitivity of gene 3 when p decreases towards 1 can be seen as a consequence of the fact that when the sigmoids in the rate equation for x_3 are slackened, a certain variation in the inputs gives smaller variation in the response. However, for all the models except Model 14 there are parameter sets for

which the system is not robust, which shows that robustness is generally only present for a certain range of parameter values. The decisive factor is how easily a perturbation shifts the position of the stable point away from the switching domain. If this happens, the gene's status in the regulatory system is changed: it is no longer an active regulator and its robustness is lost. Accordingly, the gene that was regulated is now either off or constitutively on unless it is still effectively regulated by the other gene.

For all 14 models a decrease in p reduces robustness of both singular variables, defined as the minima of CV_1^k and CV_2^k over all 81 parameter sets (Fig. 5). But even at $p = 1$ there is less variation in the equilibrium values than in the perturbed parameter values. The cases of highest
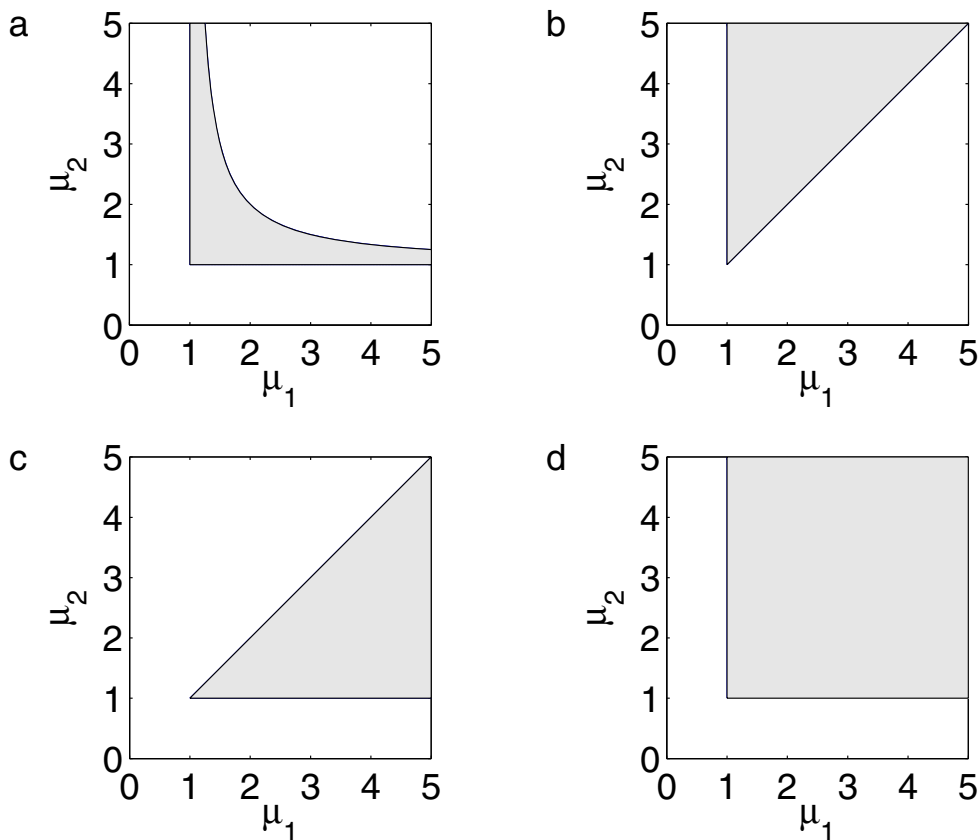

Figure 5
The coefficient of variation as function of the Hill coefficient for the most robust parameter sets for gene 1 and 2 across all 14 models. (a)-(b): Minimum of CV_1^k and CV_2^k, respectively, over all 81 parameter sets. (c): CV_{minmax} as function of the Hill coefficient for each of the 14 models. An explanation for the remarkably high values for Models 1, 2, and 11 for high Hill coefficient values is given in the text. (d) The ratio of minimum of CV_2^k to minimum of CV_1^k as function of the Hill coefficient for each of the 14 models.

robustness for each of the 14 models and all steepness values show large variation among the models (Fig. 5). But for all models, even for the least robust Models 1 and 2, robustness increases with increasing Hill exponent, and is always smaller than CV_{uni}. In Fig. 5c the differences among the models are accentuated for high Hill exponents. For each model the number CV_{minmax} plotted along the vertical axis is computed as follows: for each of the 81 parameter sets first compute $CV_{max}^k = \max\{CV_1^k, CV_2^k\}$, then find $CV_{minmax} = \min\{CV_{max}^k\}$. For each model this procedure selects the parameter set for which both CV_1 and CV_2 are small, giving CV_{minmax} as a measure of the highest robustness when both genes 1 and 2 are taken into account.

Models 1, 2, and 11 have distinctly larger values of CV_{minmax} than the other models. This behaviour can be explained as a consequence of the shape of the parameter space domain Ω_{SSP} in (μ_1, μ_2)-space, where $\mu_j = \alpha_j / \gamma_j$, $j = 1$, 2. This is the parameter domain for which both x_1^* and x_2^* attain threshold values in the step function limit, in other words, the domain in which there is threshold robustness in both variables. For Models 1, 2, and 11 Ω_{SSP} is concave, while for all the other models it is either wedge-shaped or rectangular (Fig. 6). In the concave domains of Models 1, 2, and 11 there is no point giving threshold robustness in both variables for the full parameter perturbation range of 50% up and down (see the

Figure 6
The shaded areas are the robustness domains Ω_{SSP} in the (μ_1, μ_2)-plane for (a) Models 1, 2, 11, (b) Models 3, 13, (c) Model 6, (d) Models 4, 5, 7–10, 12, 14. For parameter values in Ω_{SSP} both x_1^* and x_2^* are singular variables and approach their threshold values in the step function limit. Then they exhibit threshold robustness for all parameter perturbations which leave the perturbed values inside Ω_{SSP}.

Methods section, which also contains the derivation of this result), contrary to all the other models where threshold robustness is obtained for all parameter values sufficiently far from the boundary. With this result in mind it is reasonable to expect a drastically reduced robustness for these three models compared to the rest.

A comparison of CV_1 and CV_2 for all models show a distinct difference in robustness of x_1 and x_2 for Model 2 and Model 11 (Fig. 5d). We can explain this difference by how the character of the stationary point varies over the parameter space (Fig. 7 for Model 2). One can see that x_1 is singular for all $\mu_2 > 1$ independent of μ_1, while the domain in which x_2 is singular is much smaller and with a strongly narrowing band. In this band, all points are close to the boundary, and robustness in x_2 is very easily lost. Accordingly, the probability of having a perturbed point in parameter space in which the singular state is preserved is much less for x_2 than for x_1, just as seen in Fig. 5. For Model 11 the situation is similar.

To further illustrate the difference between the least robust Models 1, 2, and 11 and the rest of the models, we computed for each realisation the number of perturbations N for which the stable point is a SSP with both x_1 and x_2 as
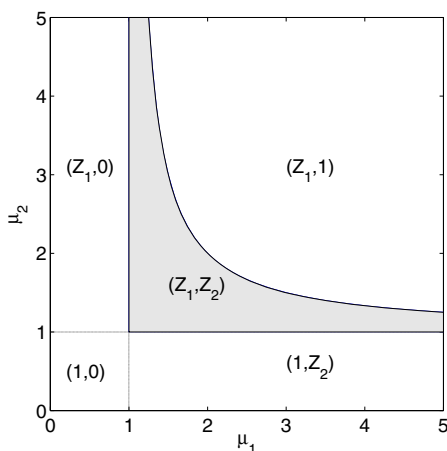
Figure 7
In the case of very steep sigmoids, the (μ_1, μ_2) space of Model 2 is divided into 5 domains, each domain comprising the parameter values giving a particular type of SSP. For example, in the domain denoted $(Z_1, 1)$, x_1 is at its threshold and is singular, thus $Z_1 \neq 0$, 1, while x_2 is above its threshold, thus $Z_2 = 1$. Only in the shaded domain are both variables singular and actively regulating. For steep, but not infinitely steep response functions the relations are approximately true.

active regulators. We consider x_j an active regulator if Z_j^* lies in the interval $\langle 0.05, 0.95 \rangle$. Among all 81 parameter sets the highest observed N was 37, 37, and 35 for Models 1, 2, and 11 respectively. For all other models one can always find a parameter set with $N = 50$, i.e. for which all perturbations render a SSP with a high degree of robustness in both x_1^* and x_2^*.

Discussion
A number of different sources of robustness in cellular function and biochemical networks are discussed in the literature (see e.g. [35,36]). Considered as as a systemic property of a developmental or functional unit in an organism, robustness has been explained as a consequence of both negative and positive feedback and several other network properties [30,37,38], as a consequence of network topology and connectivity [39-41] or of modularity or redundancy of the network [42,43]. Conversely, complexity has been seen as a consequence of selection for robustness rather than the other way round [44,45].

Distinguishing the phenomenon of robustness from homeostasis and stability which concern the system's ability to maintain a stable state, Kitano defines robustness as "a property that allows a system to maintain its functions against external perturbations [37,46]." The function maintained by threshold robustness is the gene's or gene product's ability to act as an active regulator of itself or another gene despite parameter perturbations, which can be seen as consequences of external noise.

A variable x_j exhibits threshold robustness when it is a singular (also called switching) variable of a SSP, in other words, when the stationary value x_j^* approaches the threshold θ_j as the steepness of the associated sigmoid function tends to infinity. In that case the stationary value is locked to the threshold whatever the value of the other parameters (as long they are not perturbed so strongly that x_j is no longer a singular variable in the perturbed system). Then the equilibrium value of the response function Z_j^* is neither close to 0 nor to 1, these cases corresponding to the gene in question being either constitutively off or constitutively on. We illustrated this by our analysis of Model 2 given above. For all parameter perturbations which maintain the parameter values within the shaded area in Fig. 7, both variables are singular and locked to their threshold, and robustness is preserved. Only if the equilibrium point is perturbed outside the shaded area, one or both x_j^* slip away from threshold onto one of the flatter

part of the response curve where corresponding Z_j-value approaches 0 or 1 and is rather insensitive to variations in the x_j-value. Then the *x*-variable is no longer an active regulator, the previously regulated gene now being either constantly off or constantly on, and its robustness is gone. This analysis pertains to the limiting case of very steep responses functions, but is approximately valid when the response is more gradual.

Thus, threshold robustness of a transcription factor or gene product stems from the fact that it is actively *regulating* itself or other genes in the stable equilibrium state. There is a general theorem stating that the system of equations $F(x) = g$, where $x \in R^n$, F is a differentiable function $F : R^n \to R^n$, and $F(x_0) = g_0$, can only have a differentiable solution $x = G(g)$, $x_0 = G(g_0)$, if there is a feedback loop involving all *n* variables [27,47]. When this is applied to a SSP, it follows that *for the SSP to exist there must be a feedback loop among all the singular variables, mediated by the sigmoidal terms Z_j* [27]. Thus, at least for the models investigated in the present paper, if there is a sigmoid-mediated feedback loop among a subset of the variables, and there is a SSP in which these variables are singular, and the sigmoids involved are sufficiently steep, the system will exhibit threshold robustness in this SSP for all the singular variables. Mathematically, threshold robustness is not restricted to any particular type of feedback system. Rather, it is a generic feature of GRNs with steep response functions. Due to the generality of the above-mentioned theorem and the concept and properties of SSPs [25,26], we conjecture that threshold robustness is a general property of singular stationary points. Our findings suggest that this feature is generic in a wider sense, not being dependent on response functions being steep, but it becomes weaker as the system approaches Michaelis-Menten conditions. Also, we may conclude that threshold robustness is a systemic property effectuated by the feedback loop between the sigmoidal interaction terms of the variables at threshold.

The fact that an actively regulating variable necessarily is part of a feedback loop explains how it can maintain its regulation ability. Due to the feedback and because the point is stable, a perturbation that shifts a singular variable away from its threshold, will eventually be counteracted. Thus, a singular variable, i.e. a variable with threshold robustness, is actively regulating and also itself being actively regulated. On the other hand, a feedback loop, even between sigmoidal terms, do not necessarily imply that a SSP exists. It may for certain parameter value combinations but not for others. Thus, it would not be fair to say that threshold robustness is a necessary consequence of feedback. Rather, it is a property of a SSP, and the SSP can only exist if there is a sigmoid-mediated feedback loop among its singular variables. René Thomas has suggested the reasonable conjecture that this must be a negative feedback loop [48]. Our results seem to support this conjecture and suggest that threshold robustness is maintained by negative feedback.

In our numerical simulations, we chose both thresholds fixed at $\theta_j = 1$. However, this was just a matter of convenience and not a model limitation. Without this choice, all parameter variations would still have been expressed by the same two parameters, but with the modification that $\mu_j = \alpha_j/(\gamma_j\theta_j)$. Thus, variation in the thresholds would also result in a variation of μ_j, the only additional effect being a shift in the position of the SSP as long as the perturbed parameters would not fall outside the domain in parameter space where both variables are singular. The system would still exhibit threshold robustness, but now around the shifted threshold values. For the majority of the models this domain covers the major part of the parameter space (c.f. Fig. 6). For this reason threshold robustness might also be called *adaptive robustness*.

Seen as genotype-phenotype maps the studied gene regulatory networks share some interesting features related to the genetic phenomena hidden genetic variation and neutrality. Hidden (or cryptic) genetic variation is defined as "*standing genetic variation that does not contribute to the normal range of phenotypes observed in a population, but that is available to modify a phenotype that arises after environmental change or the introduction of novel alleles* [49]." It is known that negative autoregulation hides variation in the copy numbers of genes [33]. Our results show a more general connection between negative feedback, threshold regulation and hidden genetic variation. In our simulations we fixed a certain amount of parameter variation, corresponding to genetic variation, for a range of different gene regulatory models, and showed that this variation does not result in a corresponding variation in the phenotype. Thus, in a network which exhibits threshold robustness, functional mutations are hidden for phenotypic selection. Our results imply that mutations causing changes in the maximum production rate or the relative decay rate, but keeping the threshold of regulation intact, may have almost no phenotypic signature if the regulatory dose-response relationships are steep enough. Such mutations are neutral in the sense of Wagner's definition: "*A neutral mutation does not change a well defined aspect of a biological system's function in a specific environment and genetic background* [36]." It is implicit in this definition that neutral mutations may aquire a phenotypic signature if the system conditions change. Threshold robustness also offers an explanation of how genotypic variation that is hidden under one condition may be released by for instance a mutation causing a functional change in the regulatory machinery. For instance, the hidden genetic variation

could be released by a change in the regulatory structure beyond the limit inside which robustness for one or several variables is conserved. That would turn singular components into regular components without threshold robustness and susceptible to parameter variations. Thus, a single key mutation in a regulatory structure may release a substantial amount of hidden, potential variation, as illustrated in a simulation study by Bergman and Siegal [50].

Conclusion

This paper presents a simulation study of a class of gene regulatory models in which regulation is modelled with sigmoidal response functions combined by operators mimicking Boolean functions. From a mathematical analysis it is known that when the sigmoidal response functions are very steep, the equilibrium values of the regulating agents are locked to the thresholds, thus are very insensitive to perturbations in all parameters except the threshold levels. This implies that they retain their active regulating power despite parameter perturbations. Our simulations show that this threshold robustness is preserved also for more gentle responses, and is qualitatively present even under Michaelis-Menten conditions. Even though the models investigated are simple, there are reasons to believe that they give a phenomenological description of a large number of different system in which the aggregated effects of a series of transcriptional, translational, and post-translational processes as well as protein-protein interactions and metabolic processes can be described by threshold-dominated response functions.

According to Kitano, robustness is an ability to maintain a function under noise-like perturbations. Threshold robustness is the ability of a protein or transcription factor to maintain an active regulation of a gene in homeostasis under external perturbations. The feedback loop that the system necessarily possesses in the homeostatic state, ensures that the robust members of the loop are themselves regulated and kept close to their threshold values.

Some of the 14 models investigated show a much lower degree of robustness than the rest, a fact that we have explained by a specific analysis of the models in which we compared the shape and extension of the parameter space domain in which the robustness property is preserved under parameter perturbations. For the models with lower degree of robustness, this domain is smaller and concavely shaped such that the perturbed values very easily fall outside the robustness domain. We have also seen that one variable may be considerably less robust than the other, despite all response functions being equally steep, and have offered an explanation of this phenomenon.

Threshold robustness may offer increased insight into genetic phenomena such as maintenance and release of genetic variation in evolution, but a closer investigation of these matters was beyond the scope of the present paper.

Methods
Model equations and regulatory functions
To simplify the simulations our model system

$$\dot{y}_j = \kappa_j R_j(Z_1, Z_2) - \lambda_j y_j, \quad j = 1,...,3, \qquad (4)$$

was scaled to give the scaled equations (3)

$$x_j' = \alpha_j R_j(Z_1, Z_2) - \gamma_j x_j, \quad j = 1,...,3,$$

where $\alpha_3 \in \langle 0, 1 \rangle$ and all $\gamma_j \in \langle 0, 1 \rangle$, and $Z_j = x_j^p/(x_j^p + 1)$. Details are given below.

The regulatory functions R_1 and R_2 were chosen as algebraic equivalents of Boolean functions with two inputs, subject to the requirements $\partial R_j/\partial Z_j \leq 0$, $j = 1, 2$, and $(\partial R_1/\partial Z_2)(\partial R_2/\partial Z_1) \leq 0$ to ensure that there is no positive loop in the system and accordingly no multistationarity, and a globally attracting stationary point. Excluding TRUE and FALSE which imply no regulation, and two more functions that do not fulfil the above mono-stationarity conditions, we were left with the 12 regulatory functions listed in Table 1. The algebraic equivalents are computed by the rules given in [25], X_1 and X_2 are Boolean variables, $\overline{X}_i = $ NOT X_i; Z_1 and Z_2 are the corresponding sigmoidal

Table 1: The 12 Boolean functions used and their algebraic equivalents

	Boolean function	Algebraic equivalent
1	X_1 AND X_2	$Z_1 Z_2$
2	X_1 AND $\overline{X}_2$	$Z_1(1 - Z_2)$
3	$\overline{X}_1$ AND X_2	$(1 - Z_1)Z_2$
4	$\overline{X}_1$ AND $\overline{X}_2$	$(1 - Z_1)(1 - Z_2)$
5	X_1	Z_1
6	X_2	Z_2
7	$\overline{X}_1$	$1 - Z_1$
8	$\overline{X}_2$	$1 - Z_2$
9	X_1 OR X_2	$Z_1 + Z_2 - Z_1 Z_2$
10	$\overline{X}_1$ OR $\overline{X}_2$	$1 - Z_1 Z_2$
11	X_1 OR $\overline{X}_2$	$1 - Z_2 + Z_1 Z_2$
12	$\overline{X}_1$ OR X_2	$1 - Z_1 + Z_1 Z_2$

functions as defined above. The third function $R_3(Z_1, Z_2)$ = Z_1Z_2 in all cases.

Disregarding cases that were equivalent due to the x_1-x_2 symmetry of Eq. (3), we were left with 14 different models (Fig. 3, only x_1 and x_2 are shown). For example, the scaled equation system for Model 6 is

$$
\begin{aligned}
x_1' &= \alpha_1(1 - Z_1)Z_2 - \gamma_1 x_1, \\
x_2' &= \alpha_2(1 - Z_1) - \gamma_2 x_2, \\
x_3' &= \alpha_3 Z_1 Z_2 - \gamma_3 x_3.
\end{aligned} \tag{5}
$$

Scaling
Here we outline the scaling steps to non-dimensionalise the model equations (4) and standardise the parameter ranges.

1. For $j = 1, 2$, introduce $x_j = \gamma_j/\theta_j$ to scale the thresholds for x_j to $\theta_j = 1$.

2. Assume $\lambda_j \in \langle 0, \Lambda \rangle$, $j = 1, 2, 3$, and scale the time t by introducing $\tau = \Lambda t$.

3. Assume $\kappa_3 \in \langle 0, K \rangle$ and introduce $x_3 = (K/\kappa_3)\gamma_3$.

This leads to the scaled equations (3). The new parameters are $\alpha_j = \kappa_j/(\Lambda\theta_j)$, $j = 1, 2$, $\alpha_3 = \kappa_3/K$, and $\gamma_j = \lambda_j/\Lambda$, $j = 1, 2, 3$, with the properties $\alpha_3 \in \langle 0, 1 \rangle$ and all $\gamma_j \in \langle 0, 1 \rangle$. Below we use $\mu_j = \alpha_j/\gamma_j$, $j = 1, 2$.

Sampling
This subsection describes the generation of parameter sets for Eq. (3). The parameters γ_1, γ_2, γ_3 were sampled from independent uniform distributions over $\langle 0, 1 \rangle$. Since we were interested in parameter sets that exhibit threshold regulation in the step-function limit, we did not sample the production rates directly but rather worked our way back from the values of the equilibrium values Z_1^* and Z_2^* in the limit $p \to \infty$. Granted that these lie in $\langle 0, 1 \rangle$, the steady state values of x_1 and x_2 are at their thresholds $\theta_j = 1$. From the steady state conditions of Eq. (3), the maximal production rates for gene 1 and 2 are then given by

$$
\alpha_j = \frac{\gamma_j}{R_j(Z_1^*, Z_2^*)}, \quad j = 1, 2. \tag{6}
$$

For each of the 14 models we generated 81 initial parameter sets $\phi^k = \{\alpha_{1k}, \gamma_{1k}, \alpha_{2k}, \gamma_{2k}, \alpha_{3k}, \gamma_{3k}\}$ by taking the equi-

librium values (Z_1^*, Z_2^*) from the square lattice $\{0.1, 0.2,...,0.9\} \times \{0.1, 0.2,...,0.9\}$.

Having drawn a parameter set we introduced variation in the parameters. For each parameter set ϕ^k we generated a set of 50 perturbed parameter values $\tilde{\phi}_\ell^k$, $\ell = 1,...,50$ by keeping γ_j fixed and sampling α_j of all three genes uniformly in the range $\langle \alpha_j/2, 3\alpha_j/2 \rangle$. This is justified by the fact that the stable state values only depend on the ratios μ_j. For each set of perturbed parameter values we used Matlab's odesolver to find the steady state $x_{j\ell}^{*k}$ of the three gene products in Eq. (3).

Measuring robustness towards pertubation
The robustness properties of the the three gene products in Eq. (3) were evaluated by comparing the coefficient of variation (CV) for parameters to the coefficient of variation for the steady state values. For each $k \in \{1,...,81\}$ the robustness of the steady state values of x_j against perturbations in the production rates α_j is analysed using the coefficient of variation

$$
CV_j^k = \frac{\text{std}(\{x_{j\ell}^{*k}\}_{\ell=1,...,50})}{\text{mean}(\{x_{j\ell}^{*k}\}_{\ell=1,...,50})}. \tag{7}
$$

As the coefficient of variation is invariant under scaling, the parameter CVs will be close to the CV of the uniform distribution $U(0.5, 1.5)$ which is $CV_{\text{uni}} = 1/\sqrt{12} \approx 0.288$. Since we wanted to study how the robustness depends on the steepness of the dose-response function, we carried out this procedure for $p = 1, 2, 5, 10, 20, 50, 100$.

Explaining the differences in robustness among the models
The reason for the reduced robustness of Models 1, 2, and 11 compared to the remaining models can be explained by investigating the shape and size of the robustness domain Ω_{SSP} in the (μ_1, μ_2)-plane for which both x_1 and x_2 are active regulators in the stable point in the step function limit. We find that for Models 1, 2, and 11, Ω_{SSP} is a narrow, concave stripe, while for the remaining models it is a convex domain (Fig. 6). Below we show how to derive this result for Model 1. The analyses for the rest of the models are similar.

From Fig. 3 it follows that the scaled equations of motion for x_1 and x_2 in Model 1 are

$$
\begin{aligned}
x_1' &= \alpha_1[1 - Z_1 Z_2] - \gamma_1 x_1, \\
x_2' &= \alpha_2[1 - Z_2 + Z_1 Z_2] - \gamma_2 x_2.
\end{aligned} \tag{8}
$$

Assume the stationary point is a SSP in which both variables are singular, and that p is very large. Then at equilibrium, $x_1^* \approx \theta_1 = 1$, $x_2^* \approx \theta_2 = 1$, and $0 < Z_i^* < 1$. Solving the equilibrium conditions with respect to μ_j, $j = 1, 2$ under these assumptions we find that $\mu_j > 1$ and

$$\mu_1 = \frac{1}{1-Z_1^* Z_2^*},$$
$$\mu_2 = \frac{1}{1-Z_2^* + Z_1^* Z_2^*}.$$

Upon elimination of Z_1^* this leads to

$$\mu_2 = \frac{\mu_1}{(2-Z_2^*)\mu_1 - 1}. \tag{9}$$

Then Ω_{SSP} is the domain covered by this family of curves when Z_2^* vary between 0 and 1. One easily finds Ω_{SSP} to be the domain between the three curves $\mu_1 = 1$, $\mu_2 = 1$, and $\mu_2 = \mu_1/(\mu_1 - 1)$ (Fig. 6a). As both μ_1 and μ_2 are perturbed 50% up and down in our simulations, there is no point in Ω_{SSP} in which both perturbed values are bound to stay inside Ω_{SSP}.

Authors' contributions
SWO and EP initiated the work, ABG and EP devised the experimental design, ABG ran all the simulations, carried out the data analysis and designed the figures, all authors contributed to and approved the final manuscript.

Acknowledgements

References
1. Glass L, Kauffman SA: **Co-operative components, spatial localization and oscillatory cellular dynamics.** *Journal of Theoretical Biology* 1972, **34:**219-237.
2. Glass L, Kauffman SA: **The logical analysis of continuous, non-linear biochemical control networks.** *Journal of Theoretical Biology* 1973, **39:**103-129.
3. Thomas R: **Boolean Formalization of Genetic-Control Circuits.** *Journal of Theoretical Biology* 1973, **42(3):**563-585.
4. Yuh CH, Bolouri H, Davidson EH: **Genomic *cis*-Regulatory Logic: Experimental and Computational Analysis of a Sea Urchin Gene.** *Science* 1998, **279(5358):**1896-1902.
5. Gardner T, Cantor C, Collins J: **Construction of a genetic toggle switch in *Escherichia coli*.** *Nature* 2000, **403(6767):**339-342.
6. Yuh CH, Bolouri H, Davidson EH: ***cis*-regulatory logic in the endo16 gene: switching from a specification to a differentiation mode of control.** *Development* 2001, **128(5):**617-629.
7. Guet CC, Elowitz MB, Hsing WH, Leibler S: **Combinatorial synthesis of genetic networks.** *Science* 2002, **296(5572):**1466-1470.
8. Setty Y, Mayo AE, Surette MG, Alon U: **Detailed map of a *cis*-regulatory input function.** *PNAS* 2003, **100(13):**7702-7707.
9. Kramer BP, Fischer C, Fussenegger M: **BioLogic gates enable logical transcription control in mammalian cells.** *Biotechnology and Bioengineering* 2004, **87(4):**478-484.
10. Istrail S, Davidson EH: **Logic functions of the genomic *cis*-regulatory code.** *PNAS* 2005, **102(14):**4954-4959.
11. Mayo AE, Setty Y, Shavit S, Zaslaver A, Alon U: **Plasticity of the *cis*-Regulatory Input Function of a Gene.** *PLoS Biology* 2006, **4(4):**e45.
12. Goldbeter A, Koshland DE: **An Amplified Sensitivity Arising from Covalent Modification in Biological Systems.** *PNAS* 1981, **78(11):**6840-6844.
13. Rossi FMV, Kringstein AM, Spicher A, Guicherit OM, Blau HM: **Transcriptional control: Rheostat converted to on/off switch.** *Molecular Cell* 2000, **6(3):**723-728.
14. Biggar SR, Crabtree GR: **Cell signaling can direct either binary or graded transcriptional responses.** *The EMBO Journal* 2001, **20:**3167-3176.
15. Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB: **Gene Regulation at the Single-Cell Level.** *Science* 2005, **307(5717):**1962-1965.
16. Kim J, Bates DG, Postlethwaite I, Ma L, Iglesias PA: **Robustness analysis of biochemical network models.** *IEE Proceedings – Systems Biology* 2006, **153(3):**96-104.
17. Wolf DM, Eeckman FH: **On the Relationship Between Genomic Regulatory Element Organization and Gene Regulatory Dynamics.** *Journal of Theoretical Biology* 1998, **195(2):**167-186.
18. Buchler NE, Gerland U, Hwa T: **Nonlinear protein degradation and the function of genetic circuits.** *PNAS* 2005, **102(27):**9559-9564.
19. Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, Phillips R: **Transcriptional regulation by the numbers: models.** *Current Opinion in Genetics & Development* 2005, **15(2):**116-124.
20. Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, Kuhlman T, Phillips R: **Transcriptional regulation by the numbers: applications.** *Current Opinion in Genetics & Development* 2005, **15(2):**125-135.
21. Thattai M, van Oudenaarden A: **Attenuation of noise in ultrasensitive signaling cascades.** *Biophysical Journal* 2002, **82(6):**2943-2950.
22. Hooshangi S, Thiberge S, Weiss R: **Ultrasensitivity and noise propagation in a synthetic transcriptional cascade.** *PNAS* 2005, **102(10):**3581-3586.
23. Plahte E, Mestl T, Omholt SW: **Global Analysis of Steady Points for Systems of Differential-Equations with Sigmoid Interactions.** *Dynamics and Stability of Systems* 1994, **9(4):**275-291.
24. Mestl T, Plahte E, Omholt S: **A mathematical framework for describing and analyzing gene regulatory networks.** *Journal of Theoretical Biology* 1995, **176(2):**291-300.
25. Plahte E, Mestl T, Omholt SW: **A methodological basis for description and analysis of systems with complex switch-like interactions.** *Journal of Mathematical Biology* 1998, **36(4):**321-348.
26. Plahte E, Kjøglum S: **Analysis and generic properties of gene regulatory networks with graded response functions.** *Physica D: Nonlinear Phenomena* 2005, **201(1–2):**150-176.
27. Veflingstad SR, Plahte E: **Analysis of gene regulatory network models with graded and binary transcriptional responses.** *Biosystems* 2007, **90:**323-339.
28. Cherry J, Adler F: **How to make a biological switch.** *Journal of Theoretical Biology* 2000, **203:**117-133.
29. Brazhnik P, de la Fuente A, Mendes P: **Gene networks: how to put the function in genomics.** *Trends in Biotechnology* 2002, **20(11):**467-472.
30. Becskei A, Serrano L: **Engineering stability in gene networks by autoregulation.** *Nature* 2000, **405(6786):**590-593.
31. Rosenfeld N, Elowitz MB, Alon U: **Negative autoregulation speeds the response times of transcription networks.** *Journal of Molecular Biology* 2002, **323:**785-93.
32. de Jong H: **Modeling and Simulation of Genetic Regulatory Systems: A Literature Review.** *Journal of Computational Biology* 2002, **9:**67-104.
33. Thomas R, D'Ari R: *Biological Feedback* Boca Raton, USA: CRC Press; 1990.

34. Alon U: *An Introduction to Systems Biology. Design Principles of Biological Circuits* Boca Raton, London, New York: Chapman & Hall (CRC); 2007.
35. Stelling J, Sauer U, Szallasi Z, Doyle FJ III, Doyle J: **Robustness of Cellular Functions.** *Cell* 2004, **118(6):**675-685.
36. Wagner A: *Robustness and Evolvability in Living Systems* Princeton: Princeton University Press; 2005.
37. Kitano H: **Biological robustness.** *Nature Reviews Genetics* 2004, **5(11):**826-837.
38. Venkatesh KV, Bhartiya S, Ruhela A: **Multiple feedback loops are key to a robust dynamic performance of tryptophan regulation in *Escherichia coli*.** *FEBS Letters* 2004, **563(1–3):**234-240.
39. Bhartiya S, Chaudhary N, Venkatesh KV, Doyle FJ: **Multiple feedback loop design in the tryptophan regulatory network of *Escherichia coli* suggests a paradigm for robust regulation of processes in series.** *Journal of the Royal Society Interface* 2006, **3(8):**383-391.
40. Ma W, Lai L, Ouyang Q, Tang C: **Robustness and modular design of the *Drosophila* segment polarity network.** *Molecular Systems Biology* 2006, **2:**.
41. Siegal ML, Bergman A: **Waddington's canalization revisited: Developmental stability and evolution.** *PNAS* 2002, **99(16):**10528-10532.
42. von Dassow G, Meir E, Munro E, Odell G: **The segment polarity network is a robust development module.** *Nature* 2000, **406(6792):**188-192.
43. Albert R, Jeong H, Barabasi AL: **Error and attack tolerance of complex networks.** *Nature* 2000, **406(6794):**378-382.
44. Hartwell LH, Hopfield JJ, Leibler S, Murray AW: **From molecular to modular cell biology.** *Nature* 1999, **402(6761):**C47-C52.
45. Lauffenburger DA: **Cell signaling pathways as control modules: Complexity for simplicity?** *PNAS* 2000, **97(10):**5031-5033.
46. Kitano H: **Towards a theory of biological robustness.** *Mol Syst Biol* 2007, **3:**137.
47. Thomas R, Kaufman M: **Multistationarity, the basis of cell differentiation and memory. I. Structural conditions of multistationarity and other nontrivial behavior.** *Chaos* 2001, **11:**170-179.
48. Thomas R: **Circular causality.** *IEE Proceedings Systems Biology* 2006, **153(4):**140-153.
49. Gibson G, Dworkin I: **Uncovering cryptic genetic variation.** *Nat Rev Genet* 2004, **5(9):**681-90.
50. Bergman A, Siegal ML: **Evolutionary capacitance as a general feature of complex gene networks.** *Nature* 2003, **424(6948):**549-552.