



UvA-DARE (Digital Academic Repository)

Inferring *Drosophila* gap gene regulatory network: A parameter sensitivity and perturbation analysis

Fomekong-Nanfack, Y.; Postma, M.; Kaandorp, J.A.

DOI

[10.1186/1752-0509-3-94](https://doi.org/10.1186/1752-0509-3-94)

Publication date

2009

Document Version

Final published version

Published in

BMC Systems Biology

[Link to publication](#)

Citation for published version (APA):

Fomekong-Nanfack, Y., Postma, M., & Kaandorp, J. A. (2009). Inferring *Drosophila* gap gene regulatory network: A parameter sensitivity and perturbation analysis. *BMC Systems Biology*, 3, 94. <https://doi.org/10.1186/1752-0509-3-94>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

Research article

Open Access

Inferring *Drosophila* gap gene regulatory network: a parameter sensitivity and perturbation analysis

Yves Fomekong-Nanfack, Marten Postma and Jaap A Kaandorp*

Address: Section Computational Science, Faculty of Science, University of Amsterdam, Kruislaan 403, 1098 SJ Amsterdam, the Netherlands

Email: Yves Fomekong-Nanfack - Y.FomekongNanfack@uva.nl; Marten Postma - M.Postma@uva.nl; Jaap A Kaandorp* - J.A.Kaandorp@uva.nl

* Corresponding author

Published: 21 September 2009

Received: 22 January 2009

BMC Systems Biology 2009, 3:94 doi:10.1186/1752-0509-3-94

Accepted: 21 September 2009

This article is available from: <http://www.biomedcentral.com/1752-0509/3/94>

© 2009 Fomekong-Nanfack et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Inverse modelling of gene regulatory networks (GRNs) capable of simulating continuous spatio-temporal biological processes requires accurate data and a good description of the system. If quantitative relations between genes cannot be extracted from direct measurements, an efficient method to estimate the unknown parameters is mandatory. A model that has been proposed to simulate spatio-temporal gene expression patterns is the connectionist model. This method describes the quantitative dynamics of a regulatory network in space. The model parameters are estimated by means of model-fitting algorithms. The gene interactions are identified without making any prior assumptions concerning the network connectivity. As a result, the inverse modelling might lead to multiple circuits showing the same quantitative behaviour and it is not possible to identify one optimal circuit. Consequently, it is important to address the quality of the circuits in terms of model robustness.

Results: Here we investigate the sensitivity and robustness of circuits obtained from reverse engineering a model capable of simulating measured gene expression patterns. As a case study we use the early gap gene segmentation mechanism in *Drosophila melanogaster*. We consider the limitations of the connectionist model used to describe GRN Inferred from spatio-temporal gene expression. We address the problem of circuit discrimination, where the selection criterion within the optimization technique is based of the least square minimization on the error between data and simulated results.

Conclusion: Parameter sensitivity analysis allows one to discriminate between circuits having significant parameter and qualitative differences but exhibiting the same quantitative pattern. Furthermore, we show that using a stochastic model derived from a deterministic solution, one can introduce fluctuations within the model to analyze the circuits' robustness. Ultimately, we show that there is a close relation between circuit sensitivity and robustness to fluctuation, and that circuit robustness is rather modular than global. The current study shows that reverse engineering of GRNs should not only focus on estimating parameters by minimizing the difference between observation and simulation but also on other model properties. Our study suggests that multi-objective optimization based on robustness and sensitivity analysis has to be considered.

Background

Gene regulatory networks (GRNs) play a fundamental role in body plan formation and development [1]. In most cases the published gene regulatory networks are based on experimental studies. Furthermore, various studies reveal that network dynamics depend on qualitative aspects (network structure) as well as the quantitative properties [2,3]. An additional option is to study gene regulatory networks in mathematical models and simulation studies. Qualitative and quantitative models may provide insights in the causal relationships between components of a network as well as the mechanisms behind the network dynamics [2]. Qualitative models mainly focus on the inference of the wiring diagram and the understanding of the connection between the dynamics of the systems and the network structure [4,5]. Understanding how the gene regulation leads to pattern formation requires the usage of quantitative models that allows inferring the network structure, but also simulating the spatio-temporal gene expression dynamics [6]. Many approaches have been used to model gene regulatory networks (see for a review [7]). In modelling processes from developmental biology especially the class of quantitative spatio-temporal, models of gene regulation is relevant. These type of models can potentially be linked with three-dimensional biomechanical models of morphogenesis and provide new insights into developmental biology.

Although there is a large variety of formalisms used to describe quantitative model of GRN, the most commonly used is based on differential equations. This deterministic modelling framework allows describing spatio-temporal continuous system, where protein products are continuously produced through the balance of activation or repression. Models can be used in reverse engineering studies to infer the network structure and dynamics of the system. Quantitative spatio-temporal models of gene regulation are typically characterized by a large set of unknown parameters. Without a suitable method for reverse engineering the network and estimating the unknown parameter values these models have no practical use. Because of incomplete data, imprecise information concerning the interaction mechanisms and the large set of unknown parameter values; reverse engineering of the gene regulatory network is by far a straightforward task [8].

Reverse engineering of GRN capable of simulating spatio-temporal gene expression mainly consists in revealing the GRN structure that leads to the observed pattern. Optimization is therefore important [9,10] and it is used to infer a gene network whether it is transcriptional or protein interaction. A system identification approach is used to select a model structure and by means of parameter estimation, the network topology is estimated from experi-

mental data. The optimization problem consists in minimizing the difference between simulated expression profiles and available data in order to estimate the best circuit that predicts with very good accuracy the spatio-temporal gene expression. Based on the resulting parameter-set, the network diagram is extracted and one tries to establish causal-relation of the dynamic mechanism that governs the pattern formation, using further analysis. The prediction is clearly linked to the quality and amount of the data; even with sufficient data, it is not guaranteed that the inferred network corresponds to the appropriate gene network that leads to the observed pattern [11,12]. From a theoretical point of view, this question is a matter of the structural identifiability of the model [13]. Given a data set, is it possible to uniquely infer the network? Although there exist some theoretical methods [12] to investigate the identifiability before inferring the network, when confronted with a complex mechanism characterized by a multi-dimensional parameters space, the feasibility is still analytically complex. In many situations, one is confronted with an overfitting problem that leads to parameter-set with very different quantitative values and even worse, yielding different network topologies having similar good realistic patterns.

To draw conclusions, one has to differentiate between circuits that are more likely to be biologically realistic.

It was shown that for certain experimental data that it is not possible to confirm whether the inferred model is really valid [14]. As long as the model is not contradicted by the given data, it is necessary to extend the validation test further. Especially when the model does not only predict a unique network, parameter-set discrimination can be addressed using diverse models validation approaches. The quality of the circuits can for example be quantified by measuring the parameter reliability and sensitivity, the uniqueness of the predicted network, the model robustness and the predictability of the model. In this paper we focus on the model robustness against perturbation. In another paper (Fomekong-Nanfack et al., in press.) we focus on the other aspects that can be used to quantify the quality of the circuits: methods for exploring the solution space, parameter correlation and asymptotical stability analysis.

By analyzing the robustness of the inferred network, we can test the quality of the circuits. The term robustness has various meanings, but in the current study, robustness is addressed as the ability of a system to maintain its mechanical and dynamical behaviour under perturbation. From a biological view, robustness is related to stability, homeostasis, canalization, redundancy, and plasticity [15-18], and can also be applied to dynamical process in development [19]. In simulations of dynamical develop-

mental processes such as pattern formation, one should also expect robust behaviour within a certain extent in the model [20]. In the current paper, model robustness is addressed in two different perspectives. First, we investigate the quantitative robustness of the model towards internal fluctuations in expression level. It is known that presence of noise in gene regulation can lead to phenotype variation [21,22]. There are some studies on the robustness of GRNs under the influence of molecular fluctuation [23,24] and show the importance of noise and stochastic events [25]. In most cases, deterministic models are used to infer GRNs from noisy data. It has been shown in several studies that the robustness to noise mainly depends on the network structure instead of the parameter setting [26-28]. Therefore, one way to discriminate between circuits having different gene network but exhibiting the same pattern is to analyze their behaviour under noisy conditions. Model-solutions showing more robustness or stability can be considered for further analysis. Second through simple parameter perturbation, we investigate how model-circuits behave distinctly. This analysis allows us to identify the parameters that have the most significant influence on the model and to distinguish circuits that are less sensitive to overall perturbation. The combination of these two analyses allows one to discriminate between sets of circuits that are more robust, although they are obtained from the same model description and quantitative data.

Reverse-Engineering the gap gene network

In the current paper, the case study early development along the anteroposterior (A-P) of gap gene network of the *Drosophila melanogaster* is considered. In the gene regulatory network responsible for segmentation of the *Drosophila melanogaster* embryo [29-32] for most, if not all genes involved, experimental and bioinformatics studies are available. These studies not only give information about potential interactions between the different genes, but also provide spatio-temporal information about the gene expression patterns in the embryo. The segmentation genes leading to pattern formation, which occurs during the first stages of development after fertilization is controlled by a network of genetic interactions, with a cascade like modularity [33]. At the early stage, the maternal mRNA, located at the extremes of the egg will define the anterior-posterior axis of the embryo. Following fertilization, the process begins with the diffusion of maternal morphogen factors: Bicoid (Bcd), Nanos (Nos), Caudal (Cad) and maternal Hunchback (Hb), and the activation of the torso receptor on the poles. At cycle 13 (after 13 nuclear divisions), the gradient of these maternal gene products in the embryo promotes the transcription of the first zygotic genes during cycle 14: zygotic *hb*, *giant* (*gt*), *Krüppel* (*Kr*), *knirps* (*kni*) and *tailless* (*tll*); each forming expression domains in specific regions of the embryo

[34]. This set of genes; the so-called gap genes-are defined in broad domains along the anterior-posterior axis. Their transcription factors regulate the expression of another group of zygotic genes that comprises seven pair-rule genes, which form a pattern of narrow stripes along the anterior-posterior axis. Finally, the pair rule genes regulate the formation of fourteen stripes by segment polarity genes. For review about the detailed mechanism, see [35-37].

For modelling the segmentation mechanism of the early *Drosophila melanogaster* embryo, two main formalisms have been proposed: logical formalism (tackling qualitative aspects) proposed by Sánchez and Thieffry [38], and continuous models proposed by Mjolsness et al. [39] used to obtain quantitative dynamics of a system. Following this formalism, Reinitz and co-workers formulated the problem as an inverse problem [40]. Given a complete mathematical model and sufficient accurate quantitative data, the parameters in the model can be estimated by optimization techniques, i.e., by fitting the model to the data. Except for box constraints, only little experimental information is used to constrain the parameter values in the model. This inverse modelling formulation of the problem leads to different gene circuits describing different aspect of the segmentation gene mechanism [41-44].

Using the gene circuit method proposed by Reinitz *et al* [40], Jaeger et al. [42,45] have inferred a network by means of reverse engineering that can reproduce the measured spatial and temporal gene expression patterns. The gap gene model involves seven different genes, *bcd*, *cad*, *hb*, *gt*, *Kr*, *kni* and *tll*. The experimental data used to fit the model were obtained from the FlyEx database, where an extensive amount of accurate quantified spatio-temporal expression data for all genes is stored [46,47]. A connectionist description was used to model the gene regulatory network. The number of parameters in this framework mounts up to 66 different unknowns for a network of six genes. To estimate these parameters the model was fitted to detailed spatio-temporal data using parallel simulated annealing (PLSA) [48-50], which is a technique that performs a global parameter search. Ten different gene circuits were obtained using this reverse engineering approach. Compared to the variance in the experimental data, all solutions fitted the dataset spatially and temporally accurately. Furthermore, the model reproduced the experimentally observed dynamic shift of the expression profiles and it was shown that diffusion is not the cause of the shift. Compared with experimental knowledge, for most interactions in the gene regulatory network the sign, i.e. activation, repression or no-interaction was reasonably well reproduced. Analysis of the 10 circuits suggested asymmetric repression between the gap genes, which was hypothesized to be the cause of the observed domain

shift. This result has also been obtained by Perkins et al. [43] where parameter estimation was performed using a three-step strategy. Following [42], Fomekong-Nanfack et al. [51] employed a different optimization technique to find parameter sets that accurately fitted the experimental dataset and lead to the similar networks as proposed by Jaeger et al. [45]. In search for a computational more efficient method, they developed a hybrid optimization algorithm [51]. Typically this approach first performs a global search based on evolutionary strategy followed by a local search. With this method, another 91 circuits were obtained that fitted the experimental dataset accurately. Recently, Ashyraliyev et al. [52] showed that the circuits obtained by Fomekong-Nanfack et al. can be further improved using Levenberg-Maquardt.

The gene circuit method previously used [40,42,43,51] does not make any assumption about the network structure. All gene-to-gene interactions are assumed to be plausible and minimizing the difference between observation and simulation drives the inference. The criterion for selecting acceptable circuits was based on a low root mean square error (RMS) and visual inspection of the simulated gene expression profiles (no major pattern defects noticeable). In the current case, a large set of circuits can be used to simulate the A-P patterning of the *Drosophila melanogaster*. Using a local sensitivity analysis, determinability of some parameters were studied in [52], where it was suggested that one could confirm on the nature of some biological interactions for parameters that were shown to be identifiable. Using a large set of parameters as base value, it is possible to examine the robustness of the model. Zak et al. [23] showed how input perturbations and stochastic gene expression influence the identifiability of a specific regulatory network given gene expression profiles and prior structural knowledge. Using as starting point the parameters-set obtained in [42,51] as base value, we discuss how does input perturbation and stochastic simulation allow a model-based robustness analysis of a model that leads to multiple circuits.

Results

Gap gene circuits

The analysis is based on 101 circuits each characterized by 66 parameters obtained from [42,51] using different optimization techniques. All circuits with a RMS value smaller than 12 (expression level measured in units of fluorescence level) were labeled as good fits and were selected for further analysis. In Fig. 1, the simulated expression patterns at gastrulation time are shown. The profiles (light grey) and the average profile (colour solid line) for each gene are shown together with the real data (dashed line). All circuits reproduce the gene expression patterns without any major defects (mainly amplitude variation).

Although the profiles are well defined, the circuit-parameters show a rather different picture. In Fig. 2, the distribution of the 66 parameters is shown. The grey line represents the different 101 values of each circuits and the blue line their average. From this figure, we see that the distributions vary from parameter to parameter. None of the non-regulatory parameters such as the production rate, the diffusion coefficient or the decay time constant seems to be very consistent from circuit to circuit. In many cases, the promoter rate and the diffusion hit the upper boundary. Many parameters show a strong tendency to cluster at a specific value or location (specifically those around zero such as W_{gt}^{Hb} , W_{hb}^{Kr} or W_{gt}^{Kni}). In some cases, we see a very broad distribution around the mean, especially for the parameters describing strong repression. In other cases regulatory parameters show both positive and negative interactions (Bcd_{ill} , W_{ill}^{Hb} , Bcd_{kmi}), representing different network topologies.

Robustness towards fluctuations

Gene transcription, degradation and diffusion inherently are stochastic processes, which lead to fluctuations in gene expression levels. If expression levels are high, these fluctuations generally do not affect the time evolution of the system. In these cases the deterministic model and stochastic model will yield a similar result. However if the system is non-linear and the fluctuations occur at an early time in development, where levels are still low, fluctuations may lead to different patterns. From a biological point of view this may not be favorable because gene pattern defects can lead to aberrations in development, hence development should be robust towards modest transient disturbances [53,54]. Here we investigate if the circuits obtained from the optimization, which are based on a deterministic model, are robust towards fluctuations.

Analysis of fluctuations in the bicoid gradient [55] suggests that the expression level of bicoid is about five fold higher than the fluorescence unit. In our analysis we assume that the expression level of all genes is five fold higher than the fluorescence unit. We have used the Gillespie algorithm [56] to introduce fluctuations in the gene expression levels.

For each circuit we performed 100 stochastic simulations (see methods) and analyzed the quality of the final expression pattern at $T = 68.1$ min. In each run, by comparing the deterministic model with the stochastic model using a RMS cut-off criterion, we counted how many of domains *cad*, *hb*-anterior (*hba*), *hb*-posterior (*hbp*), *Kr*, *gt*-anterior (*gta*), *gt*-posterior (*gtp*), *kni* and *tll* were considered to be correct. For each circuit we also calculated an

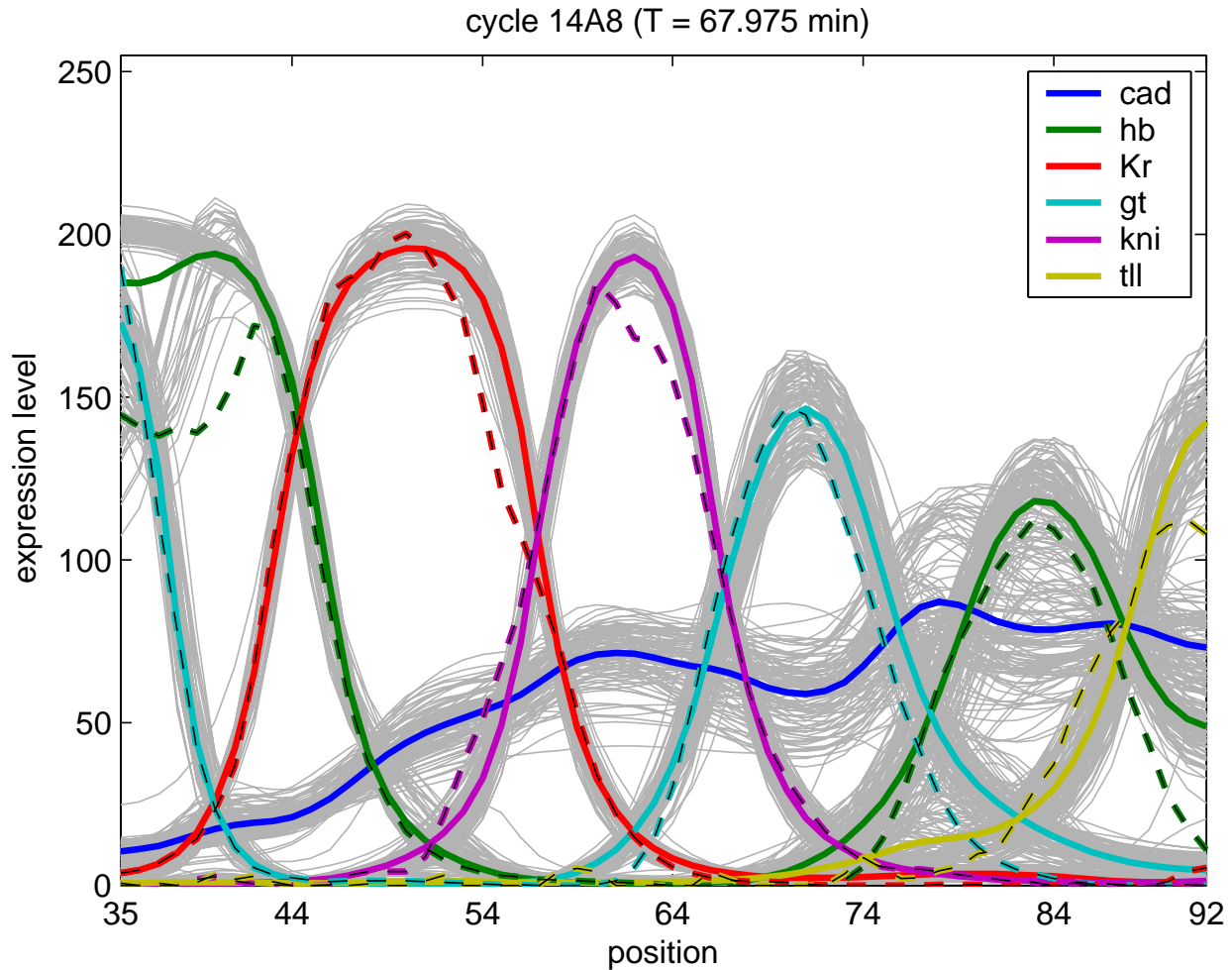


Figure 1

Simulated expression profiles of the 101 gap gene circuits at the final time points. The x-axis corresponds to 35-92% of the A-P position and the y-axis describes the expression level in fluorescence units. The experimentally measured expression profiles are plotted with coloured dashed lines. They represent average quantitative gene expression data obtained from fluorescent immunohistochemistry stained *Drosophila* blastoderm embryos [95] followed by successive image-processing operations [46,96]. The simulated profiles obtained from different circuits for each gene are shown in light grey and the average profiles at a specific time point are plotted using coloured solid lines.

overall score by counting the runs where all expression domains formed correctly [see Tab.1 in Additional file 1, and Additional files 2, 3 w]. On average the domains Cad (99%) and Hb-anterior (93%) show the best scores. The domains of *kni* (86%), *Kr* (80%) and *gt*-posterior (81%) have an intermediate score. The domains *gt*-anterior (66%), *hb*-posterior (70%) and particularly *tll* (47%) have a low score. Most circuits have a very low overall score; only the top 25% of the circuits have an overall score higher than 35%. Most of the circuits have a score lower than 20%.

If all domains except one develop correctly the overall score is determined by that domain score, however if two or more independent domains have scores lower than 100%, the overall score will be lower than the individual domain scores. For example, if all domains would have a score of 99% and are independent the overall score would be $0.99^8 * 100\% = 92\%$. However we also observe interactions between domains, which leads to a higher overall score than would be expected if they behave independently. In these cases, if one domain yields a low score for a particular run, it is likely that another domain also yields a lower score for the same run. This is typically observed

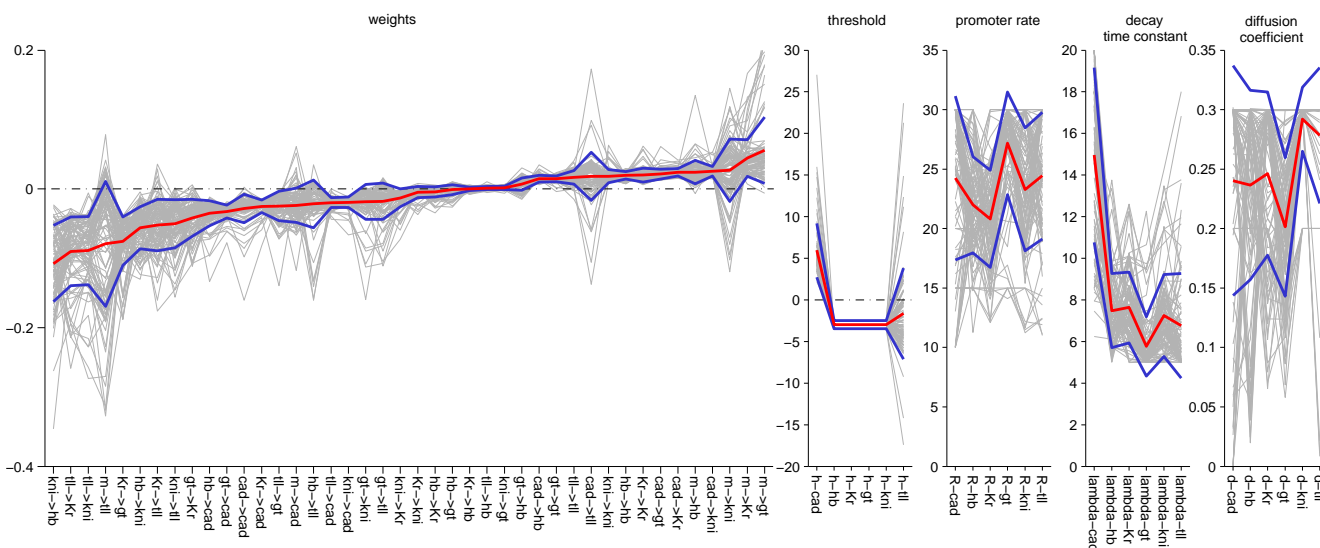


Figure 2
Distribution of the 66 parameters obtained from the 101 gene circuit. The solid red line represents the average value, the solid blue lines represent the standard deviations and the light grey lines are the different individual circuit parameters. Parameters are sorted according to their mean value. Most parameters show a strong tendency to cluster around a particular value, defining the type of interactions. However, some of them have a very broad distribution around their mean and in a few cases, they show all different types of interactions i.e. activation, repression or no interaction.

between adjacent domains. Notable examples are *Kr* with anterior *gt*, anterior *hb* with anterior *gt*, *kni* with posterior *gt*, and posterior *gt* with posterior *hb* and *tll*.

In Fig. 3 the patterns of eight different circuits are shown. In these graphs the dashed lines represent the final profiles obtained from the deterministic model and the solid lines the average profiles obtained from the individual stochastic simulations calculated from the individual stochastic runs, which are shown in grey, [see movies in Additional file 4,5,6,7 where stochastic simulation of some circuits are shown]. In Fig. 4 the time-evolution of the expression level of different gene combinations at a particular nucleus position are shown for the same circuits as in Fig. 3. For comparison, the deterministic model is plotted using a solid red line and all stochastic runs (100) using light coloured lines. At the end of these lines a dot is shown, which represents the final concentration at $T = 68.1 \text{ min}$. The end points were used to extract two clusters (with k-means clustering), which are represented by the blue and green colour.

In Fig. 3A, B the final pattern formed during the stochastic simulations are shown for circuit 48 and 11. Circuit 48 has an overall pattern score of 62% and nr 11 has an overall score of 34%. In circuit nr 11 all domains except *Kr* (44%) and anterior *gt* (35%) score 100%. *Kr* and anterior *gt* show a strong interaction, if anterior *gt* disappears then *Kr* expands into the anterior *gt* domain. In circuit 48 this

interaction between *Kr* and *Gt* at the anterior of the embryo is not present. In this circuit especially *tll* is not well defined. Fig. 4A-B shows the trajectory of *Kr* and *gt* at nucleus 35, for circuit 48 in almost all runs *gt* and *Kr* develop correctly, however in circuit 11 there are two possible outcomes of the stochastic run. Here *gt* and *Kr* can develop correctly, however *gt* may also disappear and completely repressed by *Kr*. In this circuit there are two pathways for the system to evolve, which may lead to two stable points or a single stable point but two pathways. Jaeger et al. [45] suggested that non-overlapping gap gene are mutually exclusive and have mutual repression. This result was confirmed for *gt* and *Kr* by Ashyraliyev et al. [52]. Circuit 48 and 11 both show this strong mutual repression but we believe that the bad stochastic score of anterior *gt* and *Kr* might be caused by the weak repression of *gt* by *Tll*.

In Fig. 3C, D the final pattern formed during the stochastic simulations are shown for circuit 2 and 41. Both circuits have a very low overall score of 12% and 0% respectively. In circuit 2, a very low score for *hb-p*, *gt-p* and *tll* mainly causes this and for circuit 41 all domains except *cad* are not well defined. In Fig. 4C, D the trajectories at nucleus 35 of *hb* and *gt* are shown. For circuit 2 the pathway is well defined, in most runs both *gt* and *hb* evolve similar to the deterministic model. However in circuit 41 the pathways are not well defined at all and show variability. In circuit 2, we see that *Hb* represses *gt* while *Gt* weakly activates *hb*,

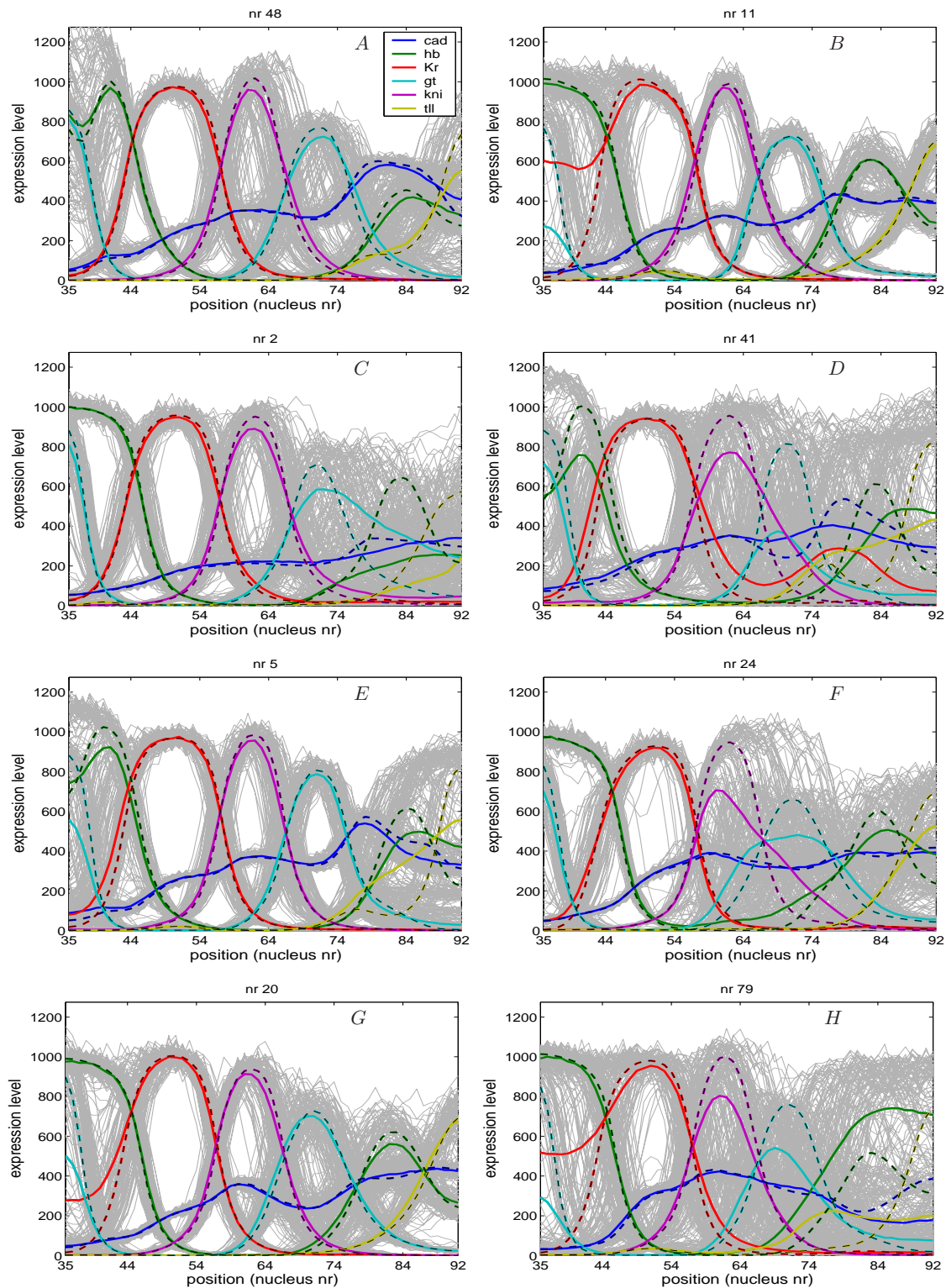


Figure 3
Stochastic simulation of eight different circuits. In these graphs the dashed lines represent the final profiles obtained from the deterministic model and the solid lines the average profiles obtained from the stochastic simulations calculated from the 100 individual stochastic runs, which are shown in grey.

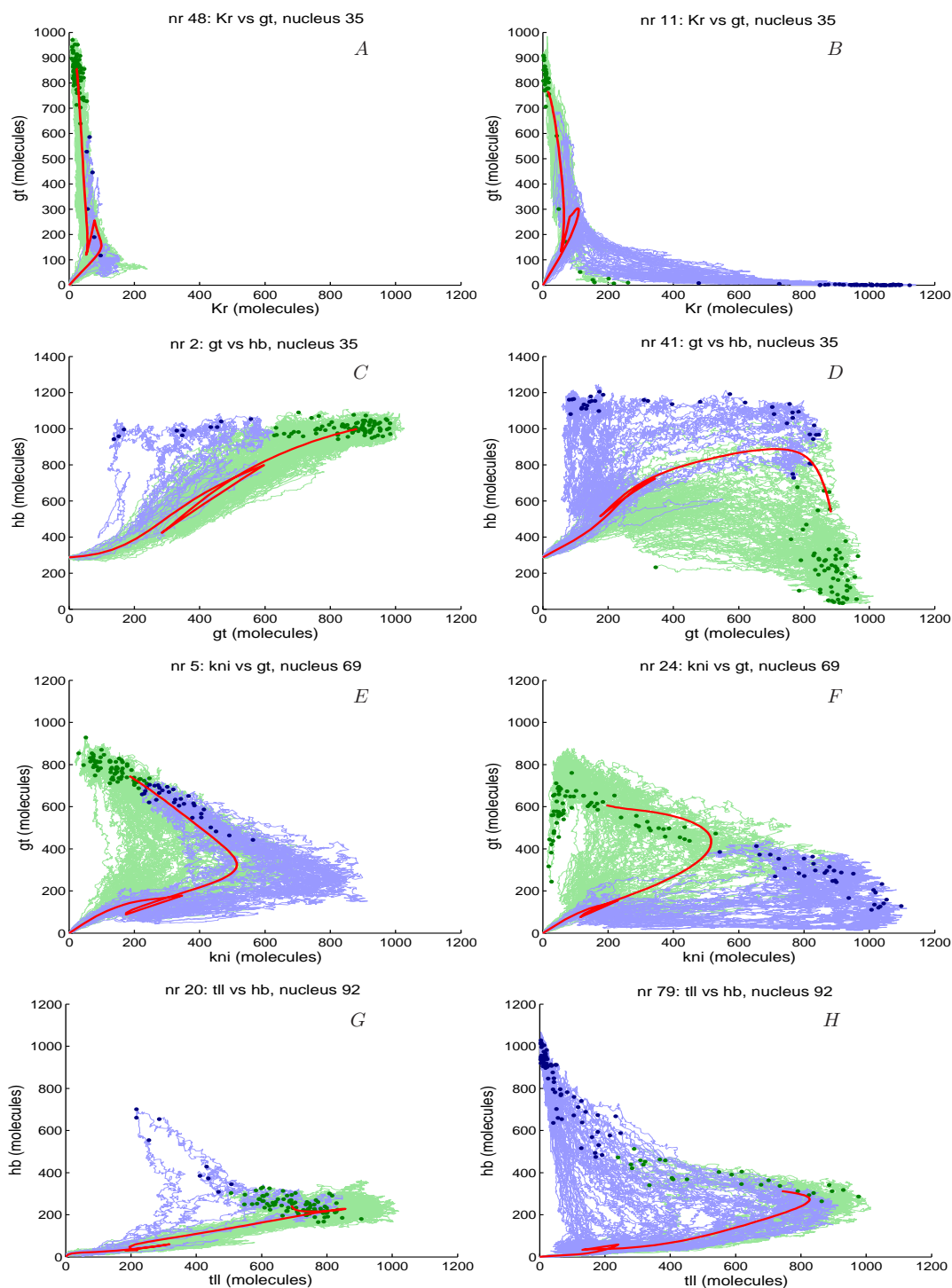


Figure 4
Phase portrait of stochastic simulation at a particular nuclear position. The time-evolution of the expression level of different gene combinations at a particular nuclear position is shown for the same circuits as in Fig.1. For comparison, the deterministic model is plotted using a solid red line and all stochastic runs (100) with light coloured lines. At the end of these lines a dot is plotted, which represents the final concentrations at T = 68.1 min. These end points were used to extract two clusters (with k-means clustering), which are represented by blue and green colours. The left panels show circuits with final points and pathways very close to deterministic model and the right panels show less well defined circuits.

and inversely for circuit 41. It is suggested that overlapping gap genes do not activate each other [45,52]. The poor robustness of these 2 circuits is therefore a consequence of the wrong connectivity of these 2 interactions.

In Fig. 3E, F the final pattern formed during the stochastic simulations are shown for circuit 5 and 24. Circuit 5 has a very low score of 10% and circuit 24 a score of 40%. Although circuit 5 has a low overall score both the *kni* and posterior *gt* domain score 100%, circuit 24 has a lower score for these domains. In Fig. 4E, F the trajectories are shown for *kni* and *gt* at nucleus 69. In circuit 5 the final points of the stochastic runs are very close to the final point of the deterministic run. Although the final points are well defined, the pathway, which is a shift of both domains, to these points is quite variable. In circuit 24 the same phenomenon is observed, only here *Kni* in some cases completely suppresses *gt*. Both circuits show *gt* activation by *Kni* (weak), which seems to be a wrong interaction [45,52]. Circuit 5 shows repression of *hb* by *Gt* and weak activation of *gt* by *Hb* while circuit 24 shows the inverse mechanism. Based on the identifiability analysis obtained by Ashyraliyev et al. [52], it is suggested that *Gt* does not repress *hb* and *Hb* does not regulate *gt*. However, the determinability of these parameters has a very poor confidence and qualitative conclusion on these interactions is still ambiguous.

In Fig. 3G, H the final pattern formed during the stochastic simulations are shown for circuit 20 and 79. Circuit 20 has an overall score of 50% and circuit 79 has a very low score of 2%. In circuit 20 all domains except *Kr* and anterior *gt* are well defined. In circuit 79 all domains except *cad* are not well defined. In Fig. 4G, H the trajectory of *hb* and *tll* at nucleus 92 are shown. In circuit 20 most final points are very close to the deterministic model, however in circuit 79 almost all points are far away from the deterministic model. In this circuit *tll* domain is repressed by *Hb* and completely disappears and the *hb* domain continues to grow. In some runs in circuit 20 (blue trajectories) we see the same tendency of continuous *tll* repression combined with *hb* increase. Circuit 79 shows very inconsistent regulatory mechanism with respect to available literature [31,32,57,58]. *gt* and *hb* show mutual activation, and *Kni* activates *kr*. These interactions seem to be the wrong regulatory mechanism, leading to stochastic instability and a very low pattern score.

Correlation between robustness and parameters

In Fig. 5 the correlation between the eight different expression domains scores and all parameters are shown. The correlation pattern reveals that circuits with higher promoter rates, diffusion coefficients and higher degradation rates are more robust towards fluctuations. This suggests that higher rates tend to increase robustness of the circuits.

Except for W_{cad}^{Bcd} all maternal inputs correlate negatively with pattern robustness. This suggests that strong maternal inputs tend to reduce the robustness of the circuit. All negative inputs on *cad* show a positive correlation, which suggests that strong *Cad* input weights tend to reduce robustness. Furthermore, the inputs of *Cad* on all genes except *tll* show a strong negative correlation, hence strong weights reduce robustness [In Fig. 1 Additional file 1, the correlations is shown separately for circuits with promoter threshold $H = -2.5$ and $H = -3.5$].

Robustness towards parameter perturbation

If a parameter in one of the circuits (considered as a model input) is perturbed by a certain amount the formation of the gene expression pattern (considered as the model output) will be perturbed as well. The amount by which a parameter can be increased and decreased before a certain pattern error (RMS) is reached can be used as a measure for pattern sensitivity with respect to that parameter. In order to get a better understanding of the model's sensitivity, for each circuit and each parameter in that circuit the sensitivity interval (SI) was calculated (see methods). From a biological point view pattern formation should be robust towards small perturbations in the parameters [59].

In Fig. 6C SI versus parameter value for D_{gr} , W_{hb}^{Kni} and W_{gt}^{Kni} are plotted. For most circuits the lower value for the diffusion coefficient reaches the value zero, hence the lower bound of SI is equal to the parameter value, meaning that without diffusion the pattern still forms correctly. This phenomenon is also observed for all other diffusion coefficients. Furthermore, the upper bound of SI does not scale with the parameter value; hence circuits with very similar diffusion coefficients can have very different SIs. Non-scaling behavior is observed for most parameters, another example is W_{gt}^{Kni} shown in Fig. 6B. For this parameter, both the upper and lower bound do not scale with parameter value, also here we observe that circuits with very similar parameter values can have very different SIs; this was observed for most parameters. For some parameters the SI does however scale with parameter value. For example all decay time constants show scaling, but also some weights. Fig. 6A shows the plot of W_{kni}^{Hb} , here both the upper and lower bound of the SI decreases with smaller absolute parameter values. Hence the circuits become more sensitive towards perturbations when the parameter is smaller in magnitude [see Additional file 8 where all parameters SIs are shown].

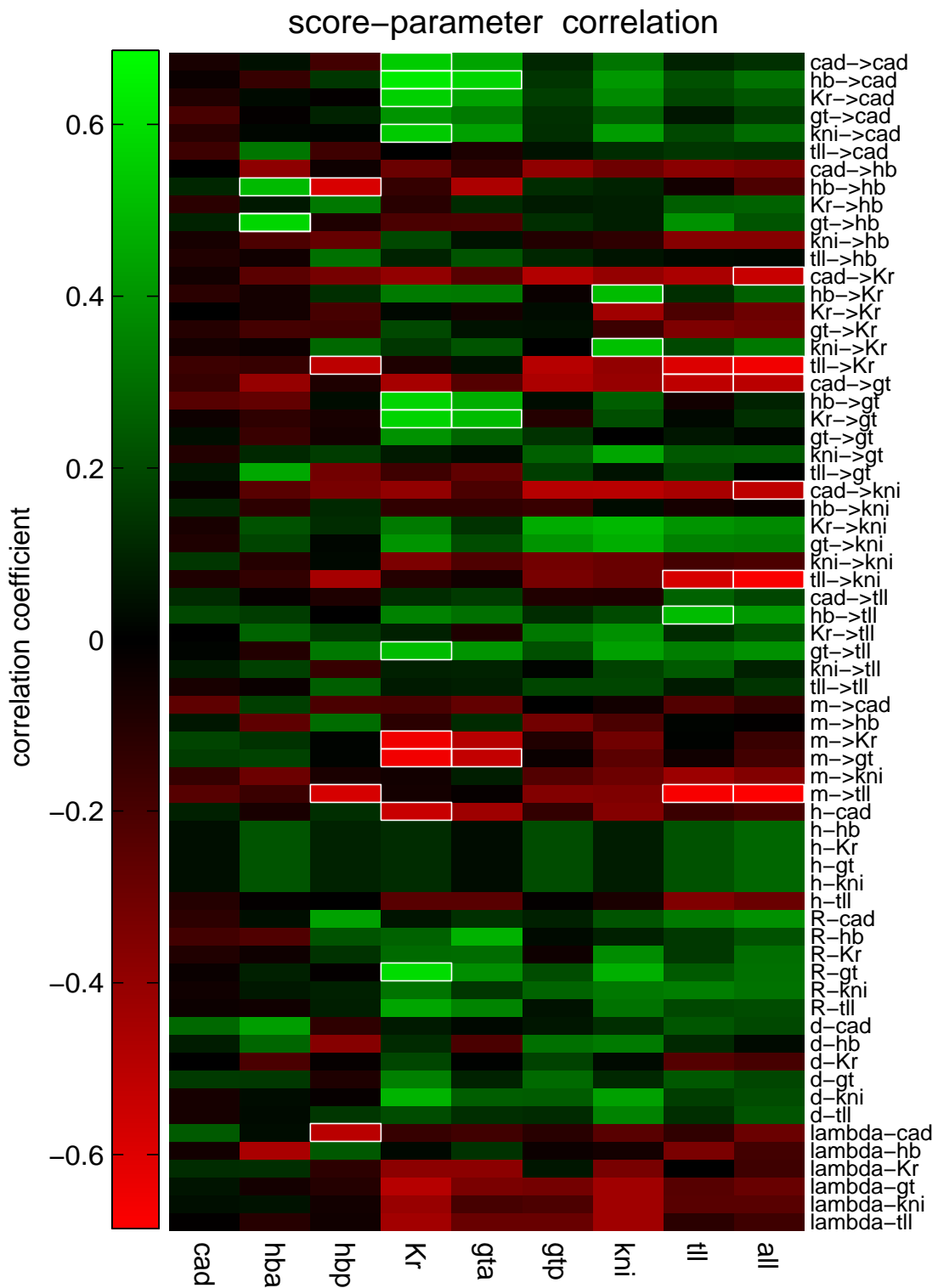


Figure 5
Correlation between the score of the eight different expression domains obtained from stochastic simulations and all parameters. Bright green represents a strong positive correlation and bright red a strong negative correlation. Squares in white borders are the most significant correlations.

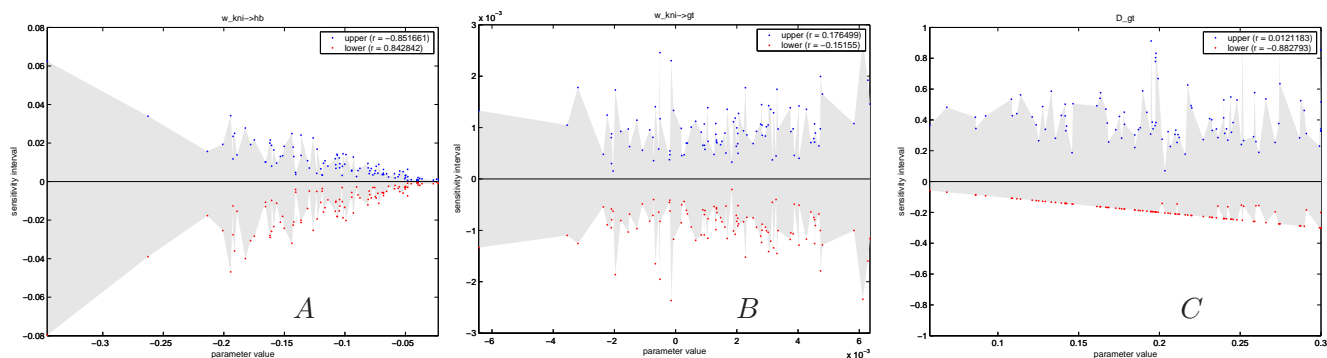


Figure 6

Distribution of sensitivity interval versus parameter value. (A) Plot of W_{kni}^{Hb} sensitivity interval. Here both the upper and lower sensitivity interval scale with absolute parameter values. Hence the parameter becomes more sensitive when it is smaller in magnitude. (B) W_{kni}^{Gt} sensitivity shows non-scaling behaviour, which is also observed for most other parameters. Furthermore, circuits with very similar parameter values can have very different sensitivities; this was also observed for most parameters. (C) Diffusion sensitivity illustrated by D_{gt} . For most circuits the lower value for the diffusion coefficient reaches zero, hence the lower sensitivity interval is equal to the parameter value, meaning that without diffusion the pattern still forms correctly.

In Fig. 7 the correlation matrix, calculated on the basis of the SIs is shown. The correlation pattern shows blocks on the diagonal, where circuit parameters that regulate one particular gene tend to cluster together. For example, in the case of *cad*, if in a particular circuit a parameter that regulates *cad* has a large SI, it is likely that the other parameters that regulate *cad* also have a larger SI. This suggests that sensitivity appears to behave in a modular fashion, where the genes represent modules. One exception is *Kr* and *gt*, which both are found in the same cluster, suggesting these two genes have a strong correlation. This correlation is caused by their strong mutual repression [52], forcing them to act in similar way to perturbation.

Circuit sensitivity

To investigate which parameters and which circuits are more sensitive than others we have calculated the average sensitivity on a logarithmic scale for each parameter and each circuit (see methods). Then the SI of each parameter value in each circuit was plotted using an intensity plot, where the colour corresponds to $-\log SI$. The parameters and circuits were ordered in Fig. 8 according to their average sensitivity, showing the parameters with smallest SI and most sensitive circuit at the top right corner. From Fig. 8 it can be seen that the difference between the smallest and largest SIs are 4 log units. The circuits are less sensitive with respect to diffusion coefficients, promoter rates, decay time constants and promoter threshold. Circuits are most sensitive with respect to W_x^{Cad} weights followed by

the auto-regulation weights W_x^X . Because the different parameter types (diffusion, decay, promoter rate, thresholds and weights) are not on the same scale we also calculated the average relative sensitivities of the circuit parameters [see Tab. 2 in Additional file 1]. This approach can however be problematic, if the parameter values are close to zero, possibly yielding a very high relative sensitivity. Therefore these outliers were removed from the analysis. Using this measure, the least sensitive parameters are the diffusion coefficients with a relative sensitivity of about 100-200%. The thresholds are now more comparable with the weights and have a relative sensitivity of about 1%. Furthermore the W_x^{Cad} weights are still amongst the most sensitive parameters with a relative sensitivity of about 0.5-1%. From Fig. 9 we see that the average sensitivity of the circuits varies from one to another. Certain parameters in the least sensitive circuits, notably weights related to *tll* and *cad* tend to have lower sensitivities. By comparing the least sensitive circuits with the more sensitive circuits using a t-test several parameters were found that are significantly different in these groups. These are in fact the parameters that show scaling between sensitivity and parameter value W_{tll}^{Hb} , W_{tll}^{Bcd} , W_{tll}^{Gt} , W_{kni}^{Tll} , W_{Kr}^{Tll} , W_{Kr}^{Cad} , W_{gt}^{Cad} and W_{kni}^{Cad} .

The standard deviation of the circuit parameters can be quite large compared to the average parameter value

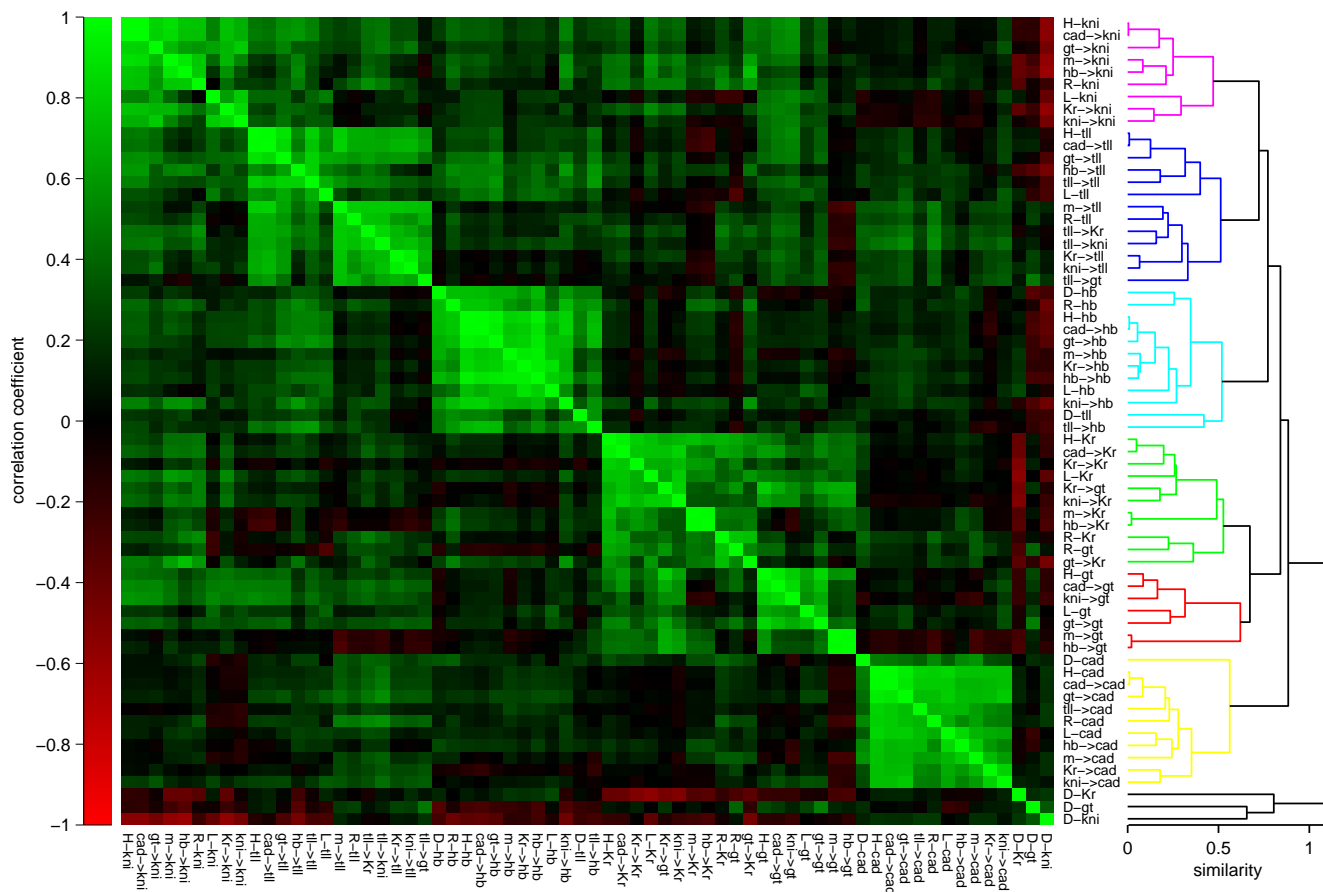


Figure 7
Parameter sensitivity intervals correlation. The correlation pattern is calculated on the basis of sensitivity intervals. The correlation matrix shows blocks on the diagonal, where circuit parameters regulating one particular gene tend to cluster together.

shown in Fig. 2. And furthermore, average parameter value and standard deviation strongly correlate ($r = 0.81$). In Fig. 9 the average $-\log$ SIs versus standard deviation is plotted. Amongst similar parameter types (weights and non-weights) no correlation is observed between SI and standard deviation. This suggests that if a parameter is very variable across the circuits, i.e. it is not well-defined [60], it is not necessary less sensitive. A notable example is W_{cad}^{Bcd} and W_{ill}^{Cad} , which both have a similar standard deviation but differ in sensitivity by 2 log units. Furthermore, we also observe that the standard deviation on average is much higher than the sensitivity interval.

Model sensitivity vs. pattern robustness

In Fig. 10 we have plotted the average sensitivity versus the overall score of the circuit. In this figure the circuits with promoter threshold $H_{hb,Kr,gt,kni} = -2.5$ are shown in

blue and the circuits with $H_{hb,Kr,gt,kni} = -3.5$ in red. The circuits with an average sensitivity higher than 2.05 are significantly less robust towards fluctuations than the circuits with an average sensitivity lower than 2.05. The circuits with $H = -3.5$ are more sensitive than the circuits with $H = -2.5$. Furthermore the robust circuits have a better overall pattern score on average; however lower sensitivity does not necessarily yield a good pattern score. This suggests that the average sensitivity is not the only determinant for robustness towards fluctuations. In Fig. 2 of Additional file 1, the correlations between parameter SIs and stochastic simulation of individual gene domain robustness is shown separately for circuits with promoter threshold $H = -2.5$ and $H = -3.5$.

Discussion

The 101 gap gene circuits were previously obtained with inverse modelling [42,51] of a seven gene (*bcd*, *cad*, *hb*, *Kr*,

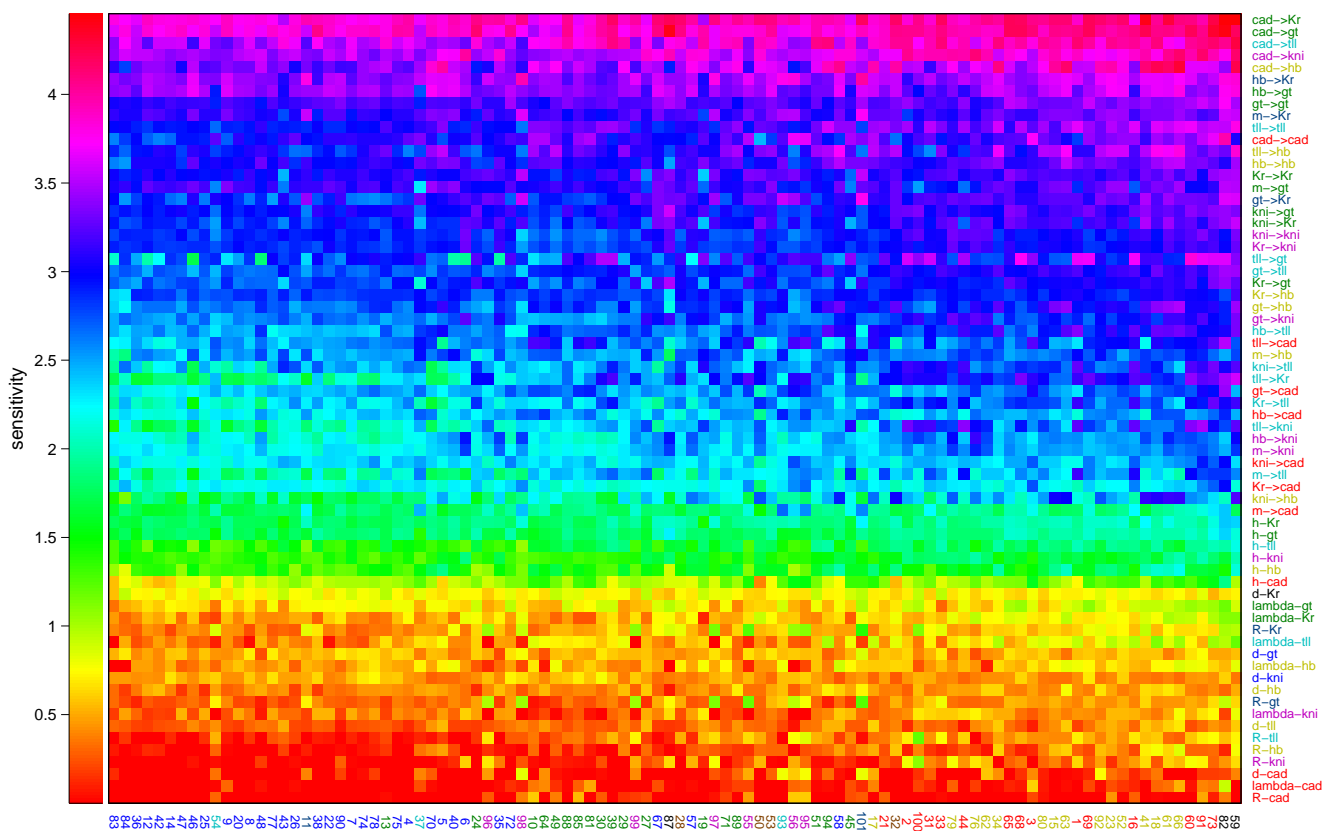


Figure 8
Relation between parameter sensitivity intervals and circuit sensitivity. Intensity plot of the sensitivity of all parameters and all circuits sorted according to the average parameter sensitivity and average circuit sensitivity. The most sensitive parameter and most sensitive circuit are shown at the top right.

gt, kni and tll) connectionist model using detailed spatio-temporal gene expression data [61]. All circuits were selected only on the basis of a low RMS value ($RMS \leq 12$), and in all cases the circuits were able to reproduce the dynamics and the patterns accurately. However, the parameter values of the circuits appeared not to be well defined, for some parameters multiple clusters were found, which represented multiple circuit topologies. Furthermore many parameters showed a single cluster with a high degree of scattering around the mean as shown in Fig. 2. Before any further tests, this result insinuates that for some parameters such as the non-regulatory parameters (showing major scattering), the notion of a global optimal parameter set makes no biological sense as suggest by Von Dassow et al. [27]. Also, the scatterings of most of the regulatory parameters are almost in a precise region of the search space, describing a precise qualitative behavior that can be linked to a potential interaction (repression, no-interaction or activation). The model always performs well and leads to very good pattern suggesting that the model is robust to the network topology

as well as variation in individual parameters. Robustness to network topology is a behavior already observed by Von Dassow et al. [27] in the case of the segment polarity of *Drosophila melanogaster*.

To further analyze the properties of the model and to classify the quality of the circuits we conducted a perturbation analysis. We analyzed the robustness of 101 *Drosophila melanogaster* gap gene circuits using two different approaches, first the robustness of pattern formation towards intrinsic fluctuations in gene expression levels was investigated using a stochastic model and secondly the sensitivity of the circuits with respect to the parameters in the model was investigated using a simple parameter perturbation technique.

Robustness towards fluctuations

Robustness of the regulatory network is expected to be a fundamental property of the network, pattern formation should therefore not be compromised by small intrinsic fluctuations in gene expression levels. Therefore, the pat-

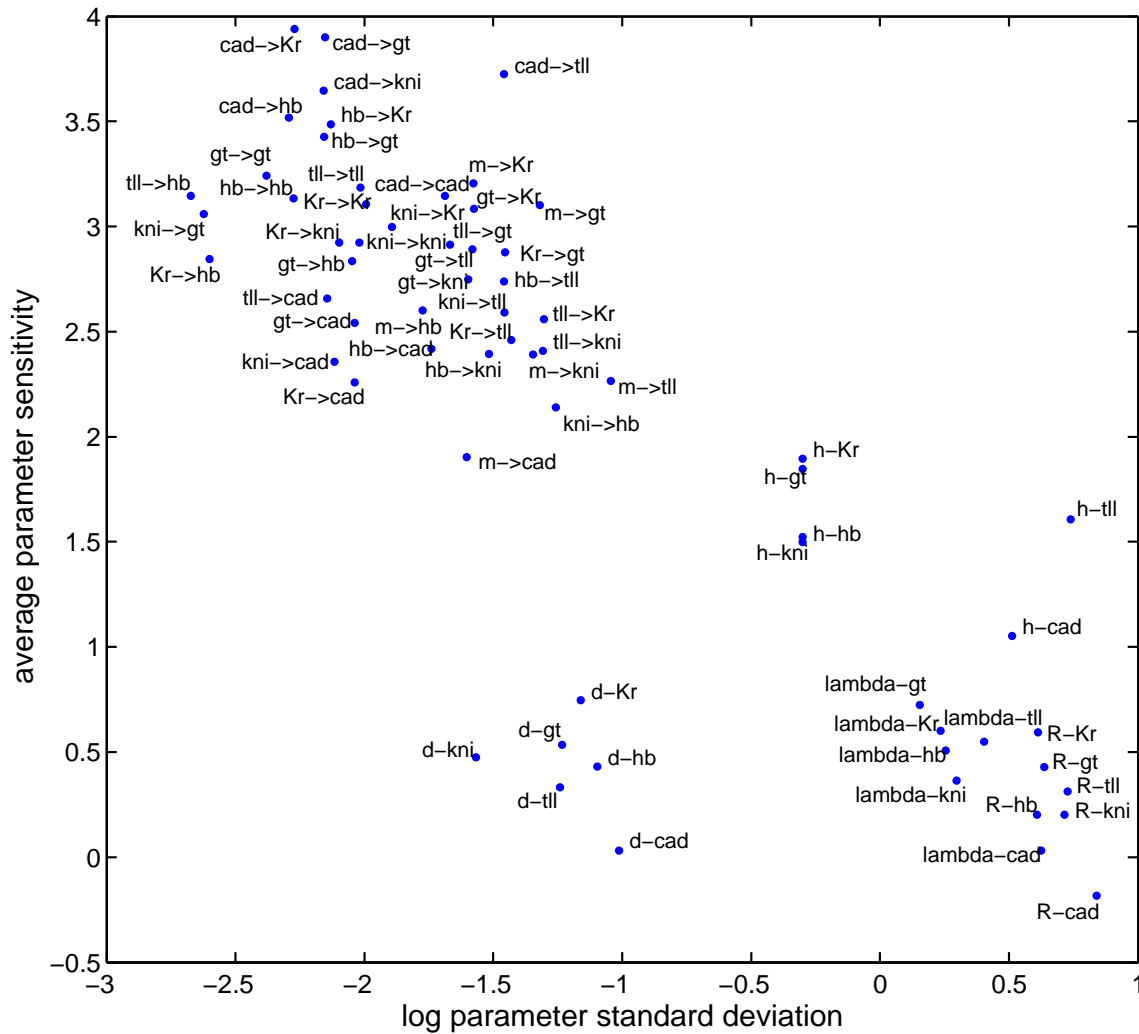


Figure 9
Relationship between average sensitivity interval and parameter standard deviation. On the y-axis, $-\log SIs$ is given and on the x-axis parameter standard deviation. The figure shows that there is no obvious correlation between SI and STD. The parameter identifiability is not necessarily linked to SI.

terns obtained from stochastic simulations should statistically be similar to the pattern obtained from the deterministic simulation for robust circuits.

Robustness of domain formation

Introducing fluctuations in the model had a profound effect on the formation of the expression patterns. Although some of the circuits showed robust formation for most domains, none of the circuits were sufficiently robust when looking at the formation of all the domains. We observe that fluctuations can lead to an increase or

decrease of domain amplitudes (see Fig. 3). Furthermore we also observe posterior and anterior boundary shifts, which leads to domain expansion, domain contraction or domain shifts. We also observe that domains completely disappear or that domains appear in other regions, where they repress other domains. Especially in a significant number of circuits the anterior domains Gt and Kr did not form robustly during the simulations. The least robust was the Tll domain at the posterior end, this domain interacted with Gt and Hb, which all showed defects in many circuits. We believe that Tll weak robustness is caused by

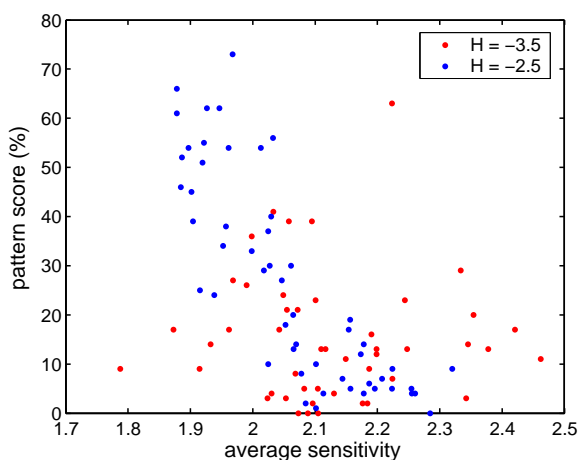


Figure 10
Scatter plot of pattern scores versus average circuit sensitivity. On the x-axis, average sensitivity of a circuit is plotted and on the y-axis the overall pattern score as percentage, which was obtained from 100 stochastic simulation runs. This figure shows that sensitive circuits are not robust towards fluctuations.

an incomplete model where gap gene are regulating *tll* which is inconsistent with experimental evidence [45,62]. However, we see that although the overall robustness to fluctuation is poor, the robustness tends to behave in a modular way, where, for a given circuit, all genes but one can have very good score. The gene with a bad stochastic simulation can be linked to a bad or wrong interaction, feature that cannot be seen from the deterministic simulation. This modular behavior shown in Fig. 11 is strongly related to the network topology as well as the interaction weights.

Looking at the evolution of the system in phase space, we found that the pathways in many cases were not well defined, which appeared to be linked to the existence of multiple attractors. The fluctuations allowed the system to jump from one attractor to another, causing the system to evolve into a pattern very different compared to the deterministic model. This suggests that by using the reverse engineering approach many circuits can be found, which are well defined for the deterministic case but not for the stochastic case. Multiple attractors can easily occur in a non-linear model with many parameters, where overfitting may lead to many connections in the circuits. Also the type of feedback loops may be crucial, e.g. *cad* has a negative auto-regulation, which reduces the effect of fluctuations, however all other genes have strong positive auto-activation [63,64]; in these cases fluctuations are amplified if there are no negative feedback loops associated with the increase of the expression level.

The effect of certain parameters on robustness

The connectionist model contains a promoter threshold for each gene, which determines if gene production is on or off when there are no other inputs. In the current gap gene model the production of *hb*, *Kr*, *gt* and *kni* can only be turned on by maternal inputs or by auto-activation, therefore the promoter threshold are set to a fixed negative value, which was either $H = -3.5$ or $H = -2.5$. Looking at the robustness of individual circuits, we find that circuits with promoter threshold set to $H = -2.5$ are more robust towards fluctuations compared to circuits with promoter threshold set to $H = -3.5$, and also their parameters are on average less sensitive. In general we find that weaker maternal inputs from bicoid and caudal increase the robustness with respect to fluctuations. Furthermore higher promoter rates, decay rates and diffusion coefficients improve the robustness towards fluctuations.

We have seen that circuits are relatively less sensitive with respect to diffusion coefficients, and with some extent, to decay time constant. Regarding the diffusion coefficient, this result confirms observations made elsewhere where it was suggested that the diffusion is not essential to explain precise gene expression pattern formation [42,65], but the diffusion term does account for the boundary nuclei effect [66]. It was also shown by Nusslein-Volhard et al. [67] that the effect of diffusion is reduced with the exponential increase of the number of nuclei. Also, Gregor et al. [68] showed that diffusion coefficient does not scale with varying embryos length. The rest of the parameters have a mixed sensitivity behavior and do not fall into a specific category [see Tab. 2 in Additional file 1]. Some of the weights are very sensitive and some are not, idem for the production rate. Although we do not have precise biological explanations regarding the difference of the sensitivity behavior of the parameters, it might be interested to experimentally investigate if such difference is a property of the connectionist model or a characteristic of the regulatory mechanism. The sensitivity information can then guide the selection of the optimal mutation targets and thereby reduce the experimental effort. This validation could be done for example by measuring mRNA degradation rate of zygotic Hb in embryos with over expressed maternal *hb*, or by measuring the binding affinity in mutants. Also one could consider inducing genetic mutation to control kinetic parameters that can be measured [69].

We also find that certain parameters that regulate posterior *tll* and also some parameters that regulate anterior *gt* and *Kr* affect robustness of the corresponding domains. Furthermore we find that robustness and parameter sensitivity are linked. Especially for the group with a lower promoter threshold we find that circuits with a lower average

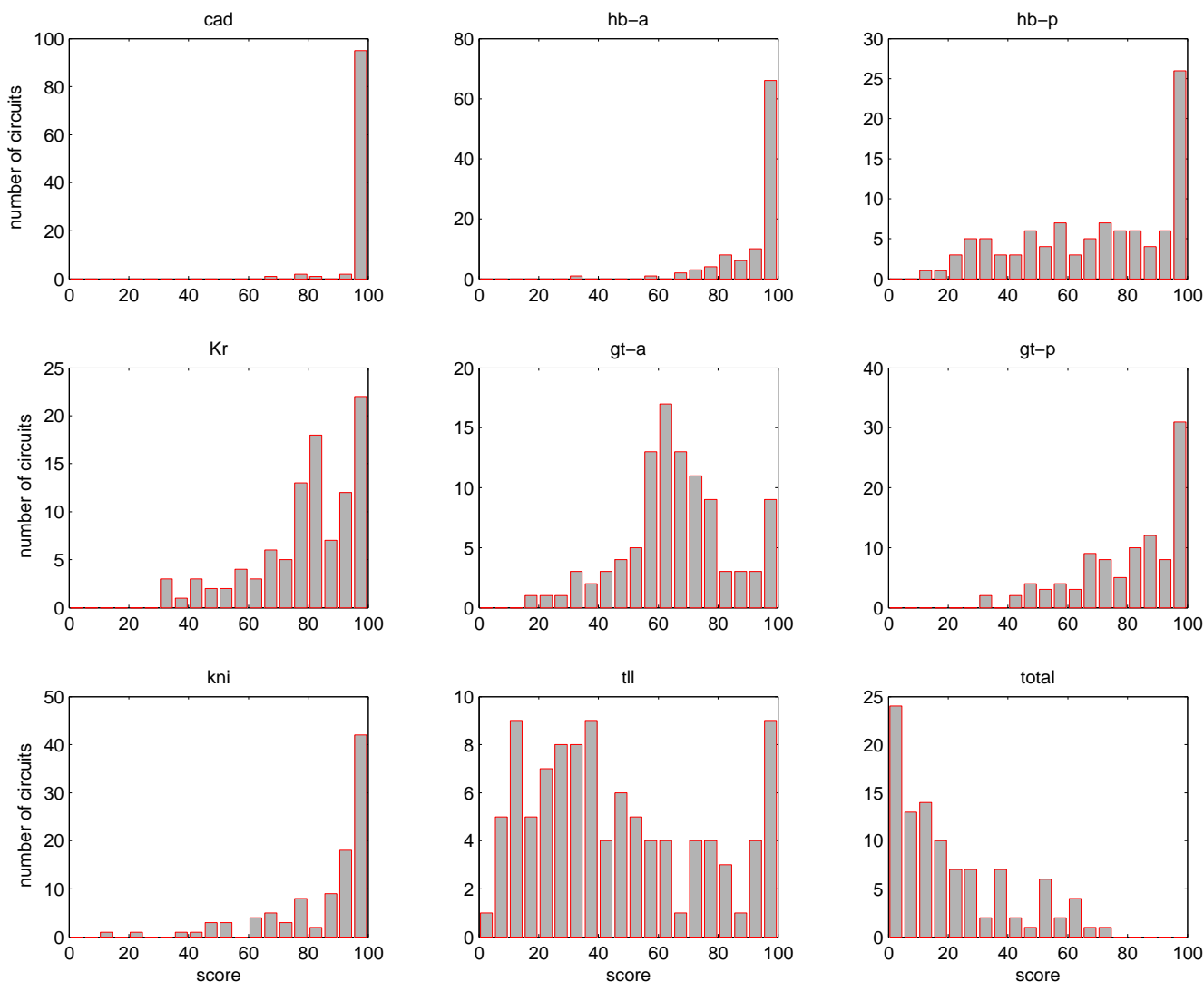


Figure 11
Robustness of overall patterns and domain formation. Each panel beside the last one shows the distribution of the circuits' robustness to gene expression fluctuation. Genes that are expressed in more than one domain (such as *hb* and *gt*) have more than one score. The score determines how precise is the stochastic simulation to the deterministic simulation. For each circuit, 100 stochastic simulations were run and score for individual gene and total pattern was calculated. On the x-axis the score is given and the y-axis determines the number of circuits with a given score. *Cad* and anterior Hb (*hb-a*) show very good score while *Tll* shows a very bad score.

sensitivity are more robust towards fluctuations. These results show that in the current gap gene model not just the network topology but to a large extent the precise value of the parameters determines robustness. We noticed that inputs of *cad* to other genes correlate negatively with robustness, these parameters are amongst the most sensitive in the model, with a $-\log SI$ in the order of 4. If we assume that fluctuations are proportional to the square root of the concentration level (this is a reasonable assumption for steady production) the relative fluctuation

level is about $1/\sqrt{n} * 100\%$, where n is the number of molecules. For $n = 100$, $n = 1000$ and $n = 10000$ the relative fluctuation level then is 10%, 3.2% and 1%. The most sensitive parameters in the model have a relative sensitivity of about 0.5 - 2% [see Tab. 2, Additional file 1]. This holds assuming that 20% of the circuit RMS is a critical cut-off.

Possible reasons for weak robustness

We conclude that the current connectionist model describing the gap gene segmentation obtained from reverse engineering techniques is not robust. Although some circuits appear more robust than others, most of the circuits are extremely sensitive towards gene expression level fluctuations and also pattern formation is very sensitive with respect to perturbation of a large number of parameters. In most cases the parameters causing this are linked to: Tll regulation, in some cases are linked to anterior Gt and Kr regulation and finally the promoter threshold (see Fig. 5). There may be a multiple reasons for weak robustness. First, the incompleteness of the model may be the cause. In the real system it is known that the terminal pathway with Tll, huckebein [70] and Torso also regulate gap genes, the latter two are missing in the current model. Furthermore, hb regulation is not only zygotic but also has a maternal component at the anterior end of the embryo that is also regulated by nanos, nanos and a maternal description *hb* mRNA are missing in the model. Furthermore, only a part of the embryo is considered, which may lead to boundary effects at the anterior end. Secondly, the current reverse engineering approach using a simple RMS optimization may lead to sensitive circuits, not only some time points are missing in the data also features in the data that are not represented in the model may be over fitted. Finally, the rather phenomenological approach of the connectionist model may also be a source of sensitivity because promoter threshold has a profound effect on robustness.

How to improve the network inference

In modelling processes from developmental biology especially the class of quantitative spatio-temporal models of gene regulation is relevant. This type of models can potentially linked with three-dimensional biomechanical models of morphogenesis and provides new insights into developmental biology. Especially the quantitative spatio-temporal models of gene regulation are characterized by a large number of unknown parameters and an (infinite) class of potential solutions. So far, very few model-based analysis method have been proposed to validate or invalidate models, especially for nonlinear and spatially distributed models [6,71,72].

The analysis presented in this paper shows the effect of missing genes, suggesting that it might be necessary to include more genes known to be involved in gap gene segmentation. Although data might not be available, one could consider combining synthetic data with observation. The hierarchical modular structure of the system should be considered within the model description [27]. Phenomenological models such as the connectionist model can generate realistic pattern, but they still fail at establishing the clear role of genetic perturbation. These

models can be improved by considering using Hill type functions [27] and/or more detailed rate equations [65]

Another improvement would be to infer the network using stochastic models that translate the fluctuations in the system. It is computationally expensive to numerically solve such systems and it will require efficient optimizations methods [73].

To prevent overfitting, one might consider reducing the number of free parameters by adding more constraints known from previous experiments. In general the inverse problem that is only based on minimizing the RMS leads to over-fitting, allowing for many solutions that all fit the dataset equally well but may not show correct behavior or properties beyond the dataset. It is therefore difficult to know, which solution is most similar to the real system. The different post-optimization analyses presented in this paper and in an accompanying manuscript ([74]) revealed some simple but efficient invalidation tests. Ideally, we would like to funnel all circuits into a pipeline of tests where the circuits are filtered and those passing all the tests are the robust and stable solutions. It may however turn out that none of the solutions will pass all the tests. Therefore we propose not only to minimize the RMS but also to incorporate other objectives into the optimization strategy. By introducing multiple objectives the solutions obtained should possess better properties and also have better defined parameters. The possible additional objectives for gene regulatory networks are:

- **Sensitivity constraints** should lead to networks that are robust towards parameter or input perturbation. This could simply be addressed by incorporating a sensitivity analysis within the optimization procedure. Using the Levenberg-marquardt method for parameter estimation, Ashryalyev et al. [52] have investigated simultaneously while estimating the parameters, their identifiability. Rodriguez-Fernandez et al. [75] have presented a hybrid method that can handle the sensitivity analysis while optimization. Using such method, we suggest attributing to each circuit a sensitivity value that determines the overall sensitivity of its parameter.
- **Noise robustness constraint** is another improvement that should lead to circuits robust towards fluctuations (internal or external). This improvement can be achieved by inferring the network using stochastic models that include the fluctuations in the system. It is computationally expensive to numerically solve such systems and it will require efficient optimizations methods [73,76]. Furthermore, It is known that patterning is insensitive to external fluctuation such as maternal gene expression. Consequently, an efficient

model should also be able to simulate gap-gene expression given noisy external Bcd expression. An efficient way to distinguish all the solutions obtained from the parameter estimation would then be to change the Bcd dosage during the optimization. It was shown by Manu et al. [77] that Bcd variation is not the cause of the precise canalization of the gap genes but its a consequence of the cross regulation between zygotic gap genes [29]. In an optimization procedure, good circuits should be able to reproduce the patterns as well as the shift without showing any major fluctuation of the gap gene expression while varying the Bcd concentration.

- **Reducing network connectivity** In our accompanying manuscript [74], we show that the long term dynamic revealed that in the presence of certain motifs, the circuits pattern converge to oscillatory attractors. This behavior is not desirable as we expect the patterns to converge to a stationary point being the steady state. Therefore, we would ideally ignore or penalized the circuits having this behavior. Also, we have shown that circuits with realistic topologies do not have these motifs and converge to the desired attractors. This could be obtained by either running the circuits beyond the data set to force asymptotic stability [66], or by adding an entropy function to prioritize circuits with the minimal connectivity. It was shown by Isalan et al. [78] that circuits where reduced connectivity are more robust to parameter variation in comparison to circuits with less connected members.

These different additional objectives would be integrated in a multi-objective optimization framework, which is an extension of the classic optimization of a single-objective function (see Handl et al. for interesting review on Multi-objective optimization [79]). Many multi-objective algorithms exist [80,81] such as method where multi objectives are transformed into one single objective [82-84], Pareto [85,86] and non-Pareto dominance approaches. To our knowledge, very few have used Multi-objective optimization for GRN inference [87,88]. Based on recent reviews [9,10], stochastic ranking evolutionary strategy (SRES) is one of the best methods for parameter estimation of biological problems. This method has been used to obtain some of the circuits analyzed in this article and obtained elsewhere [51]. The implementation of the SRES is suitable for multi-penalties where the weighted sum of penalties is used [9] as a multi-objective constraints. The strength of the ES is its intrinsic parallel nature [89]. One alternative would be to use an island based ES where the number of island is determined by the different objectives. Each island minimizes the multi-objectives, but the distribution of the weight is different from island to island. Migration between islands will spread individuals

with a particular strong property in the other islands. The main objective being the least square difference between the simulation and the data, it might be more practical to start with one population where the penalties have equal weights and later on, switch to islands once the LSE is acceptable. This would guarantee that in all islands, the individuals have at least a good RMS.

Conclusion

In this paper, we have provided a robust analysis of a model used to infer the gap gene regulatory network of *Drosophila melanogaster*. The model has been extensively used elsewhere [42,43,51] to simulate and provide some explanations concerning the regulatory mechanism that leads to precise pattern formation. Unfortunately, many assumptions were based on a limited number of circuits obtained using simulated annealing and previous researchers assumed that the model was correct with a correct topology. In this article, we have shown that the mathematical model leads to different circuits all capable of reproducing the quantitative spatio-temporal gene expression pattern. Consequently, it is difficult to decide based solely on the architecture which circuit is the correct one.

Robustness towards fluctuation has revealed that the overall gap gene domain tends to be poorly resistant to perturbation and this weak property could be related to some particular interactions predicted by the circuits. Furthermore, parameter perturbation analysis has shown that the circuits with lower sensitivity do not necessarily yield to robustness to fluctuation. The reason for these few exceptions is related to the promotor threshold and to local domain robustness, which can considerably affect the overall global robustness. Overall, the analysis shows that the network possesses modular robustness and some local properties may affect the robustness of a gene expression locally as show in Fig. 11. This feature is strongly related to the network topology as well as the interaction weights.

From a biological point of view, this paper has shown that it is difficult to relate the connectionist model with biological evidence. The model ability to simulate the gene expression does not necessarily provide meaningful information since alternative networks are predicted. Therefore, it would be interesting to test some of the different alternatives, especially when there is not yet any experimental evidence to invalidate. However, it was recently demonstrated that it is still possible to draw qualitative conclusions on the regulatory topology of the gap gene network [52]. The overall analysis has shown that based on the robustness toward gene expression fluctuations and parameter perturbations, it is possible to identify robust circuits as well as the parameters that are identifiable according to their sensitivity intervals. For a computa-

tional/system biologist, this shows that it is essential to further analyze a model prediction, when results are obtained from reverse engineering based on parameter estimation, since some of its properties may or may not invalidate the results. We have also provided some preliminary suggestions to efficiently improve the GRN inference to avoid reverse engineering that leads to circuits with different topologies, by controlling the optimization by means of multi-objectives minimization.

Methods

Inference of the Gap Gene model

The gap gene circuits analyzed in this paper were presented by Jaeger et al. [42] and Fomekong-Nanfack et al. [51]. In both cases, the inference was performed using the same quantitative data, the same model description but different parameter estimation methods.

Quantitative data used are available online in the FlyEx database <http://urchin.spbcas.ru/flyex/> or <http://flyex.ams.sunysb.edu/flyex>. The database presents a collection of spatio-temporal gene expression data obtained from fluorescently stained wild-type embryos for Eve protein and two other genes. Data were obtained by applying different image processing strategies [46]. The embryos are for different time ranging from cycle 7 to cycle 14A. In the simulation, data obtained at cycle 12 were used as initial conditions. For the genes *Kr*, *gt*, *kni*, *Tll* these are very close to zero and set to 0 in the simulations.

Mathematical model of gap gene considers the 35% to 92% of the A-P axis of an embryo. It is reduced to a one-

dimensional discrete model where nuclei are aligned horizontally. The model focuses on the development between cycle 13 and cycle 14A8, before gastrulation (71.1 min). Three rules describe the mechanism during that phase: interphase, mitosis and division [90]. Interphase and mitosis are continuous stages describing the dynamic of protein variation of a gene within a nucleus. The division is a discrete process describing the cleavage of a nucleus in two. Mitosis, arising before division, differs from interphase by the absence of protein synthesis. The resulting model is a system of 180 equations before division and 348 equations after, with a total of 66 unknown parameters written as:

$$\frac{dg_i^a(t)}{dt} = \begin{cases} R_a \Phi_a \left(\sum_{b=1}^{N_g} W_a^b g_i^b + m_a g_i^{bcd} + h_a \right) \\ -\lambda_a g_i^a \\ +D_a (g_{i+1}^a - 2g_i^a + g_{i-1}^a) \end{cases} \quad (1)$$

where N_g denotes the number of genes or gene products involved and Φ is a sigmoid function with range (0,1). $g_i^a(t)$ represents the concentration level at time t of gene a in nucleus i with $1 \leq i \leq N$ and N the number of nuclei during a cleavage cycle. The concentration, g_i^{bcd} , of the maternal gene *bicoid* is taken from experimental observations and is kept constant in time during the simulation.

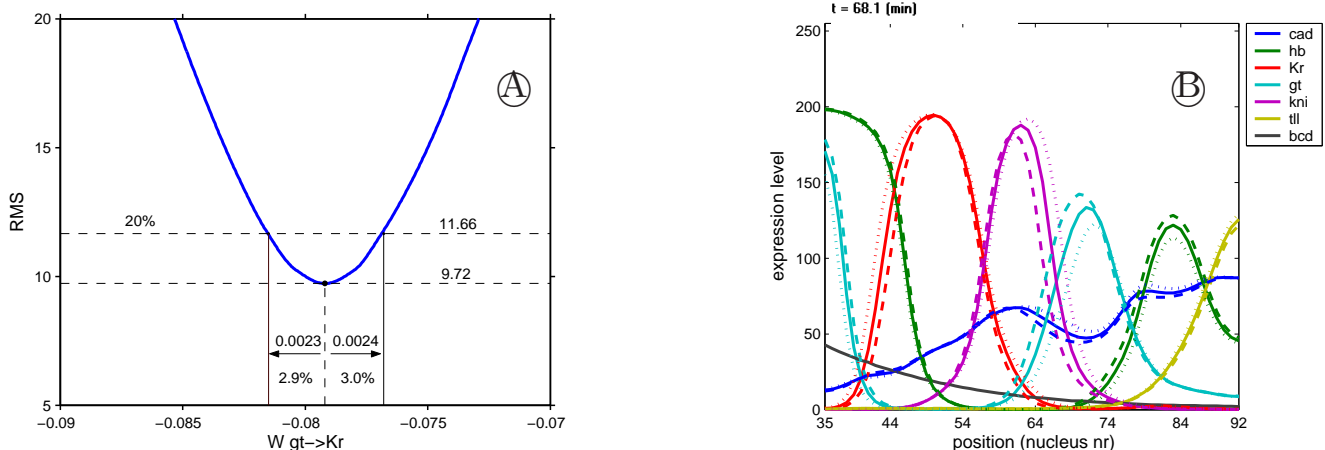


Figure 12

Parameter perturbation. For each parameter within each circuit, lower and upper parameters bound are calculated by tweaking the original value until the RMS is increased by 20%. (A) perturbation of the parameter W_{Kr}^{Gt} and in (B) simulated gene expression after parameter perturbation. The solid lines represent the original circuit's gene expression. The small dashed lines and big dashed lines are the profiles corresponding respectively to the lower and upper parameter bounds.

The parameters are: the regulatory weight matrix W_a^b , describing the influence of gene b on gene a , the production rate R_a , the activation threshold h_a for Φ , the decay rate λ_a , the diffusion coefficient D_a , and the regulatory influence bcd_a .

Parameter estimation was performed by two different strategies. Jaeger et al. [42] have used a parallel simulated annealing (PLSA) algorithm as described in [48], originally proposed by Lam [49]. The expensive computational time required by PLSA could only lead to 10 gap gene circuits with good solution's quality and patterns behavior. Later on, Fomekong-Nanfack et al. [51] proposed 101 gap gene circuits obtained using hybrid methods composed of a stochastic ranking evolution strategy [91] followed by direct search [92]. Evolution strategy is a evolution based algorithm like genetic algorithm [93] and direct search is a local search strategy suitable to solve a variety of optimization problems that are not well suited for standard optimization algorithms, including problems in which the objective function is discontinuous, non-differentiable, stochastic, or highly nonlinear. This number of solution could be obtained because of the reasonably low computational time of their method (8 h on single processor) compare to PLSA (1 to 5 days on 10 parallel CPU), but leading to the same quality of solution. In both cases, the optimization goal is to estimate the unknown parameters that minimize a scalar valued *cost-function*, by exploring the set of possible values in an allowed search space. The chosen cost-function is the least squares of the difference of the simulated and the observed data:

$$E(\theta) = \sum_{i,t} (g_i^a(t, \theta)_{model} - g_i^a(t)_{data})^2, \quad (2)$$

with θ the parameter vector to which a constraint or penalty function is added. An explicit search-space constraint is given for parameters R_a , λ_a and D_a . For the parameters W_a^B , bcd_a and h_a a collective penalty function is used ([40]) to restrict the function value of Φ to the domain $[\Lambda, 1-\Lambda]$ with Λ a small parameter (in this study taken to be 0.001). The root mean square (RMS) described by Reinitz et al. is used ([40]) as a measure of the quality of a model solution for a given set of parameters:

$$RMS = \sqrt{\frac{E(\theta)}{N_d}} \quad (3)$$

where $E(\theta)$ is given by Equation (2) and N_d is the number of data points.

Sensitivity analysis and confidence interval

If a single parameter is slightly decreased or increased, the RMS of the simulated pattern will increase because the quality of the fit to the gene expression data reduces. The amount, or range, by which a parameter can be changed before the fit becomes significantly bad, is a measure for the sensitivity of the parameter for a particular solution. For each parameter within each solution we computed the lower and upper value of a parameter where the RMS increases by 20%, which corresponds to a situation where the gene pattern becomes significantly bad as shown in Fig. 12. Hence, the lower value $\theta_{n,i} - \Delta L_{n,i}$ is computed from $f(\theta_{n,i} - \Delta L_{n,i}) = 1.2 * RMS_n$ and the upper value $\theta_{n,i} + \Delta U_{n,i}$ is computed from $f(\theta_{n,i} + \Delta U_{n,i}) = 1.2 * RMS_n$, where RMS is the original RMS value of the fit obtained from the optimization and f denotes the cost function, n denotes the solution number and i the parameter index. We define the sensitivity interval as $[\theta_{n,i} - \Delta L_{n,i}, \theta_{n,i} + \Delta U_{n,i}]$ and the sensitivity as $S_{n,i} = -\log(\frac{1}{2} \Delta L_{n,i} + \frac{1}{2} \Delta U_{n,i})$. The average sensitivity of a solution and a parameter are defined as $S_n = \sum_i^{np} S_{n,i} / N_{np}$ where np is the number of parameters and respectively $S_{\theta_i} = \sum_n^{ns} S_{n,i} / N_{ns}$ where ns is the number of circuit. The lower relative sensitivity is defined as $\theta_{n,i} / \Delta L_{n,i} * 100\%$ and the upper relative sensitivity is defined as $\theta_{n,i} / \Delta U_{n,i} * 100\%$.

Stochastic behavior of the model

We implemented a nonlinear stochastic deterministic differential equation model of the gap gene where the dynamic is driven by Gillespie stochastic simulation algorithm [56]. The master equation is based on the differential equation describing the concentration change of a gene expression in a particular nucleus. Each equation describes three different reactions: protein production, protein decay and diffusion. Only one reaction can occur at a specific time. The Gillespie algorithm simulates the system by choosing first in a probabilistic manner which reaction occurs, and then estimates when does the next reaction will be realized. To speed the process, we have used a method [94] that only calculates and updates reactions that have changed. The probabilistic nature of the algorithm imposed us to run 100 simulations for each of the 101 gap gene circuits. All simulations used the same initial condition and the number of molecules is 5 times the deterministic case (based on [55]). For simplicity, analyses are only made on the final pattern (gastrulation time) by comparing the deterministic simulation with the stochastic one.

Authors' contributions

All authors participated in the design of the study, and wrote the manuscript together. YF-N and MP conducted the analysis of data and simulations, and made the figures. All authors read and approved the final manuscript.

Additional material

Additional file 1

Additional statistics. This file (*GapGeneModelRobustnessAddFile1.pdf*) contains the material, which is not given in the paper due to the space limitation. It mainly gives additional statistical results for the stochastic simulations and the perturbation analysis.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1752-0509-3-94-S1.PDF>]

Additional file 2

Stochastic simulation of all circuits. The document (*GapGeneModelRobustnessAddFile2.pdf*) provided gives the simulation obtained from the 100 stochastic runs of each circuit obtained deterministically. Each page corresponds to a circuit and on each page; individual panels correspond to individual run. The panels with a grey background represent runs that have a defect in one or more expression domains. The runs shown in the panels with a white background are considered to be correct.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1752-0509-3-94-S2.PDF>]

Additional file 3

Robustness of all circuits. In this document (*GapGeneModelRobustnessAddFile3.pdf*), each figure shows the deterministic stochastic simulation of a circuit. In these graphs the dashed lines represent the final profiles obtained from the deterministic model and the solid lines the average profiles obtained from the stochastic simulations calculated from the 100 individual stochastic runs, which are shown in grey.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1752-0509-3-94-S3.PDF>]

Additional file 4

stochastic simulation of circuit number 5 run 52. The movie shows the spatio-temporal simulation of circuit number 5 for the 52 stochastic run.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1752-0509-3-94-S4.MP4>]

Additional file 5

stochastic simulation of circuit number 5 run 83. The movie shows the spatio-temporal simulation of circuit number 5 for the 83 stochastic run.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1752-0509-3-94-S5.MP4>]

Additional file 6

stochastic simulation of circuit number 11 run 64. The movie shows the spatio-temporal simulation of circuit number 11 for the 64 stochastic run.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1752-0509-3-94-S6.MP4>]

Additional file 7

stochastic simulation of circuit number 11 run 98. The movie shows the spatio-temporal simulation of circuit number 11 for the 98 stochastic run.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1752-0509-3-94-S7.MP4>]

Additional file 8

Absolute parameters sensitivity intervals of all parameters. The data provided in this document (*GapGeneModelRobustnessAddFile8.pdf*) represent the sensitivity interval of all the 66 parameters obtained from the 101 circuits. Each figure gives the upper and lower bound of the SI.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1752-0509-3-94-S8.PDF>]

Acknowledgements

This work was supported by the Netherlands Organization for Scientific Research, project NWO-CLS 635.100.010 (Simulation of developmental regulatory networks", <http://www.science.uva.nl/research/scs/3D-RegNet>) and by the EC (MORPHEX, NEST Contract No 043322). We used data from the FlyEx database <http://flyex.ams.sunysb.edu>. We thank Dr. Johannes Jaeger for his wise suggestions.

References

- Davidson E: *The regulatory genome - Gene regulatory networks in development and evolution* London: Academic Press; 2006.
- Kitano H: **Looking beyond the details: a rise in system-oriented approaches in genetics and molecular biology.** *Curr Genet* 2002, **41**:1-10.
- Guet CC, Elowitz MB, Hsing W, Leibler S: **Combinatorial Synthesis of Genetic Networks.** *Science* 2002, **296**:1466-1470.
- Julien Gagneur GC: **From molecular networks to qualitative cell behavior.** *FEBS letters* 2005, **579**(8):1867-1871.
- Batt G, Ropers D, de Jong H, Geiselmann J, Mateescu R, Page M, Schneider D: **Validation of qualitative models of genetic regulatory networks by model checking: analysis of the nutritional stress response in Escherichia coli.** *Bioinformatics* 2005, **21**(Suppl 1):i19-28.
- Reeves GT, Muratov CB, Schupbach T, Shvartsman SY: **Quantitative models of developmental pattern formation.** *Dev Cell* 2006, **11**(3):289-300.
- de Jong H: **Modeling and Simulation of Genetic Regulatory Systems: A Literature Review.** *J Comput Biol* 2002, **9**:67-103.
- D'haeseleer P: **Reconstructing Gene Network from Large Scale.** In *PhD thesis* University of New Mexico; 2000.
- Moles CG, Mendes P, Banga JR: **Parameter Estimation in Biochemical Pathways: A Comparison of Global Optimization Methods.** *Genome Res* 2003, **13**(11):2467-2474.
- Ashyraliyev M, Fomekong-Nanfack Y, Kaandorp JA, Blom J: **Systems biology: parameter estimation for biochemical models.** *FEBS Journal* 2008, **276**(4):886-902.
- Krishnan A, Giuliani A, Tomita M: **Indeterminacy of Reverse Engineering of Gene Regulatory Networks: The Curse of Gene Elasticity.** *PLoS ONE* 2007, **2**(6):e562.
- Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, Sethna JP: **Universally sloppy parameter sensitivities in systems biology models.** *PLoS Comput Biol* 2007, **3**(10):1871-1878.

13. Jaqaman K, Danuser G: **Linking data to models: data regression.** *Nat Rev Mol Cell Biol* 2006, **7(11)**:813-819.
14. Poolla K, Khargonekar P, Tikku A, Krause J, Nagpal K: **A Time-Domain Approach to Model Validation.** *IEEE Trans Automat Contr* 1994, **39**:951-959.
15. Carlson J, Doyle J: **Complexity and robustness.** *Proc Natl Acad Sci* 2002, **99**:2538-2545.
16. Kitano H: **Towards a theory of biological robustness.** *Mol Syst Biol* 2007, **3**:137.
17. Jen E: **Stable or robust? What's the difference?** *Complexity* 2003, **8**:12-18.
18. Wagner A: **Robustness and evolvability: a paradox resolved.** *Proc R Soc Lond B Biol Sci* 2008, **275**:91-100.
19. Stelling J, Gilles E, Doyle F III: **Robustness properties of circadian clock architectures.** *Proc Natl Acad Sci USA* 2004, **101**:13210-13215.
20. Stelling J, Sauer U, Szallasi Z, Doyle FJ, Doyle J: **Robustness of Cellular Functions.** *Cell* 2004, **118(6)**:675-685.
21. Thattai M, van Oudenaarden A: **Intrinsic noise in gene regulatory networks.** *Proc Natl Acad Sci* 2001, **98**:8614-8619.
22. Blake W, Kaern M, Cantor C, Collins J: **Noise in eukaryotic gene expression.** *Nature* 2003, **422**:633-637.
23. Daniel E, Zak GEG, Schwaber JS, Francis J, Doyle I: **Importance of Input Perturbations and Stochastic Gene Expression in the Reverse Engineering of Genetic Regulatory Networks: Insights From an Identifiability Analysis of an In Silico Network.** *Genome Research* 2003.
24. Jin Y, Lindsey M: **Stability analysis of genetic regulatory network with additive noises.** *BMC Genomics* 2008, **9(Suppl 1)**:S21.
25. Rao CV, Wolf DM, Arkin AP: **Control, exploitation and tolerance of intracellular noise.** *Nature* 2002, **420**:231-237.
26. Barkai N, Leibler S: **Robustness in simple biochemical networks.** *Nature* 1997, **387**:913-917.
27. George von Dassow EMM, Eli Meir, Odell GM: **The segment polarity network is a robust developmental module.** *Nature* 2000, **406**:188-192.
28. Kastner J, Solomon J, Fraser S: **Modeling a Hox gene network in silico using a stochastic simulation algorithm.** *Developmental Biology* 2002, **246**:122-131.
29. Jackle H, Tautz D, Schuh R, Seifert E, Lehmann R: **Cross-regulatory interactions among the gap genes of Drosophila.** *Nature* 1986, **324(6098)**:668-670.
30. Reinitz J, Levine M: **Control of the initiation of homeotic gene expression by the gap genes giant and tailless in Drosophila.** *Dev Biol* 1990, **140**:57-72.
31. Capovilla M, Eldon ED, Pirrotta V: **The giant gene of Drosophila encodes a b-ZIP DNA-binding protein that regulates the expression of other segmentation gap genes.** *Development* 1992, **114**:99-112.
32. DE C, MSG C, X W, A P, D P: **A self-organizing system of repressor gradients establishes segmental complexity in Drosophila.** *Nature* 2003, **426**:849-853.
33. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P: *Molecular Biology of the Cell*, Garland Fourth edition. 2002.
34. Rolando Rivera-Pomar HJ: **From gradients to stripes in Drosophila embryogenesis: filling in the gaps.** *Trends in Genetics* 1996, **12(11)**:478-483.
35. Akam M: **The molecular basis for metameric pattern in the Drosophila embryo.** *Development* 1987, **101**:1-22.
36. Ingham PW: **The molecular genetics of embryonic pattern formation in Drosophila.** *Nature* 1988, **335(6185)**:25-34.
37. St Johnston D, Nusslein-Volhard C: **The origin of pattern and polarity in the Drosophila embryo.** *Cell* 1992, **68(2)**:201-219.
38. Sánchez L, Thieffry D: **A logical analysis of the gap gene system.** *Theor Biol* 2001, **211**:114-141.
39. Mjolsness E, Sharp DH, Reinitz J: **A connectionist model of development.** *J Theor Biol* 1991, **152(4)**:429-453.
40. Reinitz J, Sharp DH: **Mechanism of eve stripe formation.** *Mech Dev* 1995, **49(1-2)**:133-158.
41. Sharp DH, Reinitz J: **Prediction of mutant expression patterns using gene circuits.** *Biosystems* 1998, **47(1-2)**:79-90.
42. Jaeger J, Surkova S, Blagov M, Janssens H, Kosman D, Kozlov KN, Myasnikova E, Vanario-Alonso CE, Samsonova M, Sharp DH, Reinitz J: **Dynamic control of positional information in the early Drosophila embryo.** *Nature* 2004, **430(6997)**:368-371.
43. Perkins TJ, Jaeger J, Reinitz J, Glass L: **Reverse engineering the gap gene network of Drosophila melanogaster.** *PLoS Comput Biol* 2006, **2(5)**:e51.
44. Janssens H, Hou S, Jaeger J, Kim AR, Myasnikova E, Sharp D, Reinitz J: **Quantitative and predictive model of transcriptional control of the Drosophila melanogaster even-skipped gene.** *Nat Genet* 2006, **38(10)**:1159-1165.
45. Jaeger J, Blagov M, Kosman D, Kozlov KN, Myasnikova E, Surkova S, Vanario-Alonso CE, Samsonova M, Sharp DH, Reinitz J: **Dynamic analysis of regulatory interactions in the gap gene system of Drosophila melanogaster.** *Genetics* 2004, **167(4)**:1721-1737.
46. Myasnikova E, Samsonova A, Kozlov K, Samsonova M, Reinitz J: **Registration of the expression patterns of Drosophila segmentation genes by two independent methods.** *Bioinformatics* 2001, **17**:3-12.
47. Poustelnikova E, Pisarev A, Blagov M, Samsonova M, Reinitz J: **A database for management of gene expression data in situ.** *Bioinformatics* 2004, **20(14)**:2212-2221.
48. Chu KW, Deng Y, Reinitz J: **Parallel simulated annealing by mixing of states.** *J Comput Phys* 1999, **148(2)**:646-662.
49. Lam J, Delosme JM: **An Efficient Simulated Annealing Schedule: Derivation.** In *Tech Rep 8816* Electrical Engineering Department, Yale, New Haven, CT; 1988.
50. Lam J, Delosme JM: **An Efficient Simulated Annealing Schedule: Implementation and Evaluation.** In *Tech Rep 8817* Electrical Engineering Department, New Haven, CT; 1988.
51. Fomekong-Nanfack Y, Kaandorp JA, Blom J: **Efficient parameter estimation for spatio-temporal models of pattern formation: case study of Drosophila melanogaster.** *Bioinformatics* 2007, **23(24)**:3356-3363.
52. Ashyraliyev M, Jaeger J, Blom JG: **Parameter estimation and determinability analysis applied to Drosophila gap gene circuits.** *BMC Systems Biology* 2008, **2(83)**.
53. Goodwin BC, Kauffman S, Murray JD: **Is Morphogenesis an Intrinsically Robust Process?** *Journal of Theoretical Biology* 1993, **163**:135-144.
54. Azevedo RB, Leroi AM: **A power law for cells.** *Proc Natl Acad Sci USA* 2001, **98(10)**:5699-5704.
55. Wu YF, Myasnikova E, Reinitz J: **Master equation simulation analysis of immunostained Bicoid morphogen gradient.** *BMC Syst Biol* 2007, **1**:52.
56. Gillespie DT: **Exact stochastic simulation of coupled chemical reactions.** *Journal of Physical Chemistry* 1977, **81(22)**:2340-2361.
57. Tautz D: **Regulation of the Drosophila segmentation gene hunchback by two maternal morphogenetic centres.** *Nature* 1988, **332(6161)**:281-284.
58. Eldon E, Pirrotta V: **Interactions of the Drosophila gap gene giant with maternal and zygotic pattern-forming genes.** *Development* 1991, **111(2)**:367-378.
59. Umulis D, O'Connor MB, Othmer HG: **Robustness of embryonic spatial patterning in Drosophila melanogaster.** *Curr Top Dev Biol* 2008, **81**:65-111.
60. Sun N, Sun NZ, Elimelech M, Ryan JN: **Sensitivity analysis and parameter identifiability for colloid transport in geochemically heterogeneous porous media.** *Water Resources Research* 2001, **37(2)**:209-222.
61. Myasnikova EM, Samsonova AA, Samsonova MG, Reinitz J: **Spatial registration of in situ gene expression data.** *Molecular Biology* 2001, **35(6)**:955-960.
62. Bronner G, Jackle H: **Control and function of terminal gap gene activity in the posterior pole region of the Drosophila embryo.** *Mech Dev* 1991, **35(3)**:205-211.
63. Dunlap JC: **Molecular Bases for Circadian Clocks.** *Cell* 1999, **96(2)**:271-290.
64. Cross FR, Siggia ED: **Shake It, Don't Break It: Positive Feedback and the Evolution of Oscillator Design.** *Developmental Cell* 2005, **9(3)**:309-310.
65. Alves F, Dilao R: **Modeling segmental patterning in Drosophila: Maternal and gap genes.** *J Theor Biol* 2006, **241(2)**:342-359.
66. Gursky V, Kozlov K, Samsonov A, Reinitz J: **Model with asymptotically stable dynamics for Drosophila gap gene network.** *Biophysics* 2008, **53(2)**:164-176.
67. Nusslein-Volhard C: **Gradients that organize embryo development.** *Sci Am* 1996, **275(2)**:54-55.
68. Gregor T, Bialek W, de Ruyter van Steveninck RR, Tank DW, Wieschaus EF: **Diffusion and scaling during early embryonic pat-**

- tern formation. *Proc Natl Acad Sci USA* 2005, **102(51)**:18403-18407.
69. Feng XJ, Hooshangi S, Chen D, Li G, Weiss R, Rabitz H: **Optimizing genetic circuits by global sensitivity analysis.** *Biophys J* 2004, **87(4)**:2195-2202.
 70. Weigel D, Jurgens G, Klingler M, Jackle H: **Two gap genes mediate maternal terminal pattern information in Drosophila.** *Science* 1990, **248(4954)**:495-498.
 71. El-Samad H, Prajna SAP, Doyle J, Khammash M: **Advanced methods and algorithms for biological networks analysis.** *Proc IEEE* 2006, **94**:832-853.
 72. Prajna S: **Barrier certificates for nonlinear model validation.** *Automatica* 2006, **42**:117-126.
 73. Tomshine J, Kaznessis YN: **Optimization of a stochastically simulated gene network model via simulated annealing.** *Biophys J* 2006, **91(9)**:3196-3205.
 74. Fomekong-Nanfack Y, Postma M, Kaandorp JA: **Inferring Drosophila gap gene regulatory network: pattern analysis of simulated gene expression profiles and stability analysis.** [To appear in *BMC System Biology*].
 75. Rodriguez-Fernandez M, Mendes P, Banga JR: **A hybrid approach for efficient and robust parameter estimation in biochemical pathways.** *Biosystems* 2006, **83(2-3)**:248-265.
 76. Wilkinson DJ: **Stochastic modelling for quantitative description of heterogeneous biological systems.** *Nat Rev Genet* 2009, **10(2)**:122-133.
 77. Manu , Surkova S, Spirov AV, Gursky VV, Janssens H, Kim AR, Radulescu O, Vanario-Alonso CE, Sharp DH, Samsonova M, Reinitz J: **Canalization of Gene Expression in the Drosophila Blastoderm by Gap Gene Cross Regulation.** *PLoS Biology* 2009, **7(3)**.
 78. Isalan M, Lemerle C, Serrano L: **Engineering gene networks to emulate Drosophila embryonic pattern formation.** *PLoS Biol* 2005, **3(3)**:e64.
 79. Handl J, Kell DB, Knowles J: **Multiobjective optimization in bioinformatics and computational biology.** *IEEE/ACM Trans Comput Biol Bioinform* 2007, **4(2)**:279-292.
 80. Deb K: *Multi-Objective Optimization using Evolutionary Algorithms* Wiley; 2001.
 81. Colette Y, Siarry P: *Multiobjective Optimization, Principles and Case Studies* Springer; 2003.
 82. Syswerda G, Palmucci J: **The application of genetic algorithms to resource scheduling.** In *Fourth Int Conf on Genetic Algorithms ICGA'4* Edited by: Belew R, Booker L. San Mateo, California.: Morgan Kaufmann Pub; 1991:502-508.
 83. Jakob W, Gorges-Schleuter M, Blume C: **Application of genetic algorithms to task planning and learning.** In *Parallel Problem Solving from Nature PPSN'2, LNCS* Edited by: Manner R, Manderick B. Amsterdam. North-Holland; 1992:291-300.
 84. Jones G, Brown R, Clark D, Willett P, Glen R: **Searching databases of twodimensional and three-dimensional chemical structures using genetic algorithms.** In *Fifth Int. Conf. on Genetic Algorithms* Edited by: Forrest S. San Mateo, California.: Morgan Kaufmann; 1993:597-602.
 85. Goldberg DE: *Genetic Algorithms in Search, Optimization, and Machine Learning* Addison-Wesley Professional; 1989.
 86. Bentley P, Wakefield J: *Soft Computing in Engineering Design and Manufacturing, 1997 chap. Finding acceptable Pareto-optimal solutions using multiobjective genetic algorithms* :231-240.
 87. Praveen K, Sanjoy D, Stephen W, L RJ: **A multi-objective GA-simplex hybrid approach for gene regulatory network models.** *IEEE Proceedings of the 2004 congress on evolutionary computation: CEC 2004* 2004:2084-2090.
 88. Liu PK, Wang FS: **Inference of biochemical network models in S-system using multiobjective optimization approach.** *Bioinformatics* 2008, **24(8)**:1085-1092.
 89. Lohmann R: **Application of Evolution Strategy in Parallel Populations.** In *Parallel Problem Solving from Nature - Proceedings of 1st Workshop, (PPSN 1), Volume 496 of Lecture Notes in Computer Science* Edited by: Schwefel HP, Määner R. Berlin, Germany: Springer-Verlag; 1991:198-208.
 90. Foe VE, Alberts BM: **Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in Drosophila embryogenesis.** *J Cell Sci* 1983, **61**:31-70.
 91. Runarsson TP, Yao X: **Search biases in constrained evolutionary optimization.** *IEEE Transactions on Systems, Man, and Cybernetics, Part C* 2005, **35(2)**:233-243.
 92. Lewis RM, Shepherd A, Torczon V: **Implementing generating set search methods for linearly constrained minimization.** In *Tech Rep WM-CS-2005-01* Department of Computer Science, College of William & Mary; 2005. [Revised July 2006]
 93. Holland JH: **Genetic Algorithms.** *Sci Am* 1992, **267**:66-72.
 94. Gibson M, Bruck J: **Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels.** *Journal of Physical Chemistry A* 2000, **104(9)**:1876-1889.
 95. Kosman D, Small S, Reinitz J: **Rapid preparation of a panel of polyclonal antibodies to Drosophila segmentation proteins.** *Dev Genes Evol* 1998, **208**:290-294.
 96. Myasnikova E, Kosman D, Reinitz J, Samsonova M: **Spatio-Temporal Registration of the Expression Patterns of Drosophila Segmentation Genes.** In *Proceedings of the Seventh International Conference on Intelligent Systems for Molecular Biology AAAI Press*; 1999:195-201.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

