BMC Bioinformatics

A computational pipeline for identifying kinetic motifs to aid in the design and improvement of synthetic gene circuits

Austin WT Chiang[1,2,3], Ming-Jing Hwang[1,2,3]*

Abstract

Background: An increasing number of genetic components are available in several depositories of such components to facilitate synthetic biology research, but picking out those that will allow a designed circuit to achieve the specified function still requires multiple cycles of testing. Here, we addressed this problem by developing a computational pipeline to mathematically simulate a gene circuit for a comprehensive range and combination of the kinetic parameters of the biological components that constitute the gene circuit.

Results: We showed that, using a well-studied transcriptional repression cascade as an example, the sets of kinetic parameters that could produce the specified system dynamics of the gene circuit formed clusters of recurrent combinations, referred to as kinetic motifs, which appear to be associated with both the specific topology and specified dynamics of the circuit. Furthermore, the use of the resulting "handbook" of performance-ranked kinetic motifs in finding suitable circuit components was illustrated in two application scenarios.

Conclusions: These results show that the computational pipeline developed here can provide a rational-based guide to aid in the design and improvement of synthetic gene circuits.

Background

The goal of synthetic biology is to be able to engineer biological processes and to select and put together standardized components according to a design and user-specified function and dynamics [1-5]. To this end, several depositories of biological components have been established [6-12], but the design cycles still rely very heavily on the slow and error-prone process of trying out the parts [13,14].

A fundamental problem is that we still lack clear knowledge of the factors that govern the dynamic behaviours of even the very simple circuits that are motifs of large biological networks [13-18]. Although the structure, or topology, of a biological network may largely

dictate its dynamics [19-21], the kinetic parameters (e.g. those that indicate the level of efficiency) of the involved biochemical reactions also play a role [22-27]. Thus, when performing a mathematical simulation to determine which biological components from the depositories should be chosen for a synthetically designed biological circuit with a specified dynamics, one needs to consider not only network topology, but also kinetic parameters.

In this work, we developed a computational pipeline, called Kinetic Motif and Functional Analysis (KMFA), to address this problem. By identifying the set of kinetic parameters required to produce the user-specified dynamic behaviour for a given network topology in both the presence and absence of random perturbations and using statistical analysis to identify recurrent patterns (i.e. motifs) of these kinetic parameters and understand their mechanics, KMFA provides a "handbook" of kinetic motifs in which one can look up a biological components

* Correspondence: mjhwang@ibms.sinica.edu.tw
[1]Bioinformatics Program, Taiwan International Graduate Program, Institute of Information Science, Academia Sinica, Taipei, Taiwan
Full list of author information is available at the end of the article

library to choose suitable parts for optimal performance of the designed circuit.

We first demonstrated the utility of KMFA by applying it to a well known synthetic gene circuit of a looped cascade of transcriptional inhibitions built in *Escherichia coli* [28-31]. We showed that, for this 4-gene circuit, only 2,355 (0.6%) of the 390,625 (5^8) possible combinations of the kinetic parameters could produce the prescribed steady-state concentration of each gene product as the output of the circuit under both perturbed and unperturbed conditions and that these kinetic solutions formed clusters of motifs, which could be ranked according to their relative performance, thus yielding a "handbook" of performance-ranked kinetic motifs that can be used to select library components for the circuit. We illustrated this using two scenarios: the first was to identify faulty/sub-optimal components and replace them with suitable ones to make a non-functional circuit functional, while the second was to improve the performance of an already functional circuit. KMFA is therefore a useful computational tool with a rational design capability to choose circuit components in synthetic biology research.

Methods

As shown in Figure 1, the KMFA pipeline consists of four steps. Although the concept and procedures of KMFA are generally applicable, we used a specific gene circuit to describe these steps and, later, in the Results, illustrate its utilities. This circuit was a four-gene transcriptional cascade that has been synthesized in *E coli* and has been studied both experimentally [28] and computationally [29-31].

Step 1: System setup

In this step, the system is set up to be numerically simulated in Step 2. This includes specifying the topology of the gene circuit or network, the mathematical model of the system, the target output dynamics and performance measures. Many studies have shown that deterministic modeling and stochastic modeling of biological systems have complementary benefits, the deterministic modeling providing the mean qualitative dynamics of a gene circuit and stochastic modeling the effects of random noises on the circuit [32,33]. In this study, we modeled the dynamics of the transcriptional cascade using both deterministic and stochastic simulations, the latter being used to find kinetic solutions for a robust system that could function properly under perturbations. The mathematical equations for the deterministic simulations are [29]:

$$
\begin{aligned}
\dot{x}_i(t) \quad &= f(p_i, d_i) \\
&= p_{i,0} + p_i r_i(x_j(t)) - d_i x_i(t) \\
r_i(x_j, t) &= \frac{1}{1 + \left(x_j(t)\big/1000\right)^2}
\end{aligned}
\tag{1}
$$

where $x_i(t)$ denotes the concentration of protein i and $\dot{x}_i(t)$ its rate of change at time t; $f(p_i, d_i)$ denotes the non-linear gene regulation of transcription and translation of protein i; $p_{i,0}$ is the basal production rate, p_i the production rate constant and d_i the degradation rate constant of protein i; and $r_i(x_j(t))$ denotes the Hill function of regulator protein j repressing the production of protein i at time t.

The mathematical equations for the stochastic simulations are [31]:

$$
\begin{aligned}
\dot{y}_i(t) &= g(p_i, d_i) \\
&= f(p_i, d_i) + \Delta f(p_i, d_i)n_i(t) + v_i(t) \\
&= (p_{i,0} + p_i r_i(y_j(t)) - d_i y_i(t)) \\
&\quad + (\Delta p_{i,0} + \Delta p_i r_i(y_j(t)) - \Delta d_i y_i(t))n_i(t) + v_i(t)
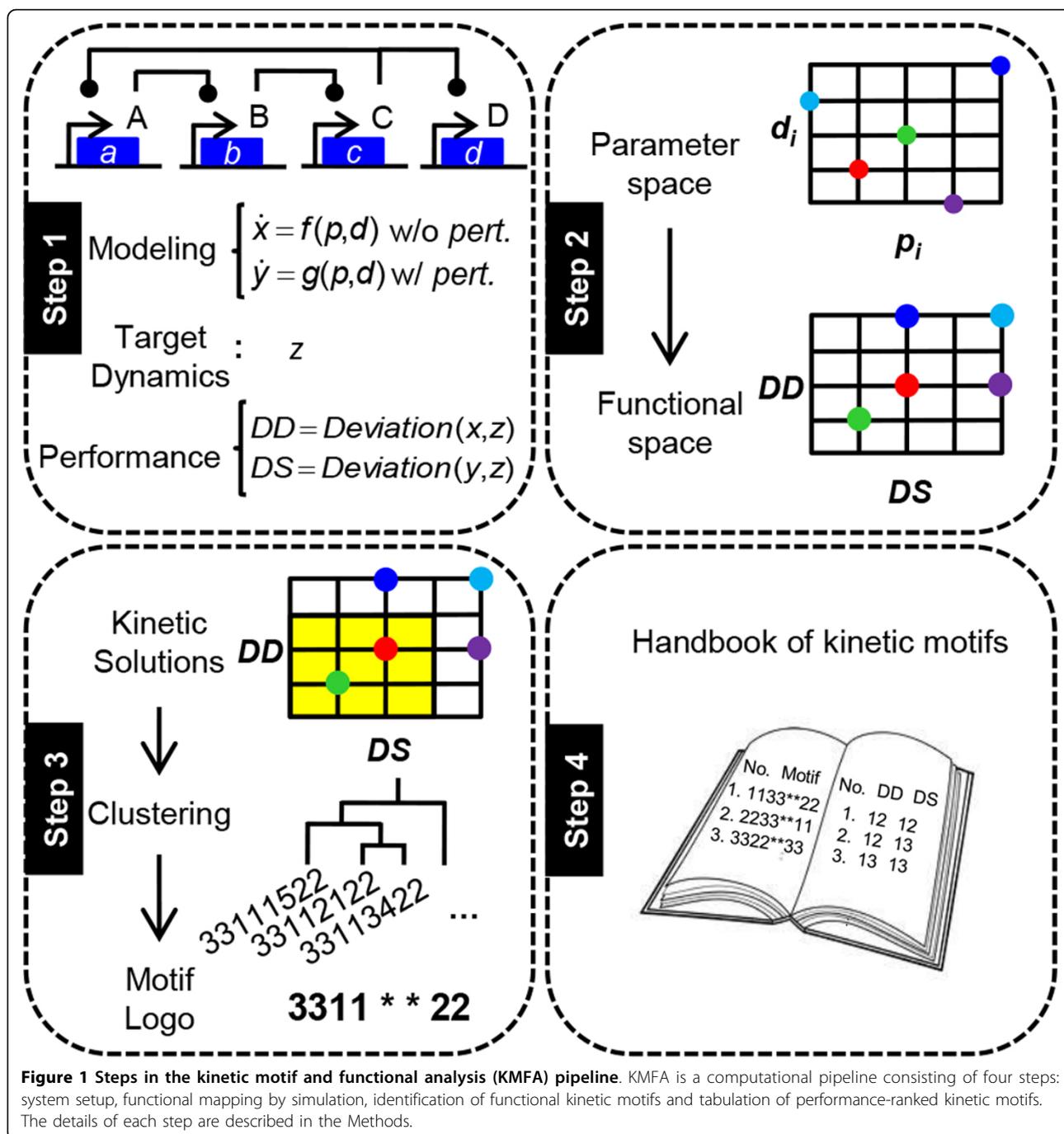\end{aligned}
\tag{2}
$$

where $y_i(t)$ denotes the concentration of protein i and $\dot{y}_i(t)$ its rate of change at time t; $g(p_i, d_i)$ denotes the non-linear gene regulation of transcription and translation of protein i influenced by both intrinsic and external random noises; $\Delta p_{i,0}$, Δp_i and Δd_i denote the respective standard deviations for the parameters of basal production rate, production rate and degradation rate; and n_i and v_i are, respectively, randomly generated intrinsic and extrinsic noises for protein i.

This cascade of transcriptional inhibitions has been shown to produce steady-state concentrations of the proteins involved [28-31]. To facilitate comparison, the target dynamics (z) were chosen to be the steady-state concentrations of protein TetR, LacI, CI and Eyfp used in Chen and Wu [31], giving z = {1000, 30000, 300, 30000 nM}. To measure the performance of a given set of kinetic parameters (p_i, d_i), i = {TetR, LacI, CI, Eyfp}, deviations from the steady-state concentrations of the target were computed for the deterministic simulations using Eq. (3) and for the stochastic simulations using Eq. (4), where DD denotes the deterministic deviation, DS the stochastic deviation and ln a logarithmic transformation function.

$$
DD = Deviation(x, z) = ln\left(\sum_i \sum_t \left|x_i(t) - z_i(t)\right|\right)
\tag{3}
$$

$$
DS = Deviation(y, z) = ln\left(\sum_i \sum_t \left|y_i(t) - z_i(t)\right|\right)
\tag{4}
$$

In order to compare with the numerical results of [31], the two measures were not normalized even though z varies from 300 to 30000 nM. However, by dictating that a viable dynamics must not deviate from the target concentration by more than 20% for each of the four proteins (see below), we reduced the possibility of the contribution from CI, which has the smallest target concentration at 300 nM, being overtaken by those of the other proteins.

Figure 1 Steps in the kinetic motif and functional analysis (KMFA) pipeline. KMFA is a computational pipeline consisting of four steps: system setup, functional mapping by simulation, identification of functional kinetic motifs and tabulation of performance-ranked kinetic motifs. The details of each step are described in the Methods.

Step 2: Simulation

For our simulations, we employed the same range of parameter values and values for the basal production rate constants, initial protein concentrations and perturbation fluctuations used by Chen and Wu [31] (summarized in Additional file 1: Table S1). Because it is impossible to enumerate all the real values of the kinetic parameters, we uniformly binned each kinetic parameter into five different rate efficiency levels, with level 1 denoting the lowest and level 5 the highest strength (efficiency) of the kinetics, and used the mean of each bin as the representative of the bin (Additional file 1: Table S2) for the simulations. The five levels may respectively correspond to "weakest", "weak", "medium", "strong", and "strongest" categories that are often used in experimental studies to characterize, say, binding efficiency of ribosome binding sites [7]. Increasing the number of levels will increase the resolution on the

transformation of parameter values but will also increase the computational cost and difficulties to identify kinetic motifs in subsequent analysis (see below). Since there are four genes in the system and each gene product (protein) is associated with two kinetic parameters, one for production and the other for degradation, there were a total of 390,625 (5^8) sets of kinetic parameters that could be used for simulations. For each of these parameter sets, Eq. (1) (deterministic) and Eq. (2) (stochastic) were simulated for a maximum of 100 time steps and differences in protein concentrations between the simulation and prescribed values (z) (Eq. (3) and Eq (4)) calculated. At any time point, if the concentration of any of the four proteins exceeded 10^6 nM, the simulation (Eq. (1) or Eq. (2)) was aborted prematurely and an extremely large deviation value of 15 was assigned. The other ill-behaved parameter sets, for which a value of 15 was assigned to their DD (Eq. (3)) or DS (Eq. (4)), were those in which at least one of the four proteins exhibited a steady-state concentration, computed as the mean for the second half of the simulation, that deviated by more than 20% from the prescribed value. Simulations with a deviation value of 15 were considered non-functional.

Step 3: Identification of kinetic motifs

All the sets of kinetic parameters that were not ill-behaved (i.e. both their DD and DS values were smaller than 15) were transformed into bin integers representing efficiency levels (see above), which were then clustered hierarchically [34] based on the similarity, computed by the Hamming distance [35], of the sequence of bin levels. This resulted in clusters of kinetic parameters, and each cluster, called a kinetic motif, could be represented by a consensus sequence logo, such as {(1, 1); (3, 3); (*, *); (3, 3)}, in which bin levels in integers of the two kinetic parameters for the production and degradation of each of the four proteins are paired in parentheses and * indicates any levels.

Step 4: Generation of a handbook of kinetic motifs

These motifs could be ranked based on their performance, i.e. how well they could produce the specified system dynamics, as measured by the deviations DD (Eq. (3)) and DS (Eq. (4)). This resulted in a table, or handbook, of performance-ranked kinetic parameter sets, which provides information that can be easily referred to in order to identify suitable genetic components to assemble a functional circuit that meets user-desired specifications.

Results

Mapping parameters to solutions

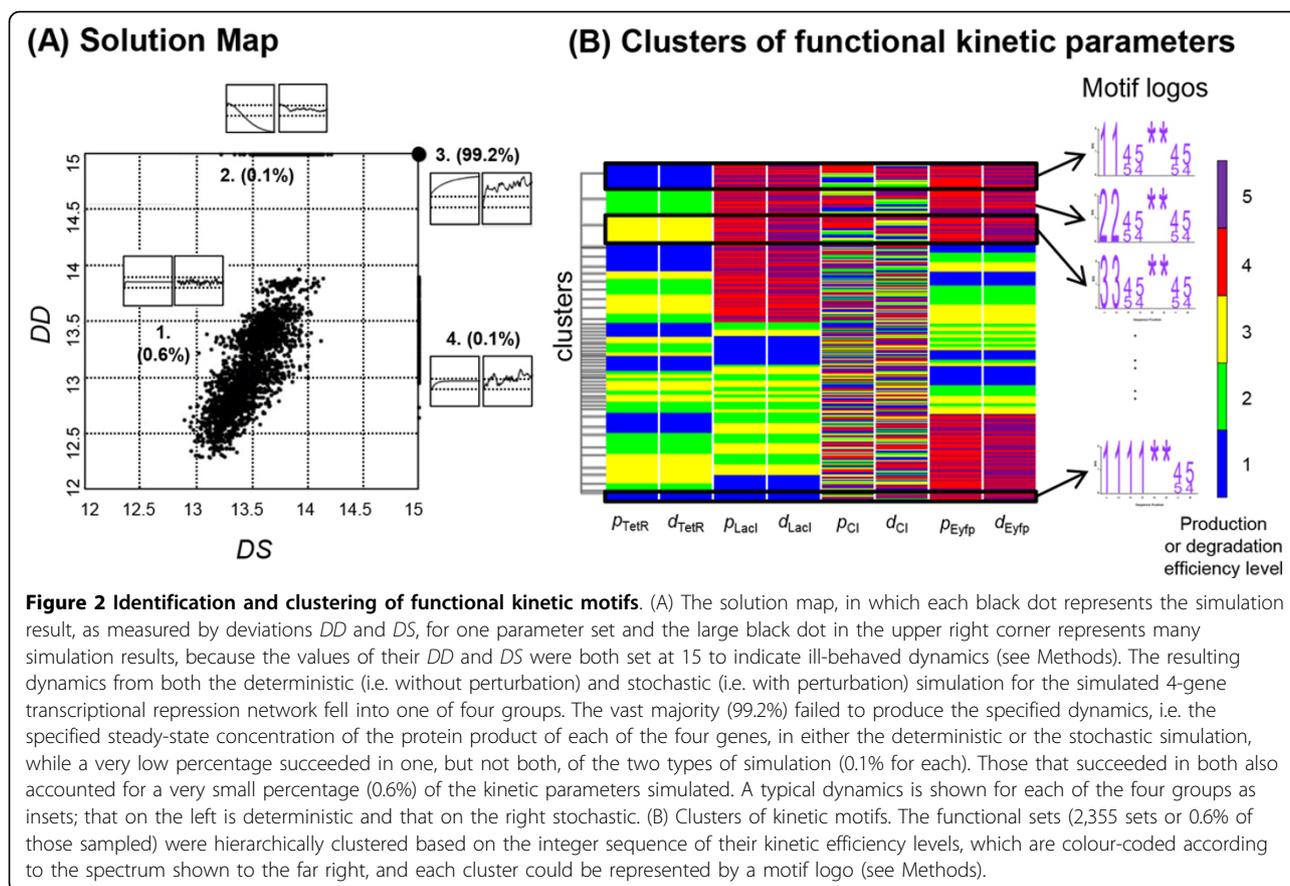As described in the Methods, we comprehensively searched the parameter space, albeit using representatives of uniformly divided bins, for sets of kinetics parameters that could produce the specified steady-state system dynamics under both conditions of with and without perturbations. Figure 2A shows that only a very small fraction (2,355 or 0.6%) of the parameter sets sampled (5^8 = 390,625) could produce the specified dynamics, while the vast majority (387,535 or 99.2%) failed to function properly under either condition. Interestingly, some parameter sets (0.1%) functioned well only without perturbations, while others (0.1%) did so only with perturbations.

Clusters of kinetic motifs

Each of the 2,355 kinetic solutions obtained in Figure 2A is a combination of 8 kinetic parameters consisting of the pair of kinetic parameters p, for the production rate, and d, for the degradation rate of each of the four proteins. The resulting 2,355 8-integer sequences were compared and grouped into 52 clusters (Figure 2B), 48 of which had at least 10 members (i.e. 10 parameter sets that could produce the specified steady-state dynamics). Similar to sequences of DNA [36] or amino acids [37-39], these clusters of kinetic parameter sets, or kinetic motifs, were represented by a motif logo made up of efficiency level integers in which the character size reflects the extent of consensus within the cluster, while '*' indicates no consensus at all (Figure 2B). These 48 clusters were tabulated in a handbook of kinetic motifs (Additional file 1: Table S3), along with their circuit structure, kinetic motif logo, number of cluster members and performance scores.

Functional association of kinetic motifs

From Additional file 1: Table S3, in which the clusters of kinetic motifs are ranked by their performance in decreasing value of the sum of their DD and DS scores, we observed that, on the whole, the top clusters tended to have high, though not necessarily the highest, production and degradation rates, i.e. there were more 3's or 4's and 5's than 1's and 2's in the motif logos, and that, conversely, levels 1 and 2 occurred more frequently in the motif logos for the clusters ranked at the bottom. This was particularly true for the 2nd (LacI) and 4th (Eyfp) proteins in the network, because they were required to have a high steady-state output concentration of 30,000 nM, compared to 1,000 nM for the 1st protein (TetR) and 300 nM for the 3rd (CI) (see Methods). Furthermore, for the same protein, its production and degradation rates seemed to be symmetry-related, i.e. level 3 degradation tended to be paired with level 3 production. Such symmetry is also visible in Figure 2B. These observations are consistent with our understanding of the dynamic behaviours of a transcription regulation system in that, in general, degradation rates determine the response time and fast degradation rates allow rapid changes in protein

Figure 2 Identification and clustering of functional kinetic motifs. (A) The solution map, in which each black dot represents the simulation result, as measured by deviations *DD* and *DS*, for one parameter set and the large black dot in the upper right corner represents many simulation results, because the values of their *DD* and *DS* were both set at 15 to indicate ill-behaved dynamics (see Methods). The resulting dynamics from both the deterministic (i.e. without perturbation) and stochastic (i.e. with perturbation) simulation for the simulated 4-gene transcriptional repression network fell into one of four groups. The vast majority (99.2%) failed to produce the specified dynamics, i.e. the specified steady-state concentration of the protein product of each of the four genes, in either the deterministic or the stochastic simulation, while a very low percentage succeeded in one, but not both, of the two types of simulation (0.1% for each). Those that succeeded in both also accounted for a very small percentage (0.6%) of the kinetic parameters simulated. A typical dynamics is shown for each of the four groups as insets; that on the left is deterministic and that on the right stochastic. (B) Clusters of kinetic motifs. The functional sets (2,355 sets or 0.6% of those sampled) were hierarchically clustered based on the integer sequence of their kinetic efficiency levels, which are colour-coded according to the spectrum shown to the far right, and each cluster could be represented by a motif logo (see Methods).

concentration and can, therefore, minimize the response time to stimulation [40,41]. Furthermore, to maintain a given steady state concentration, proteins with a rapid degradation rate also require a high production rate, the final concentration being determined by both rates [16,42]. Inspection of individual solutions with the highest degradation rate (level 5) showed that they tended to exhibit high fluctuations in protein concentrations before reaching the final steady-state concentrations, suggesting that the highest degradation rate (level 5) may not always be the most desirable for the designed circuit, which may also explain why the symmetry between protein production and degradation tended to break down a little when levels were very high (4 or 5; Additional file 1: Table S3).
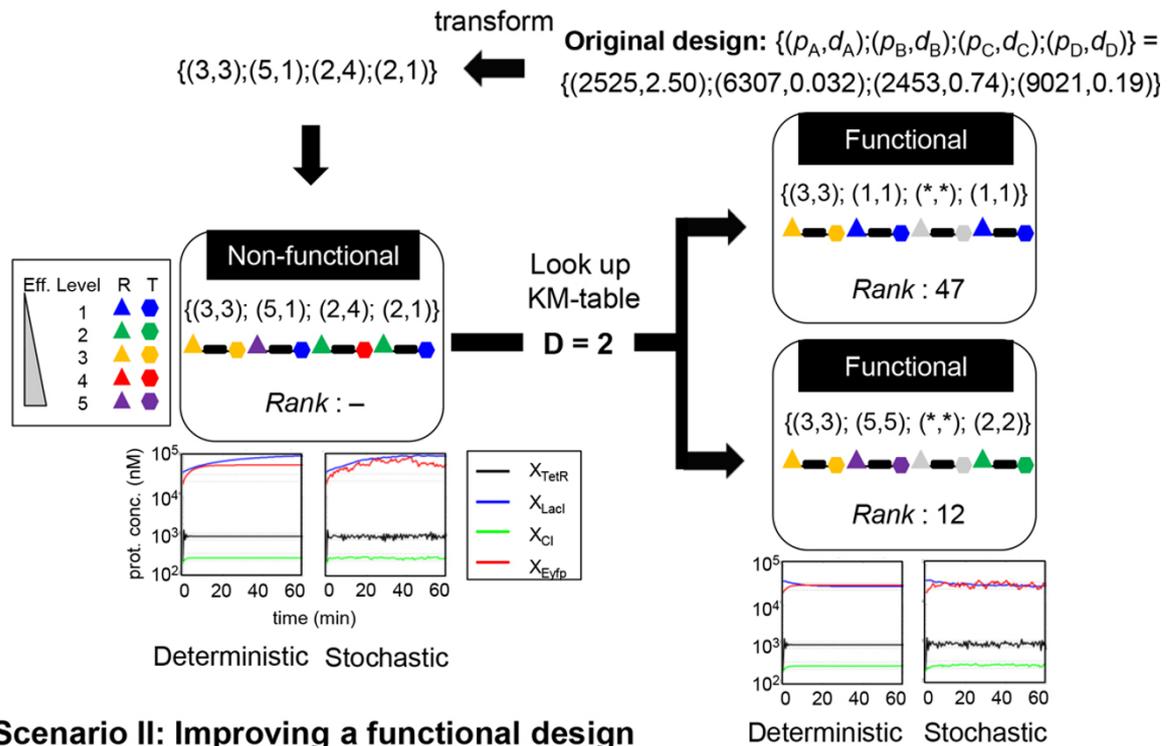
Another interesting observation was that the third protein (CI) did not show a preference for a particular efficiency level for either of its two kinetic parameters (Figure 2B and Additional file 1: Table S3). As shown in Additional file 1: Fig. S1, this can be explained by examining the repression network and its rate equations. Briefly, the high initial and steady-state concentrations of protein LacI, which represses the *cI* gene, renders the contribution of gene regulation to CI production negligible and, thus, the magnitude of its production kinetic parameter, p_{CI}, inconsequential; also, as a consequence,

at the steady state, all the values within the range allowed for its degradation kinetic parameter, d_{CI}, would lead to a CI concentration within 20% of the specified value, thus meeting the required condition for a functional solution (see Methods).

Application scenario I: making a non-functional design functional

Given a non-functional circuit, it is, at present, not easy to determine the reasons for the failure to function or to identify the faulty/sub-optimal component that needs to be replaced or corrected to salvage the design. The handbook of kinetic motifs and associated properties (Additional file 1: Table S3) is a great aid to solving this problem, as illustrated by the following example. As shown in Figure 3A, the original network was composed of components with kinetic parameters {(p_{TetR}, d_{TetR}); (p_{LacI}, d_{LacI}); (p_{CI}, d_{CI}); (p_{Eyfp}, d_{Eyfp})} with values equivalent to efficiency levels of {(3, 3); (5, 1); (2, 4); (2, 1)} that could not produce the specified dynamics for the 2nd (LacI) and 4th (Eytp) proteins. Comparing this sequence of kinetic levels with those tabulated in Additional file 1: Table S3, we found two motifs with a sequence that differed at only two of the 8 positions (i.e. Hamming distance = 2), namely motif No. 47 [{(3, 3);
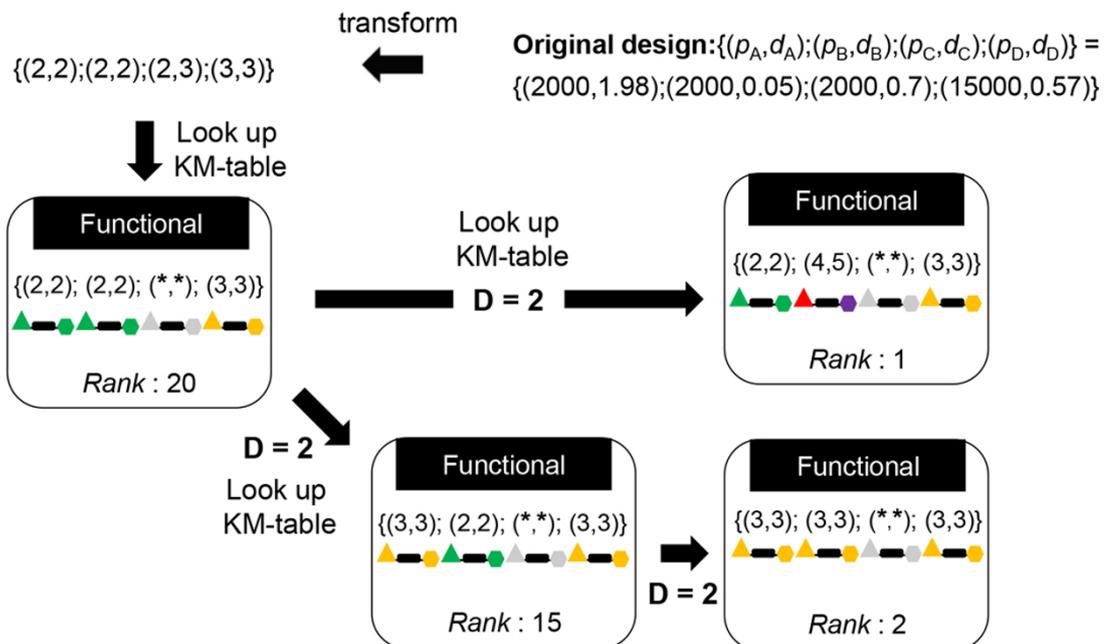
Figure 3 Two scenarios demonstrating the application of KMFA. (A) Scenario I: Making a non-functional design functional. If we have a non-functional original design, as indicated by the dynamics of the deterministic and stochastic simulation shown, we can transform its kinetic parameters into efficiency levels and look in the handbook of kinetic motifs (Additional file 1: Table S3) for a similar, but functional, design to replace some of the ribosome binding sites (RBS, denoted as R and represented by triangles) and/or protein degradation tags (PDT, denoted as T and represented by hexagons) with others with a suitable efficiency level to render the modified design functional. In this example, two different routes of minimal changes (Hamming distance (D) of replacement = 2) were found. (B) Scenario II: Improving a functional design. In this example, the original, already functional, design is that reported by Chen and Wu [31], the ranking of which in our performance table (Additional file 1: Table S3) can be improved by selecting components with an RBS and/or PTD that can work with other components to generate a better system performance (i.e. better performance ranking in Additional file 1: Table S3). All notations and symbols are the same as those in (A).

(1, 1); (*, *); (1, 1)}] and motif No. 12 [{(3, 3); (5, 5); (*, *); (2, 2)}]. Note that, in this comparison, we ignored the 3rd protein (CI) because, as mentioned above, it does not have a preference for these kinetic levels. To modify the original design into one of these two motifs, we can change the efficiency level of the kinetic parameters by using a different ribosome binding site (RBS) for one or more of the genes and thereby changing their protein production level or by using a different protein degradation tag (PDT) to alter the protein degradation rate. Specifically, we can either replace the RBS in the original design with one with a protein production efficiency level of 1 for the 2nd and 4th proteins to change the original design into one that would behave like motif No. 47 or we can replace the PDT in the original design with one with protein degradation efficiency levels of 5 for the 2nd protein and 2 for the 4th protein, making it resemble motif No. 12. Both routes of modification would render the system functional, although the second would lead to a system with a better performance. Note that the key rule in making these changes is to maintain symmetry between the protein production and degradation rates for the same protein, as discussed above.

Application scenario II: improving a functional design
In the second scenario (Figure 3B), the original circuit was already functional and, in fact, has been shown to be robust under perturbations by Chen and Wu [31]. However, it was ranked in the middle part of Additional file 1: Table S3 (motif No. 20) and therefore could still be improved. There is more than one way of improving the performance of the original design. The two proposed routes of change illustrated in Figure 3B would both lead to a superior performance ranking in the "handbook" of this circuit: the first involves increasing the efficiency level for both the RBS and PDT for the 2nd protein (LacI), while the second requires the two steps of replacing the RBS and PDT for the 1st protein (TetR), then those for the 2nd protein (LacI).

Discussion
Mathematical modelling and simulations are being increasingly used to help design simple biological circuits to achieve user-specified functions or certain network behaviours in synthetic biology [5,43-47]. Much work has been focused on finding suitable network topologies [19-21], or kinetic parameters that can produce the desired system dynamics, given a network topology [22-26,29-31]. In this work, we have developed a computational pipeline, KMFA, for the latter. Compared to previous studies, a distinctive feature of KMFA is that, by comprehensively mapping the parameter space and functional space (Figure 1, Step 2), the identified multiple sets of functional kinetic parameters can be clustered and

their motifs (i.e. their recurring combinations) identified (Figure 2B) and analysed to reveal and understand significant functional associations (see Results, Figure 2).

In this work, the network was simulated with and without perturbations, the former being modeled by a stochastic process with randomly introduced intrinsic and extrinsic noises (Eq. (2)) and the latter by a deterministic process (Eq. (1)). The results showed that there are certain situations (i.e. certain combinations of the kinetic parameters) in which the system may function well under perturbations, but not in the absence of perturbations. Although it is well recognized that noise is an integral part of normal biological functions and how biological systems evolve to be robust [48-53], our results suggest that, to design a 'truly robust' biological system, both conditions (with and without noise) need to be considered. Although further analysis is required to elucidate the mechanism for the effect of perturbation observed in this study, we speculate that perturbation alters the "background" protein production and degradation rates to allow the circuit to achieve and maintain target dynamics. Indeed, several studies have shown that gene circuits (e.g., switches and oscillations) rely on noise to achieve desired functions and would fail without them [54-57].

As a proof-of-principle study, KMFA has been designed in the present work to exhaust all possible kinetic parameter combinations, which limits its application to very small systems such as the transcriptional cascade serving as a demonstrating example above. However, this limitation can be lifted by integrating with approaches such as Monte-Carlo sampling [58], Latin hypercube sampling [59], and others [29,60] that have been developed to efficiently search for parameter solutions from high dimensional spaces. Indeed, several large systems (e.g. the Fas apoptotic pathway [61]) have been studied using some of these methods. Replacing our time-consuming simulation step (see Figure 1) with these efficient parameter-searching methods, one can collect the parameter solutions for an interesting function/dynamics of a large system, and then apply KMFA on those parameter solutions to find kinetic motifs.

An important finding of this study is that, as in studies of network topologies [19-21], only a very small percentage of kinetic parameters can be functional, i.e. produce the desired system dynamics (Figure 2A). Obviously, network topology and kinetic parameters are intertwined, and the kinetic motifs (Additional file 1: Table S3) identified here will probably be specific not only for the specific arrangement of the components of the circuit, but also for the specified output protein concentrations. Nevertheless, for a given circuit design, the "handbook" of kinetic motifs (Additional file 1: Table S3) will be very useful in deciding which biological components available from a biological

parts library, such as BioFab [10] and the MIT Registry of Standard Biological Parts [7], should be used, as illustrated by the two application scenarios presented in Figure 3. Thus, the handbook could be used to standardize the biological parts (e.g. by classifying them into 5 efficiency levels) and, via ranking all possible combinations of the parts, standardize the circuit design process. The handbook could furthermore be used to help synthesize required parts of a specified efficiency level not yet collected in the library, especially the ribosome binding sites for which a relationship between binding efficiency and binding site sequence has been elucidated [62]. As more and more genetic components (e.g. promoters, ribosome binding sites, terminators and protein degradation tags) are deposited and characterized, mathematical simulations using a computational pipeline, such as KMFA, will make the synthetic design and engineering of biological circuits more efficient and also more rationale-based. Furthermore, the general applicability of the KMFA approach will allow simulations of different network topologies and different functions (see Additional file 1 for additional illustration on an 'AND'-gate circuit [63,64]) in comparative studies to uncover intriguing common/distinct principles of biological networks, particularly those pertaining to kinetic parameters, which have received less attention than network topologies.

Conclusions

Knowledge about how to choose suitable components for a designed gene circuit is required for efficient research in synthetic biology. For a given topology of a designed gene circuit, the computational pipeline, KMFA, developed here, has produced a "handbook" of performance-ranked kinetic motifs that can serve as a user guide to allow the selection and matching up of different genetic components to achieve user-specified system functions. In addition to being a useful aid for the synthetic design of biological circuits, KMFA can also be used to elucidate the intertwined relationship of the trinity of prototype biological circuits: topology, kinetics of the parts and function.

Additional material

Additional file 1: Table S1: Ranges and values of parameters used for simulations in this study. **Table S2**: RBSs and PDTs of different efficiency levels used in this study for the structural arrangement RBS-gene-PDT. **Table S3**: Handbook of kinetic motifs for the transcriptional repression cascade circuit simulated. **Figure S1**: Analytical analysis of the kinetic parameters of protein CI. **Additional illustration**: The case of an 'AND'-gate circuit.

Authors' details
[1]Bioinformatics Program, Taiwan International Graduate Program, Institute of Information Science, Academia Sinica, Taipei, Taiwan. [2]Institute of BioMedical Informatics, National Yang-Ming University, Taipei, Taiwan. [3]Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan.

References
1. Endy D: **Foundations for engineering biology.** *Nature* 2005, **438(7067)**:449-453.
2. Guido NJ, Wang X, Adalsteinsson D, McMillen D, Hasty J, Cantor CR, Elston TC, Collins JJ: **A bottom-up approach to gene regulation.** *Nature* 2006, **439(7078)**:856-860.
3. Canton B, Labno A, Endy D: **Refinement and standardization of synthetic biological parts and devices.** *Nat Biotechnol* 2008, **26(7)**:787-793.
4. Ellis T, Wang X, Collins JJ: **Diversity-based, model-guided construction of synthetic gene networks with predicted functions.** *Nat Biotechnol* 2009, **27(5)**:465-471.
5. Slusarczyk AL, Lin A, Weiss R: **Foundations for the design and implementation of synthetic genetic circuits.** *Nat Rev Genet* 2012, **13(6)**:406-420.
6. Edelman GM, Meech R, Owens GC, Jones FS: **Synthetic promoter elements obtained by nucleotide sequence variation and selection for activity.** *Proc Natl Acad Sci USA* 2000, **97(7)**:3038-3043.
7. MIT Registry of Standard Biological Parts. [http://partsregistry.org].
8. BioBricks Foundation. [http://bbf.openwetware.org/].
9. Murphy KF, Balazsi G, Collins JJ: **Combinatorial promoter design for engineering noisy gene expression.** *Proc Natl Acad Sci USA* 2007, **104(31)**:12726-12731.
10. BioFab: International Open Facility Advancing Biotechnology. [http://www.biofab.org/].
11. Constante M, Grunberg R, Isalan M: **A biobrick library for cloning custom eukaryotic plasmids.** *PloS one* 2011, **6(8)**:e23685.
12. Galdzicki M, Rodriguez C, Chandran D, Sauro HM, Gennari JH: **Standard biological parts knowledgebase.** *PloS one* 2011, **6(2)**:e17005.
13. Lu TK, Khalil AS, Collins JJ: **Next-generation synthetic gene networks.** *Nat Biotechnol* 2009, **27(12)**:1139-1150.
14. Porcar M, Danchin A, de Lorenzo V, Dos Santos VA, Krasnogor N, Rasmussen S, Moya A: **The ten grand challenges of synthetic life.** *Syst Synth Biol* 2011, **5**:(1-2):1-9.
15. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U: **Network motifs: simple building blocks of complex networks.** *Science* 2002, **298(5594)**:824-827.
16. Alon U: **An introduction to systems biology: design principles of biological circuits.** Boca Raton, FL: Chapman & Hall/CRC; 2007.
17. Shinar G, Feinberg M: **Structural sources of robustness in biochemical reaction networks.** *Science* 2010, **327(5971)**:1389-1391.
18. Tyson JJ, Novak B: **Functional motifs in biochemical reaction networks.** *Annu Rev Phys Chem* 2010, **61**:219-240.
19. Brandman O, Meyer T: **Feedback loops shape cellular signals in space and time.** *Science* 2008, **322(5900)**:390-395.

20. Ma W, Trusina A, El-Samad H, Lim WA, Tang C: **Defining network topologies that can achieve biochemical adaptation.** *Cell* 2009, **138**(4):760-773.
21. Cotterell J, Sharpe J: **An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients.** *Mol Syst Biol* 2010, **6**:425.
22. de Jong H: **Modeling and simulation of genetic regulatory systems: a literature review.** *Journal of computational biology: a journal of computational molecular cell biology* 2002, **9**(1):67-103.
23. Li ZF, Osborne MR, Prvan T: **Parameter estimation of ordinary differential equations.** *Ima J Numer Anal* 2005, **25**(2):264-285.
24. Ingram PJ, Stumpf MP, Stark J: **Network motifs: structure does not determine function.** *BMC genomics* 2006, **7**:108.
25. Szallasi Z, Stelling J, Periwal V: **System modeling in cell biology: from concepts to nuts and bolts.** Cambridge, Mass.: MIT Press; 2006.
26. Chou IC, Voit EO: **Recent developments in parameter estimation and structure identification of biochemical and genomic systems.** *Mathematical biosciences* 2009, **219**(2):57-83.
27. Palsson B: **Systems biology: simulation of dynamic network states.** Cambridge, UK; New York: Cambridge University Press; 2011.
28. Hooshangi S, Thiberge S, Weiss R: **Ultrasensitivity and noise propagation in a synthetic transcriptional cascade.** *Proc Natl Acad Sci USA* 2005, **102**(10):3581-3586.
29. Batt G, Yordanov B, Weiss R, Belta C: **Robustness analysis and tuning of synthetic gene networks.** *Bioinformatics* 2007, **23**(18):2415-2422.
30. Chen BS, Chang CH, Lee HC: **Robust synthetic biology design: stochastic game theory approach.** *Bioinformatics* 2009, **25**(14):1822-1830.
31. Chen BS, Wu CH: **A systematic design method for robust synthetic biology to satisfy design specifications.** *BMC Syst Biol* 2009, **3**:66.
32. Scott M, Hwa T, Ingalls B: **Deterministic characterization of stochastic genetic circuits.** *Proc Natl Acad Sci USA* 2007, **104**(18):7402-7407.
33. Twycross J, Band LR, Bennett MJ, King JR, Krasnogor N: **Stochastic and deterministic multiscale models for systems biology: an auxin-transport case study.** *BMC Syst Biol* 2010, **4**:34.
34. Eisen MB, Spellman PT, Brown PO, Botstein D: **Cluster analysis and display of genome-wide expression patterns.** *Proceedings of the National Academy of Sciences of the United States of America* 1998, **95**(25):14863-14868.
35. Hamming RW: **Error Detecting and Error Correcting Codes.** *At&T Tech J* 1950, **29**(2):147-160.
36. D'haeseleer P: **How does DNA sequence motif discovery work?** *Nat Biotechnol* 2006, **24**(8):959-961.
37. Schneider TD, Stephens RM: **Sequence logos: a new way to display consensus sequences.** *Nucleic Acids Res* 1990, **18**(20):6097-6100.
38. Bork P, Koonin EV: **Protein sequence motifs.** *Curr Opin Struct Biol* 1996, **6**(3):366-376.
39. Bailey TL, Williams N, Misleh C, Li WW: **MEME: discovering and analyzing DNA and protein sequence motifs.** *Nucleic Acids Res* 2006, **34**(Web Server):W369-373.
40. Rosenfeld N, Alon U: **Response delays and the structure of transcription networks.** *J Mol Biol* 2003, **329**(4):645-654.
41. Sayut DJ, Sun L: **Slow activator degradation reduces the robustness of a coupled feedback loop oscillator.** *Mol Biosyst* 2010, **6**(8):1469-1474.
42. Schimke RT, Doyle D: **Control of enzyme levels in animal tissues.** *Annu Rev Biochem* 1970, **39**:929-976.
43. Endy D, Brent R: **Modelling cellular behaviour.** *Nature* 2001, **409**(6818):391-395.
44. Kaznessis YN: **Models for synthetic biology.** *BMC systems biology* 2007, **1**:47.
45. Zheng Y, Sriram G: **Mathematical modeling: bridging the gap between concept and realization in synthetic biology.** *Journal of biomedicine & biotechnology* 2010, **2010**:541609.
46. Liang J, Luo Y, Zhao H: **Synthetic biology: putting synthesis into biology.** *Wiley interdisciplinary reviews Systems biology and medicine* 2011, **3**(1):7-20.
47. MacDonald JT, Barnes C, Kitney RI, Freemont PS, Stan GB: **Computational design approaches and tools for synthetic biology.** *Integrative biology: quantitative biosciences from nano to macro* 2011, **3**(2):97-108.
48. Gonze D, Halloy J, Goldbeter A: **Robustness of circadian rhythms with respect to molecular noise.** *Proc Natl Acad Sci USA* 2002, **99**(2):673-678.
49. Kitano H: **Biological robustness.** *Nat Rev Genet* 2004, **5**(11):826-837.
50. Ciliberti S, Martin OC, Wagner A: **Robustness can evolve gradually in complex regulatory gene networks with varying topology.** *PLoS Comput Biol* 2007, **3**(2):e15.
51. Lapidus S, Han B, Wang J: **Intrinsic noise, dissipation cost, and robustness of cellular networks: the underlying energy landscape of MAPK signal transduction.** *Proc Natl Acad Sci USA* 2008, **105**(16):6039-6044.
52. Lander AD: **Pattern, growth, and control.** *Cell* 2011, **144**(6):955-969.
53. Munsky B, Neuert G, van Oudenaarden A: **Using gene expression noise to understand gene regulation.** *Science* 2012, **336**(6078):183-187.
54. Hasty J, Pradines J, Dolnik M, Collins JJ: **Noise-based switches and amplifiers for gene expression.** *Proc Natl Acad Sci USA* 2000, **97**(5):2075-2080.
55. Paulsson J, Berg OG, Ehrenberg M: **Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation.** *Proc Natl Acad Sci USA* 2000, **97**(13):7148-7153.
56. Lewis J: **Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator.** *Curr Biol* 2003, **13**(16):1398-1408.
57. Stewart-Ornstein J, Weissman JS, El-Samad H: **Cellular noise regulons underlie fluctuations in Saccharomyces cerevisiae.** *Mol Cell* 2012, **45**(4):483-493.
58. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E: **Equation of State Calculations by Fast Computing Machines.** *J Chem Phys* 1953, **21**(6):1087-1092.
59. Iman RL, Helton JC, Campbell JE: **An Approach to Sensitivity Analysis of Computer-Models. 1. Introduction, Input Variable Selection and Preliminary Variable Assessment.** *J Qual Technol* 1981, **13**(3):174-183.
60. Zamora-Sillero E, Hafner M, Ibig A, Stelling J, Wagner A: **Efficient characterization of high-dimensional parameter spaces for systems biology.** *BMC systems biology* 2011, **5**.
61. Hua F, Hautaniemi S, Yokoo R, Lauffenburger DA: **Integrated mechanistic and data-driven modelling for multivariate analysis of signalling pathways.** *Journal of the Royal Society, Interface / the Royal Society* 2006, **3**(9):515-526.
62. Salis HM, Mirsky EA, Voigt CA: **Automated design of synthetic ribosome binding sites to control protein expression.** *Nature biotechnology* 2009, **27**(10):946-950.
63. Terzer M, Jovanovic A, Choutko A, Nikolayeva O, Korn A, Brockhoff D, Zurcher F, Friedmann M, Schutz R, Zitzler E, et al: **Design of a biological half adder.** *IET Synthetic Biology* 2007, **1**:(1-2):53-58.
64. Chen BS, Hsu CY, Liou JJ: **Robust design of biological circuits: evolutionary systems biology approach.** *Journal of biomedicine & biotechnology* 2011, **2011**:304236.