

# Emergence of Switch-Like Behavior in a Large Family of Simple Biochemical Networks

Dan Siegal-Gaskins<sup>1,2\*</sup>, Maria Katherine Mejia-Guerra<sup>2</sup>, Gregory D. Smith<sup>3</sup>, Erich Grotewold<sup>2</sup>

**1** Mathematical Biosciences Institute, The Ohio State University, Columbus, Ohio, United States of America, **2** Department of Molecular Genetics and Plant Biotechnology Center, The Ohio State University, Columbus, Ohio, United States of America, **3** Department of Applied Science, The College of William and Mary, Williamsburg, Virginia, United States of America

## Abstract

Bistability plays a central role in the gene regulatory networks (GRNs) controlling many essential biological functions, including cellular differentiation and cell cycle control. However, establishing the network topologies that can exhibit bistability remains a challenge, in part due to the exceedingly large variety of GRNs that exist for even a small number of components. We begin to address this problem by employing chemical reaction network theory in a comprehensive *in silico* survey to determine the capacity for bistability of more than 40,000 simple networks that can be formed by two transcription factor-coding genes and their associated proteins (assuming only the most elementary biochemical processes). We find that there exist reaction rate constants leading to bistability in ~90% of these GRN models, including several circuits that do not contain any of the TF cooperativity commonly associated with bistable systems, and the majority of which could only be identified as bistable through an original subnetwork-based analysis. A topological sorting of the two-gene family of networks based on the presence or absence of biochemical reactions reveals eleven minimal bistable networks (i.e., bistable networks that do not contain within them a smaller bistable subnetwork). The large number of previously unknown bistable network topologies suggests that the capacity for switch-like behavior in GRNs arises with relative ease and is not easily lost through network evolution. To highlight the relevance of the systematic application of CRNT to bistable network identification in real biological systems, we integrated publicly available protein-protein interaction, protein-DNA interaction, and gene expression data from *Saccharomyces cerevisiae*, and identified several GRNs predicted to behave in a bistable fashion.

**Citation:** Siegal-Gaskins D, Mejia-Guerra MK, Smith GD, Grotewold E (2011) Emergence of Switch-Like Behavior in a Large Family of Simple Biochemical Networks. PLoS Comput Biol 7(5): e1002039. doi:10.1371/journal.pcbi.1002039

**Editor:** Jorg Stelling, ETH Zurich, Switzerland

**Received:** November 23, 2010; **Accepted:** March 21, 2011; **Published:** May 12, 2011

**Copyright:** © 2011 Siegal-Gaskins et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by NRI Grant 2007-35318-17805 from the USDA CSREES, DOE Grant DE-FG02-07ER15881 and NSF grant DBI-0701405 to EG, NSF Grant DMS-0443843 to GDS, and an NIH T32 training grant from the Division of Human Cancer Genetics to DSG. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: dsg@mbi.osu.edu

## Introduction

Bistability—the coexistence of two stable equilibria in a dynamical system—is responsible for the switch-like behavior seen in a wide variety of cell biological networks, such as those involved in signal transduction [1], cell fate specification [2–4], cell cycle regulation [5], apoptosis [6–8], and in regulating extracellular DNA uptake (competence development) [9]. Evidence for bistable networks has been found in experimental observations of the hysteretic (i.e., history dependent) response to stimuli that is commonly associated with bistability [10,11], for example in the Cdc2 activation circuit in *Xenopus* egg extracts [12,13] and in the lactose utilization network in *E. coli* [14]. Complementing experimental analyses, mathematical tools such as bifurcation theory can be used to determine if a particular network—written as a set of ordinary differential equations (ODEs)—is bistable [15]. However, because the dynamical behavior of a network is dependent on the values of the system parameters (e.g., reaction rates), and the number of parameters required for an accurate description of even simple systems is typically large and uncertain, new bistable circuit architectures tend to be identified only slowly and on a network-by-network basis.

Chemical reaction network theory (CRNT), which gives conditions for the existence, multiplicity, and stability of steady states in systems of nonlinear ODEs derived from mass-action kinetics [16–18], offers a novel framework for the rapid identification of network topologies with the capacity for bistability (herein referred to as bistable networks). Importantly, CRNT is applicable without specific knowledge of the system parameters. This ability to study network characteristics in a parameter-free context is particularly beneficial in cell and developmental biology, given the high level of uncertainty in parameter values [19]. As a result, CRNT has found a number of biological applications [20–23]. Still, considering its potential for large-scale analyses, the use of CRNT has been fairly limited.

Here, we apply CRNT to reaction network models representing a broad class of small gene regulatory networks (GRNs): those consisting of two transcription factor (TF)-coding genes and their associated proteins. Our comprehensive parameter-free survey resulted in the identification of 36,771 bistable GRN architectures (out of a total of 40,680), including eleven without the TF cooperativity typically associated with switch-like circuits. Approximately 40% of the bistable systems were confirmed as such using existing computational tools, with the remainder identified

## Author Summary

Switch-like behavior is found across a wide range of biological systems, and as a result there is significant interest in identifying the various ways in which biochemical reactions can be combined to yield a switch-like response. In this work we use a set of mathematical tools from *chemical reaction network theory* that provide information about the steady-states of a reaction network irrespective of the values of network rate constants, to conduct a large computational study of a family of model networks consisting of only two protein-coding genes. We find that a large majority of these networks (~90%) have (for some set of parameters) the mathematical property known as bistability and can behave in a switch-like manner. Interestingly, the capacity for switch-like behavior is often maintained as networks increase in size through the introduction of new reactions. We then demonstrate using published yeast data how theoretical parameter-free surveys such as this one can be used to discover possible switch-like circuits in real biological systems. Our results highlight the potential usefulness of parameter-free modeling for the characterization of complex networks and to the study of network evolution, and are suggestive of a role for it in the development of novel synthetic biological switches.

through the novel concept of network ancestry, in which the presence of a bistable subnetwork can under certain conditions be used to establish bistability in a larger network if the condition that the two network structures have an identical stoichiometric subspace is met (see the following section on CRNT basics). Despite its large size, the entire two-gene bistable network family can be understood as descended from a set of only eleven minimal bistable networks, that is, bistable networks that do not contain within them a smaller bistable subnetwork and, as a consequence, are rendered monostable by the removal of one or more network reactions. Using experimental protein-protein interaction, protein-DNA interaction, and gene expression data from *Saccharomyces cerevisiae*, we demonstrate how a general theoretical survey of this kind has unique predictive power to identify bistable modules in organisms that have not been fully explored from a functional genomics perspective. Our results are further suggestive of a role for parameter-free modeling in simplifying the study of complex regulatory networks, understanding network evolution, and designing new synthetic biological circuits.

## Results

### Two-gene network construction

As done previously [23], we assume classical chemical kinetics and specify gene regulatory networks (GRNs) as sets of elementary biochemical reactions. For a network consisting of  $N$  transcription factor genes  $X_i$  and associated proteins  $P_i$  ( $i = 1, \dots, N$ ), the essential reactions are basal protein production ( $X_i \rightarrow X_i + P_i$ ) and degradation ( $P_i \rightarrow \emptyset$ ). Networks may also contain protein dimerization reactions ( $P_i + P_j \rightleftharpoons P_i P_j$ ), binding of both TF monomers and dimers to the gene promoters ( $X_i + P_j \rightleftharpoons X_i P_j$  and  $X_i + P_j P_k \rightleftharpoons X_i P_j P_k$ ), and protein production from a bound gene ( $X_i P_j \rightarrow X_i P_j + P_i$  and  $X_i P_j P_k \rightarrow X_i P_j P_k + P_i$ ). For reactions of this last type, under our parameter-free framework, the rate of protein production from a bound gene is unspecified and thus may be either higher or lower than the basal rate but cannot be zero. For simplicity, we assume that the promoter of each gene may only be bound by a single monomer or dimer species at any given time,

or they may remain unoccupied. We further assume that, while degradation is considered for monomeric TFs, all TF dimers are stable to proteolytic degradation; the validity of this assumption and its implications are discussed below.

A variety of networks may be constructed by combining these reactions, subject to certain logical constraints (e.g., the presence of a dimer-promoter binding reaction requires the inclusion of the dimer formation reaction) and with the requirement that every network includes the necessary basal TF production and degradation reactions. In the two-gene case ( $N = 2$ ), there are 4 essential reactions and 23 additional reactions (Table 1) that may be combined to form 40,680 different networks. The total number of networks is smaller than might be expected (i.e., less than  $2^{23}$ ) as a result of reaction dependencies (Table 1) and network symmetries; for example, the network consisting of reactions  $k$ ,  $q$ , and  $w$  is functionally equivalent to the that with reactions  $i$ ,  $l$ , and  $r$ , and as a result we do not include the latter and other symmetric networks like it in the total.

It should be noted that within this set of two-gene networks there are a small number for which there is no coupling between the two genes. Given that there are twelve possible one-gene networks for both  $X_1/P_1$  and  $X_2/P_2$  independently (see [23]), the total number of unique decoupled two-gene networks is  $12(12+1)/2 = 78$ , the number of distinct pairs of one-gene circuits. The presence of 78 decoupled two-gene networks was verified by searching through the full list of 40,680 networks for those lacking the basic coupling reactions  $b$ ,  $c$ ,  $j$ ,  $n$ , and  $o$  (Table 1).

### Chemical reaction network theory basics

Given the centrality of CRNT to our analysis, we provide here a primer on the relevant aspects of the theory and illustrate them with the rudimentary two-gene network that consists of only the essential basal protein production and degradation reactions (Figure 1).

At the heart of the theory is the concept of network *complexes*, formally the chemical species or linear combinations of species which occur on either side of a chemical equation. A reaction network can be visualized as a directed graph where each of these complexes appears only once at the heads and/or tails of reaction arrows. A collection of complexes connected by arrows is referred to as a *linkage class*. The complexes and linkage classes for our rudimentary network are highlighted in Figure 1 in yellow and with dashed lines, respectively.

Every complex in the network can also be represented as a vector in an appropriate vector space; in a network of  $N$  species, the complex vectors lie in  $\mathbb{R}^N$ . Reactions also have associated vectors (termed *reaction vectors*), which are constructed by subtracting the reactant complex vectors from the product complex vectors. The size of the largest linearly independent set of reaction vectors is the *rank* of the network, and the set of all possible linear combinations of reaction vectors (i.e., their span) is referred to as the *stoichiometric subspace* of the network. This subspace plays an important role in setting boundaries on the system behavior: although the species' concentrations may evolve with time, they are ultimately constrained within surfaces that are parallel translations of the stoichiometric subspace. Exactly which surface (or *stoichiometric compatibility class*) the concentrations are constrained to is defined by the initial conditions.

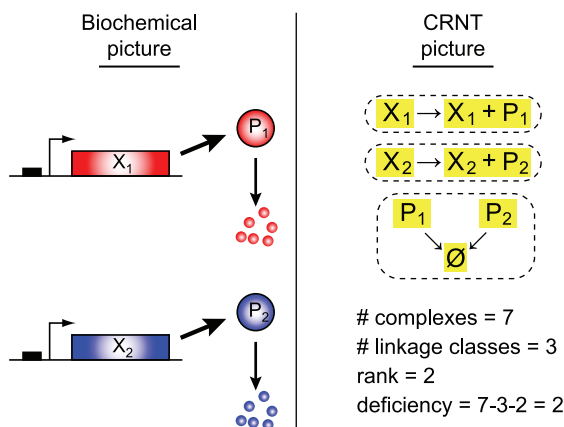
For a system with  $n$  complexes,  $l$  linkage classes, and rank  $s$ , the network *deficiency*  $\delta$  is defined as  $\delta = n - l - s$ . A number of theorems regarding the stability properties of networks are based on the deficiency, including the deficiency zero and deficiency one theorems [16,17].

**Table 1.** Reactions combined to generate the 40,680 unique networks of two genes and two gene products.

Reaction label $r_i$	Reaction	Dependencies	Biochemical process
*	$X_1 \rightarrow X_1 + P_1$	–	gene $X_1$ basal protein production
*	$X_2 \rightarrow X_2 + P_2$	–	gene $X_2$ basal protein production
*	$P_1 \rightarrow \emptyset$	–	protein $P_1$ degradation
*	$P_2 \rightarrow \emptyset$	–	protein $P_2$ degradation
<i>a</i>	$X_1 + P_1 \rightleftharpoons X_1 P_1$	–	binding of $P_1$ to the $X_1$ promoter
<i>b</i>	$X_1 + P_2 \rightleftharpoons X_1 P_2$	–	binding of $P_2$ to the $X_1$ promoter
<i>c</i>	$X_2 + P_1 \rightleftharpoons X_2 P_1$	–	binding of $P_1$ to the $X_2$ promoter
<i>d</i>	$X_2 + P_2 \rightleftharpoons X_2 P_2$	–	binding of $P_2$ to the $X_2$ promoter
<i>e</i>	$X_1 P_1 \rightarrow X_1 P_1 + P_1$	<i>a</i>	production of $P_1$ from a $P_1$ -bound gene
<i>f</i>	$X_1 P_2 \rightarrow X_1 P_2 + P_1$	<i>b</i>	production of $P_1$ from a $P_2$ -bound gene
<i>g</i>	$X_2 P_1 \rightarrow X_2 P_1 + P_2$	<i>c</i>	production of $P_2$ from a $P_1$ -bound gene
<i>h</i>	$X_2 P_2 \rightarrow X_2 P_2 + P_2$	<i>d</i>	production of $P_2$ from a $P_2$ -bound gene
<i>i</i>	$P_1 + P_1 \rightleftharpoons P_1 P_1$	–	homodimerization of $P_1$
<i>j</i>	$P_1 + P_2 \rightleftharpoons P_1 P_2$	–	heterodimerization of $P_1$ and $P_2$
<i>k</i>	$P_2 + P_2 \rightleftharpoons P_2 P_2$	–	homodimerization of $P_2$
<i>l</i>	$X_1 + P_1 P_1 \rightleftharpoons X_1 P_1 P_1$	<i>i</i>	binding of $P_1 P_1$ dimer to the $X_1$ promoter
<i>m</i>	$X_1 + P_1 P_2 \rightleftharpoons X_1 P_1 P_2$	<i>j</i>	binding of $P_1 P_2$ dimer to the $X_1$ promoter
<i>n</i>	$X_1 + P_2 P_2 \rightleftharpoons X_1 P_2 P_2$	<i>k</i>	binding of $P_2 P_2$ dimer to the $X_1$ promoter
<i>o</i>	$X_2 + P_1 P_1 \rightleftharpoons X_2 P_1 P_1$	<i>i</i>	binding of $P_1 P_1$ dimer to the $X_2$ promoter
<i>p</i>	$X_2 + P_1 P_2 \rightleftharpoons X_2 P_1 P_2$	<i>j</i>	binding of $P_1 P_2$ dimer to the $X_2$ promoter
<i>q</i>	$X_2 + P_2 P_2 \rightleftharpoons X_2 P_2 P_2$	<i>k</i>	binding of $P_2 P_2$ dimer to the $X_2$ promoter
<i>r</i>	$X_1 P_1 P_1 \rightarrow X_1 P_1 P_1 + P_1$	<i>i, l</i>	production of $P_1$ from a $P_1 P_1$ -bound gene
<i>s</i>	$X_1 P_1 P_2 \rightarrow X_1 P_1 P_2 + P_1$	<i>j, m</i>	production of $P_1$ from a $P_1 P_2$ -bound gene
<i>t</i>	$X_1 P_2 P_2 \rightarrow X_1 P_2 P_2 + P_1$	<i>k, n</i>	production of $P_1$ from a $P_2 P_2$ -bound gene
<i>u</i>	$X_2 P_1 P_1 \rightarrow X_2 P_1 P_1 + P_2$	<i>i, o</i>	production of $P_2$ from a $P_1 P_1$ -bound gene
<i>v</i>	$X_2 P_1 P_2 \rightarrow X_2 P_1 P_2 + P_2$	<i>j, p</i>	production of $P_2$ from a $P_1 P_2$ -bound gene
<i>w</i>	$X_2 P_2 P_2 \rightarrow X_2 P_2 P_2 + P_2$	<i>k, q</i>	production of $P_2$ from a $P_2 P_2$ -bound gene

\*These reactions occur in every network.  
doi:10.1371/journal.pcbi.1002039.t001

Advanced deficiency theory (ADT) [18] is required for networks with a deficiency greater than one. The ADT algorithm, detailed

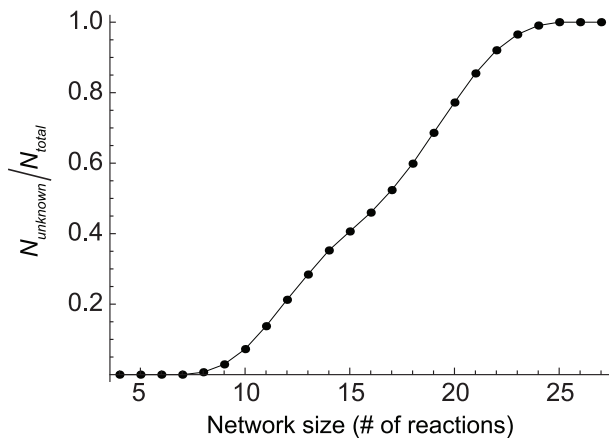


**Figure 1. Rudimentary two-gene network consisting of only basal protein production and degradation.** In the 'CRNT picture', complexes are highlighted in yellow and linkage classes are identified with dashed lines. Labeling scheme is adopted from [77].  
doi:10.1371/journal.pcbi.1002039.g001

in [24] and implemented in the Chemical Reaction Network Toolbox software package (<http://www.chbmeng.ohio-state.edu/~feinberg/crntwin/>), constructs and attempts to solve systems of equalities and inequalities that are based on the network structure. If no solutions (which together with the equality and inequality systems are known as 'signatures' of the reaction network) can be found, then the network does not have the qualitative capacity to support multiple steady states. However, if signatures can be found, then the network can support multiple steady states, and the Toolbox will produce example rate constants and associated steady states consistent with the mass-action ODE description of the network, as well as report the stability characteristics of the steady states. It should be emphasized that ADT cannot guarantee bistability even if the network does support multiple steady states, as they may be unstable. Nevertheless, with its substantial analytical power and ease of use, ADT has played a role in a number of recent studies [23,25–28].

### Preliminary bistable network identification

All of the two-gene networks modeled were found by the Chemical Reaction Network Toolbox (herein referred to as simply the Toolbox) to have a deficiency of two or more, necessitating the use of ADT in their analyses. Screening the Toolbox-generated ADT analysis reports, we determined that of the 40,680 networks

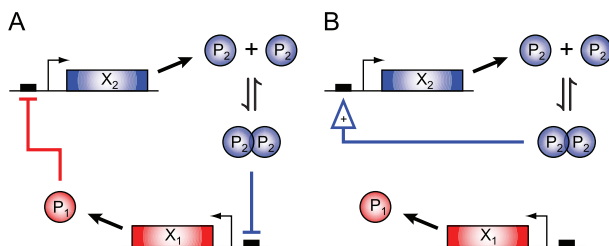


**Figure 2. Fraction of networks which cannot have their stability established by advanced deficiency theory (ADT), as a function of network size.**

doi:10.1371/journal.pcbi.1002039.g002

surveyed, 18,352 (~45%) have the capacity for multiple steady states, with 14,721 of these being confirmed as bistable with example rate constants (see Materials and Methods for a description of the screening procedure). Only 2,654 networks (~6.5%) cannot be bistable regardless of the parameter values. For the remaining 19,674 networks, ADT could neither establish nor rule out the capacity for multiple steady states, and as a result we refer to these as 'unknown' networks. It is noteworthy that the fraction of networks of a given size (that is, a given number of reactions) that are unknown increases as the size increases; for example, >90% of networks with more than 21 reactions, and all networks with more than 24 reactions, are unknown (Figure 2). As expected, the stabilities of the decoupled two-gene networks are the same as the constituent one-gene systems previously studied [23].

The two smallest bistable networks identified exhibit canonical switch topologies (Figure 3). In the double negative feedback circuit shown in Figure 3A, we find that dimerization of only one of the TFs is sufficient for bistability. The autoregulatory positive feedback network shown in Figure 3B is an example of a decoupled two-gene network, with bistability in the concentration of one TF only. We note that while CRNT does not take into account the strength of the regulation in determining a network's capacity for multiple steady states, the fact that an autoregulatory circuit requires positive feedback in order to achieve bistability is



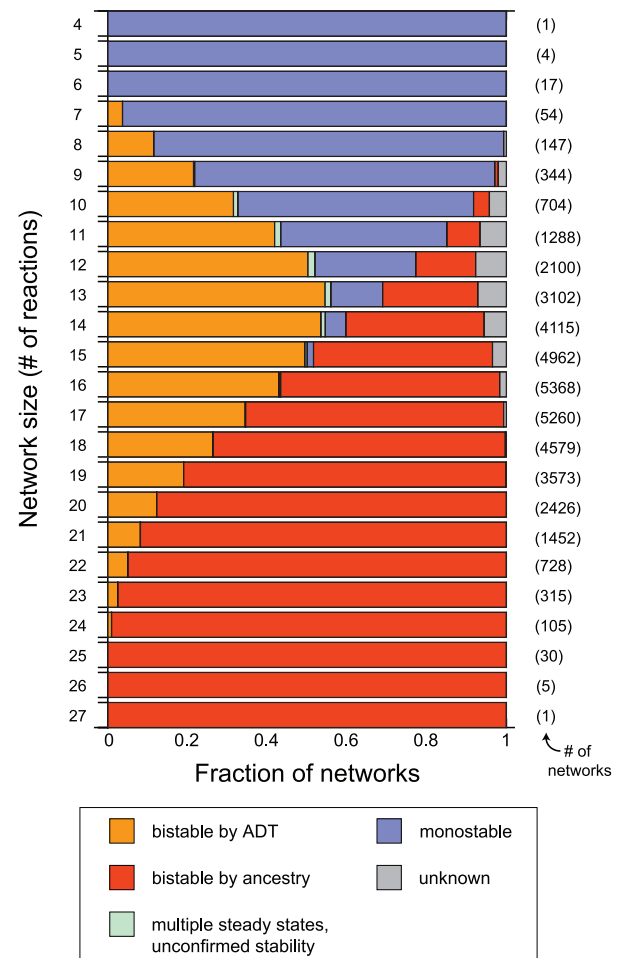
**Figure 3. The smallest two-gene bistable networks found with ADT.** (A) A double negative feedback circuit, in which dimerization of only one of the TFs is sufficient for bistability. (B) An autoregulatory positive feedback circuit. The two genes are uncoupled and the bistability is in the concentration of one TF only. In both (A) and (B), degradation of the TF monomers is not shown.

doi:10.1371/journal.pcbi.1002039.g003

well-established (see, e.g., [29,30]). Bistability via positive autoregulation has also been demonstrated experimentally with synthetic gene circuits in both prokaryotes [31] and eukaryotes [32].

### Identifying bistability through network ancestry

The bistable networks shown in Figure 3, each containing seven reactions, can be 'grown' into new eight-reaction networks through the addition of reactions from Table 1: reactions *a*, *b*, *d*, *g*, *i*, *j*, *q*, or *t* to the circuit shown in Figure 3A, and reactions *a*, *b*, *c*, *d*, *i*, *j*, or *n* to the circuit shown in Figure 3B. In all cases, the new larger networks were also confirmed by the Toolbox to be bistable. We may then ask: is bistability, once established in a 'parent' network of *N* reactions, guaranteed in any 'descendant' network of *N* + 1 reactions? ADT alone is not sufficient to answer this question, since systems were less likely to be characterizable as they increased in size (Figure 2). However, CRNT does provide a basis for establishing bistability in networks which contain subnetworks known to be bistable: if following the addition of a reaction the stoichiometric subspace of the

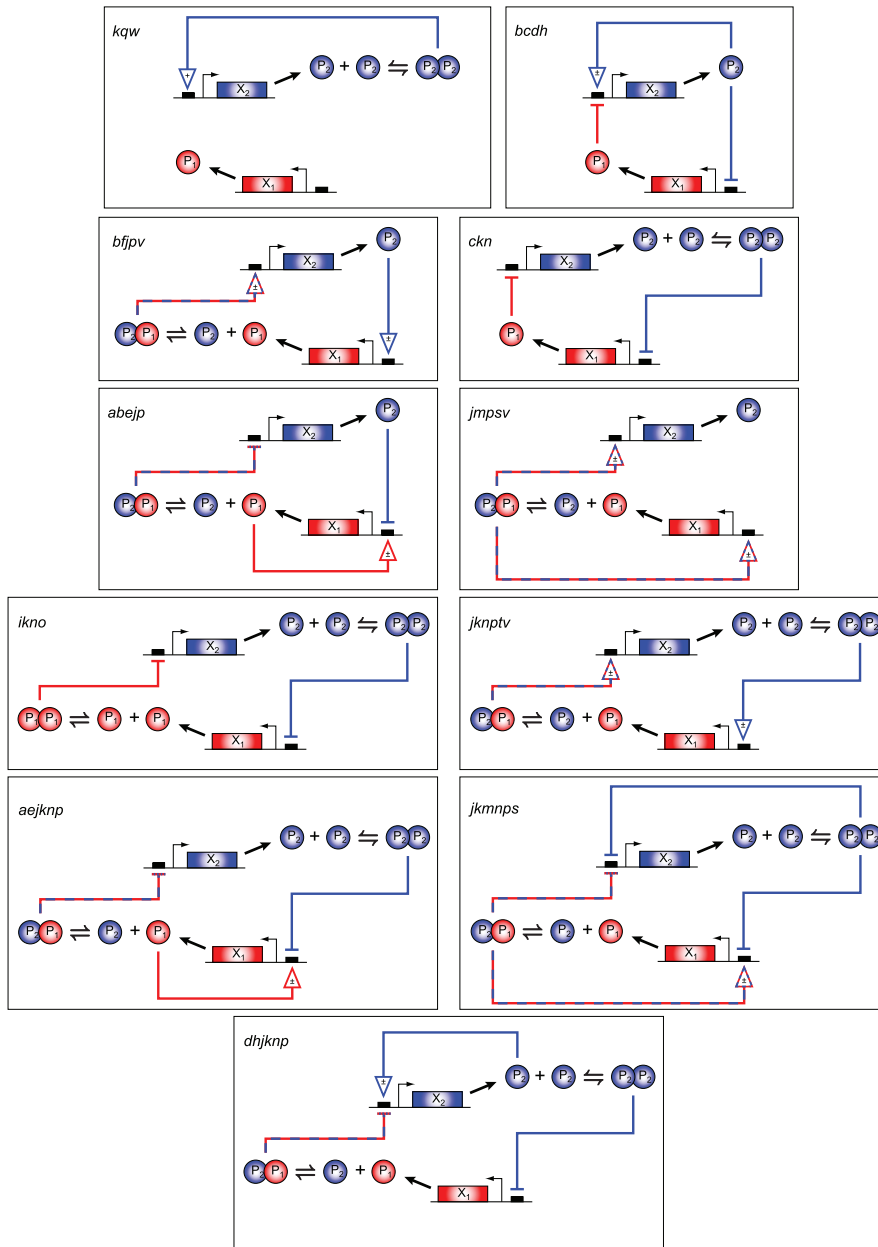


**Figure 4. The fraction of networks of each size that were established as bistable by ADT, bistable by network ancestry, having multiple steady states with unconfirmed stability, monostable, or with an unknown capacity for multiple stable steady states.** Network size is determined only by the number of reactions (from Table 1) that are present. The total number of networks of each size is shown in parentheses.

doi:10.1371/journal.pcbi.1002039.g004

descendant network is identical to that of the parent, then the larger network is also bistable for some set of parameter values. As an intuitive example, one can imagine a situation in which a reaction is added to an existing network, that the surface containing the dynamical trajectories of the network species' concentrations is not changed as a result of the addition, and that the added reaction has only a very small rate constant. In this case we should not expect a change from whatever qualitative phenomena were there before. Thus, if the parent network had two stable equilibria, the descendant network will also have two stable equilibria. Example reactions that do not result in a change in the stoichiometric subspace if added include protein produc-

tion from a TF-bound gene ( $X_i P_j \rightarrow X_i P_j + P_i$ , since the reaction vectors can be written as linear combinations of the vectors associated with  $X_i + P_j \rightleftharpoons X_i P_j$ ,  $P_i \rightarrow \emptyset$ , and  $P_j \rightarrow \emptyset$ ). Beginning with the 14,721 known bistable networks and using this 'ancestry'-based method, we identified an additional 22,050 bistable networks. Some of these networks had been previously found by the Toolbox to have the capacity for multiple steady states, but for which no example parameter sets leading to stable equilibria were given. The number of networks of each type—bistable by ADT, bistable by ancestry, multiple steady states with unconfirmed stability, monostable, or unknown—are shown as a function of network size in Figure 4.



**Figure 5. Minimal bistable networks.** Only 11 of the 36,771 bistable networks identified lose bistability by the removal of any network reaction. That is, only 11 of the bistable networks contain no subset of reactions which is also bistable. Dashed-and-colored lines indicate regulation by heterodimer. Horizontal bars represent purely-repressive TF binding, and arrows indicate TF production from a bound gene (at a non-zero rate that may be either higher or lower than the basal rate). Degradation of the TF monomers is not shown.  
doi:10.1371/journal.pcbi.1002039.g005

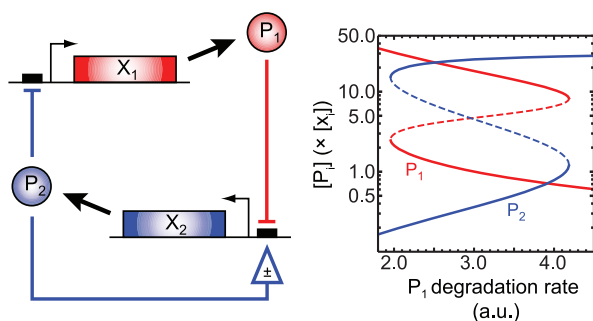


## Minimal bistable networks

Of the 36,771 bistable systems identified, only eleven do not contain within them a smaller subnetwork that is also bistable. For these eleven networks, the removal of any single reaction would result in a loss of bistability. We refer to these networks as *minimal bistable networks* (MBNs). Named according to the reaction labels in Table 1, the MBNs are: *kqv*, *ckn*, *bcdh*, *ikno*, *jmpsv*, *bfjpv*, *abejp*, *jknptv*, *jknmps*, *dhjknv*, and *aejknv*. The two networks shown in Figure 3 are minimal (*kqv* and *ckn*); the full set is shown in Figure 5. Arrows containing the symbol ( $\pm$ ) are used in the figure and all that follow to emphasize that, in assessing a network's capacity for multiple steady states, CRNT does not distinguish between up-regulation and down-regulation that results in reduced but non-zero expression. With the exception of *bcdh* (discussed in more detail in the following section), all of these networks contain one or more of the TF dimerization reactions common in bistable GRNs [23]. It can also be seen that each MBN contains feedback loops that for some parameter sets will be made positive, a characteristic shown to be generally necessary for multiple steady states in a system of ODEs [33].

## Cooperativity-free switches

Although cooperativity in gene regulation—via either the non-independent binding of TFs to multiple promoter sites or the multimerization of TFs into functional units—is an important component of some bistable networks [34,35], it is not necessary for bistability. Indeed, a number of recent mathematical models of GRNs have shown deterministic bistability without cooperativity of any kind [36–38]. Among the 40,680 two-gene networks are 45 lacking cooperativity, and of these 31 were found to be monostable, eight were identified as bistable directly by ADT, and three more were identified as bistable by network ancestry. All of the bistable networks lacking cooperativity can be derived from the MBN *bcdh*, which is shown in Figure 6 along with a bifurcation diagram showing the existence of two stable equilibria (and an unstable equilibrium) for a range of  $P_1$  degradation rate constants. The complete set of cooperativity-free bistable networks is shown in Figure 7. An essential feature of these circuits is the competitive binding of  $P_1$  and  $P_2$  to the  $X_2$  promoter. Similar competitive or sequestration-type processes have been found to be key components in some switch-like systems [36–40].



**Figure 6. Example of a bistable network lacking cooperativity.** TF  $P_2$  plays a dual role as an activator of  $X_2$  and a repressor of  $X_1$ . The bifurcation plot shows the stable (solid lines) and unstable (dashed lines) steady state protein concentrations, in units relative to the DNA concentration, for one set of parameter values as a function of the  $P_1$  degradation rate. The network ODEs and parameter values are given in the Supplementary Text S1.  
doi:10.1371/journal.pcbi.1002039.g006

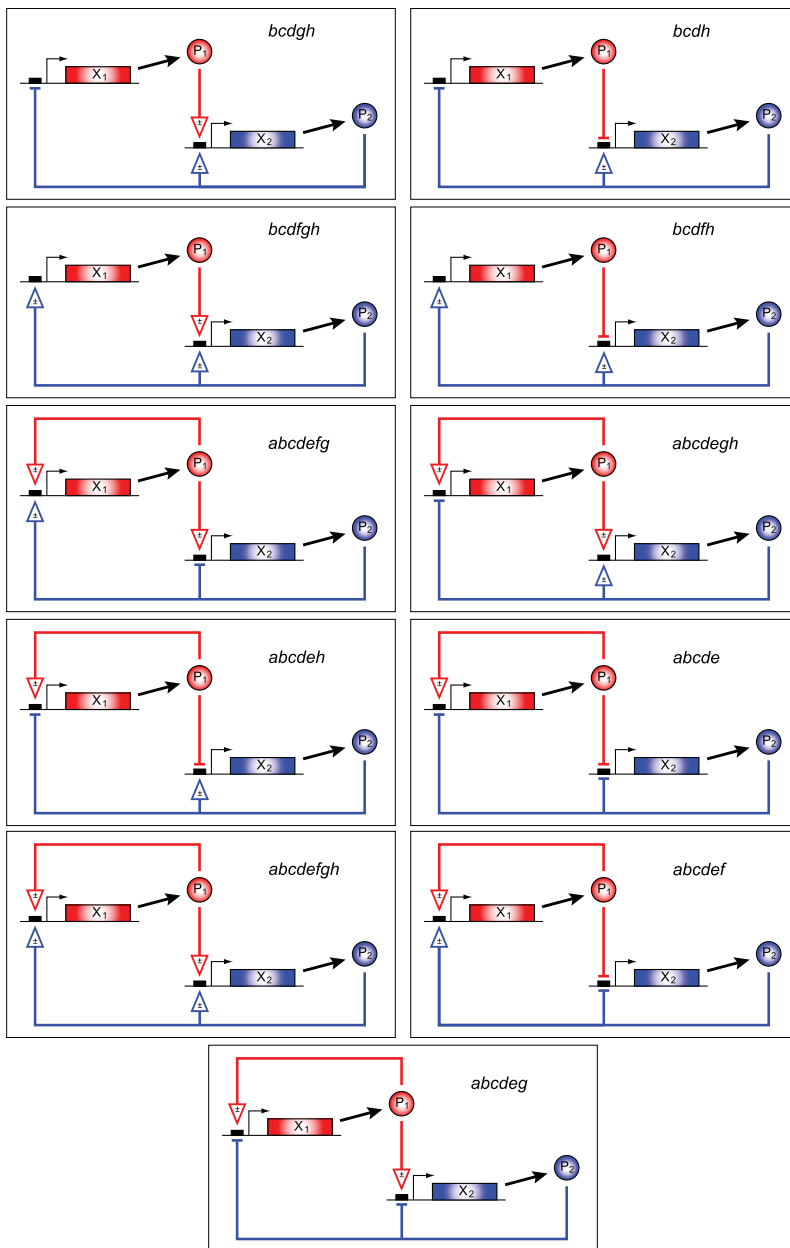
## Two-gene networks in *S. cerevisiae*

To investigate how an *in silico* network topology survey such as this can be used to better understand experimental results, we searched for real biological examples of the bistable networks identified in this study in the model organism *S. cerevisiae*. To our knowledge, there is no single database that contains *S. cerevisiae* GRN architecture, thus we combined protein-protein and protein-DNA interaction data with gene expression data to establish the large-scale empirical network shown in Supplementary Figure S1. Included in this network are 148 TFs participating in 205 protein-protein interactions (61 heterodimerization and 144 homodimerization reactions), along with 1,249 interactions between 139 TFs and 208 genes (37 ‘self-binding’ and 1,212 ‘cross-binding’ reactions). To establish which of the two-gene bistable circuits are present in the yeast network, it was first necessary to ‘translate’ the bistable models from their ideal, theoretical description (that distinguishes between and allows for each elementary reaction) into a format that is more amenable to experimental data mining; see Supplementary Text S1. We were then able to identify in the yeast data a total of 1,289 two-gene GRNs, twelve of which have topologies consistent with members of the MBN set (Table 2). Examples of these are highlighted in the next section.

## Discussion

The idea of studying theoretical network models generated via ‘random wiring’ was suggested at least fifty years ago by Monod and Jacob [41]. Only recently, with the development of powerful computational tools, have a variety of simple gene regulatory and metabolic network topologies been studied with surveys over large ranges of parameter space [42,43]. Parameter-free techniques such as CRNT are particularly well-suited for general surveys aimed at bistable network discovery, as they may more definitively answer questions regarding a mass action system’s ability to support multiple steady states. For example, using only the advanced deficiency theory (ADT) algorithm implemented in the Chemical Reaction Network Toolbox we were able to establish that  $\sim 36\%$  of the 40,680 possible unique two-gene networks are bistable for at least some sets of network parameters, another  $\sim 9\%$  have the capacity for multiple steady states (which may or may not be stable), and only  $\sim 6.5\%$  are monostable regardless of the network parameters.

As network size and complexity increases, the ability of ADT to draw conclusions becomes limited (Figure 2). One method put forward as a way to extend the usefulness of CRNT to larger networks involves the analysis of simpler subnetworks corresponding with elementary flux modes of the system [25]. We have introduced a complementary subnetwork analysis method for identifying bistability, termed network ancestry, which requires only a topological sorting of the networks based on the presence or absence of individual reactions followed by inspection of the network reaction vectors. If the parent network is determined to be bistable, and if the reaction vectors of the bistable parent and unknown descendant have the same span (i.e., the networks have an identical stoichiometric subspace), then the descendant is also bistable. As a result of network ancestry, we were able to identify an additional 22,050 networks with previously unknown stability as bistable,  $\sim 54\%$  of the total (Figure 4). We emphasize that a change in the size of the stoichiometric subspace does not in and of itself imply that bistability will be lost; however, from a purely topological perspective, it may not be obvious what the effect of the change may be. Our network ancestry method may thus be considered a relatively conservative one for establishing bistability in large networks.



**Figure 7. Bistable networks without cooperativity.** Of the 45 two-gene networks lacking dimerization, 11 were identified as bistable either directly by advanced deficiency theory analysis or via network ancestry. All the dimer-free bistable networks shown here can be derived from the minimal bistable network *bcdh* through the addition of reactions from Table 1. Horizontal bars represent purely-repressive TF binding, and arrows indicate TF production from a bound gene (at a non-zero rate that may be either higher or lower than the basal rate). Degradation of the TF monomers is not shown.

doi:10.1371/journal.pcbi.1002039.g007

The assumption of mass action kinetics is an important aspect of CRNT. Consequently, Michaelis-Menten and Hill-type expressions are not used in our CRN approach, as they require approximations to mass action that cannot be validated in a parameter-free context. In addition, it was recently demonstrated for a generic two-protein interaction network that bistability present under the ‘inconsistent’ assumption of Michaelis-Menten kinetics is lost when the system is ‘unpacked’ into its fundamental chemical steps [44]. For our two-gene networks, the Michaelis-Menten and CRN descriptions could be approximately equivalent only for specific parameters, and only if those parameters

were such that 1) the DNA-binding reactions reach their equilibria much more quickly than other reactions in the network, and 2) the equilibrium concentrations of any dimer species were proportional to the product of their constituent monomer concentrations [45].

In addition to the inherent consistency of CRN models, the mathematical theory applicable to deterministic CRNs offers significant computational advantages over other methods, in particular stochastic simulation. Furthermore, many deterministically bistable networks have been shown to retain two long-lived states when their models are reformulated to take stochasticity into

**Table 2.** Two-gene networks found in *S. cerevisiae* that have topologies consistent with members of the minimal bistable network set.

Bistable model*	X <sub>1</sub>	X <sub>2</sub>
<i>ckn</i>	PDR1 (YGL013C)	RPN4 (YDL020C)
<i>bcdh</i>	FHL1 (YPR104C)	MSN4 (YKL062W)
<i>bcdh</i>	HMS1 (YOR032C)	YAP6 (YDR259C)
<i>bcdh</i>	IXR1 (YKL032C)	PHD1 (YKL043W)
<i>bcdh</i>	RPN4 (YDL020C)	YAP1 (YML007W)
<i>bcdh</i>	FKH1 (YIL131C)	FKH2 (YNL068C)
<i>jknpv</i>	MTH1 (YDR277C)	RGT1 (YKL038W)
<i>aejkn</i>	OAF1 (YAL051W)	PIP2 (YOR363C)
<i>abejp</i>	NRG1 (YDR043C)	RIM101 (YHL027W)
<i>abejp</i>	IFH1 (YLR223C)	RAP1 (YNL216W)
<i>bfpjv</i>	KSS1 (YGR040W)	CST6 (YIL036W)
<i>bfpjv</i>	OPI1 (YHL020C)	INO2 (YDR123C)

\*Model names refer to the constituent reactions as labeled in Table 1.  
doi:10.1371/journal.pcbi.1002039.t002

account [37,44,46,47]. Still, as biochemical noise has been shown to drive some systems to exhibit switch-like behavior not predicted by deterministic models [37,47–50], it should be considered in any complete study of a specific network of interest. For models already formulated as CRNs, stochastic simulation is relatively straightforward (see, e.g., [44,51]).

We attempted to capture the most prevalent and basic biochemical processes involved in transcriptional regulation in our network model construction, but our formalism is by no means exhaustive. One mechanism not included and through which networks can achieve the nonlinearity required for bistability is the direct degradation of TF dimers ( $P_i P_j \rightarrow \emptyset$ ) [38]. Given that dimerization regularly protects against proteolysis (see, e.g., [52,53]), its exclusion from our reaction set is reasonable. Furthermore, for most of the networks analyzed here, the addition of a dimer degradation reaction would have no effect on their capacity for bistability: since the reaction vector for  $P_i P_j \rightarrow \emptyset$  can be written as a linear combination of the vectors associated with reactions  $P_i + P_j \rightleftharpoons P_i P_j$ ,  $P_i \rightarrow \emptyset$  and  $P_j \rightarrow \emptyset$ , any descendant network grown from a bistable parent via the addition of a dimer degradation reaction would have the same stoichiometric subspace and would be bistable as a result of network ancestry.

There remains a large amount of additional biological detail which could be incorporated in future surveys, including post-translational modification, multiple promoter binding sites, and the location of regulatory elements relative to the genes (which has been shown to play a role in network bistability [54]). However, any increase in the level of detail would result in an increase in the combinatorial complexity and the size of the survey. For example, whereas the set of one-gene networks are constructed using combinations of 5 different reactions [23], and our two-gene networks using 23 reactions (Table 1), the addition of a third gene alone would lead to 60 different reactions that could be ‘wired’ together. With the current version of the Toolbox taking (at best) many seconds to import, analyze, and export the results for every network model, it is perhaps not an ideal software package for surveys significantly larger than this one. New software implementations of CRNT continue to be developed (e.g., [55]), and we anticipate that future programs will

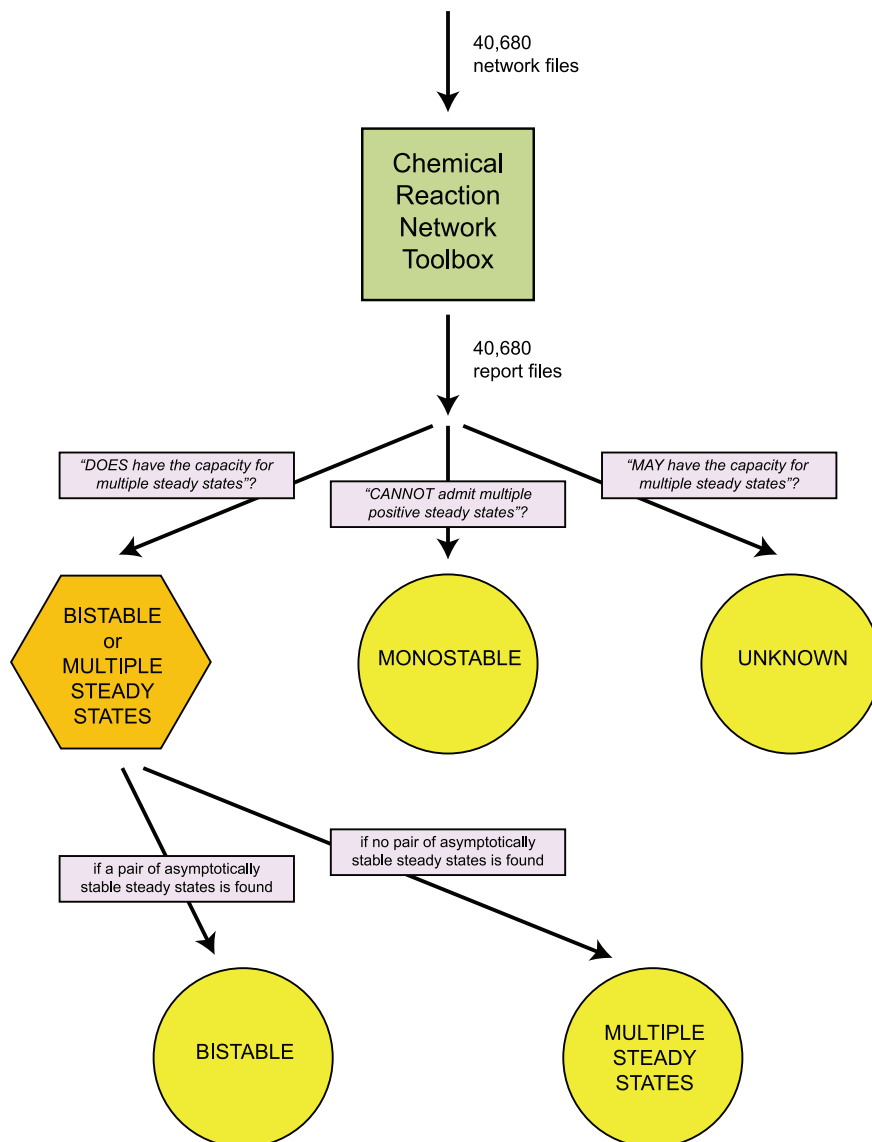
allow for even more comprehensive computational studies. In the meantime, network ancestry offers an attractive solution to the problem of scalability and applicability of CRNT to more complex networks: once all fundamental chemical reactions involved in any network of interest are identified, one could assemble the minimal network topologies covering all possible unique stoichiometric subspaces and probe that smaller set of networks for bistability. In essence, network ancestry allows for the reduction of the problem of determining a large network’s qualitative capacity for bistability to one of identifying the minimal bistable subnetworks within it.

There is a strong biological motivation to consider individual networks as parents and descendants with a topological ordering: rather than appearing *de novo*, modern GRNs grow from ancestor network kernels through mechanisms such as gene duplication and the accretion of protein domains [56–59]. Domain accretion, for example the acquisition of a DNA-binding domain by a monomer (modeled in this work by the addition of one of the promoter binding reactions *a*, *b*, *c*, or *d*), has been proposed to be particularly important for eukaryotic evolution [60,61]. And there is evidence suggesting an even more direct role for bistability in evolution: it is the primary requirement for epigenetic inheritance mechanisms known to have important evolutionary effects [62,63], and can also lead to increased population fitness in stressful or changing environments [64,65] by driving an increase in phenotypic heterogeneity [66]. Thus, the eleven MBNs identified here (Figure 5), which differ from monostable networks by just a single reaction, may represent an interesting class of networks from the standpoint of evolutionary biology, as it may be that similarly-minimal networks have played an important role in functional development and/or speciation.

We used the results of our *in silico* analysis to motivate a search of—and add functional context to—existing yeast protein-DNA and protein-protein interaction data, and in doing so were able to identify a number of two-gene systems with topologies consistent with bistability. For example, the FKH1 and FKH2 genes (and their associated proteins Fkh1p and Fkh2p, which compete for target promoter occupancy [67]) compose a network with a topology similar to the MBN *bcdh* (Table 2). FKH1 and FKH2 belong to the pervasive winged-helix/forkhead (FOX) family of TFs and are essential for proper regulation of the yeast cell cycle [68]. Other FOX genes have previously been shown to be involved in important biological functions including cell cycle regulation and cell differentiation [69], two processes for which GRN bistability has been implicated [2–5].

Additional gene pairs of interest include NRG1/RIM101 and OAF1/PIP2, which are components of GRNs with topologies similar to that of MBNs *abejp* and *aejkn*, respectively. The Rim101p and Nrg1p proteins, both identified previously as transcriptional repressors, are components in an extracellular pH-responsive differentiation pathway in yeast [70]. Further evidence suggestive of bistability in this system can be found in *C. albicans*, in which Rim101p and Nrg1p homologs regulate the morphological switch [71] associated with a dramatic change in the pathogen’s virulence [72]. Oaf1p and Pip2p, on the other hand, are involved in the production of peroxisomal proteins in the presence of fatty acids [73], and have been shown to be involved in the coordination of two different transcriptional responses to oleate [74]. We emphasize that while the two-gene networks identified through our analysis are not guaranteed to be bistable, their known topologies and functions make them excellent bistable network candidates, providing powerful hypotheses for further experimentation. The same approach may be used





**Figure 8. Screening networks for different steady state behaviors.** Networks are initially screened by the content of analysis reports produced by the Chemical Reaction Network Toolbox. The networks designated ‘multiple steady states’ are those determined by ADT to have the capacity for multiple steady states but for which no example pair of asymptotically stable steady states could be found by the program. Bistable networks are those for which an example pair of asymptotically stable steady states was reported. The complete sorting procedure is described in Materials and Methods.

doi:10.1371/journal.pcbi.1002039.g008

to provide guidance or functional context to any system for which the necessary interaction data is available.

High-throughput parameter-free analysis holds potential, not just as a tool for the study of natural systems, but also as a design aid in the growing field of synthetic biology [75,76]. For example, a survey such as this can provide inspiration for the development of new bistable switches and a library of models to draw from; already we have proposed a set of novel bistable networks that lack cooperativity and which may be particularly good designs as a result (e.g., because they do not require any ‘extra’ engineering of dimerization domains). At the very least, such a broad application of CRNT may be used to rule out (possibly large numbers of) designs incapable of bistability. CRNT can be similarly used to rule out circuits without the capacity for sustained oscillations [16] or those which cannot exhibit ‘absolute concentration robustness’ [77].

It is worth emphasizing that the region of parameter space supporting bistability in any individual network cannot be determined via parameter-free techniques alone. For example, it may be that the necessary parameter values lie outside the range of biological reality or are difficult to engineer, or that the size of the bistable region of parameter space is exceedingly small. However, in many large-scale studies, such as those that resulted in the yeast data sets used in this work, a high degree of biochemical detail is simply nonexistent. While this lack of quantitative detail can make some analyses of biological networks challenging, it also opens up opportunities for parameter-free studies to provide experimental guidance and new functional insights [78]. Once identified, potentially interesting network architectures may be analyzed in more detail, with rate constants chosen, for example, by Monte Carlo sampling of parameter space.

## Materials and Methods

### Two-gene network construction

Two-gene networks were generated in MATLAB (2009a, The MathWorks, Inc.) by first enumerating all possibilities and then removing one network from each symmetric pair (defined by two functionally-equivalent networks which can be made identical through a simple change of component subscripts). The heterodimers  $P_1P_2$  and  $P_2P_1$  were assumed to be equivalent.

### Chemical reaction network theory analysis and network screening

Advanced deficiency theory analysis was done using a preliminary version of the Chemical Reaction Network Toolbox v2.0 (<http://www.chbmeng.ohio-state.edu/~feinberg/crntwin/>) made available to us by M. Feinberg and automated with AutoIt v3 (<http://www.autoitscript.com/autoit3/index.shtml>).

Networks were screened based on the content of the analysis reports generated by the Toolbox. These reports, though unique to each network, all contain one of three statements: either the network “DOES have the capacity for multiple steady states”, “CANNOT admit multiple positive steady states”, or “MAY have the capacity for multiple steady states”. Networks with reports containing one of the latter two statements were labeled monostable and unknown, respectively. If a network was determined by ADT to have the capacity for multiple steady states, the analysis report also contained one (or more) example set(s) of rate constants and the associated pair(s) of distinct steady states. However, each steady state may be either asymptotically stable, unstable, or with a stability that is “left undetermined”. Only those networks that could support multiple steady states and for which an example pair of asymptotically stable steady states was given were deemed to be bistable networks. This is not to imply that multiple steady state networks without such an example are not bistable, only that we were unable to confirm their bistability with ADT. The screening procedure is shown schematically in Figure 8.

Network ancestry and minimal bistable network analysis was done using MATLAB. Parent and descendant network pairs were found by simple comparison of the networks stoichiometric subspaces and their constituent reactions (descendant networks contain all the same reactions as their parents plus one additional reaction). Cooperativity-free networks were identified by their lack of dimerization reactions, since by construction, the model genes do not have two TF binding sites that could be occupied simultaneously and there are no multi-protein complexes larger than dimers.

Additional data analysis was done with MATLAB and Mathematica (Wolfram Research, Inc.). The bifurcation plot shown in Figure 6 was generated using XPPAUT (<http://www.math.pitt.edu/~bard/xpp/xpp.html>).

## References

- Pomeroy JR (2008) Uncovering mechanisms of bistability in biological systems. *Curr Opin Biotechnol* 19: 381–8.
- Lai K, Robertson MJ, Schaffer DV (2004) The sonic hedgehog signaling system as a bistable genetic switch. *Biophys J* 86: 2748–57.
- Laslo P, Spooner CJ, Warmflash A, Lancki DW, Lee HJ, et al. (2006) Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell* 126: 755–66.
- Huang S, Guo YP, May G, Enver T (2007) Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Dev Biol* 305: 695–713.
- Cross FR, Archambault V, Miller M, Klovstad M (2002) Testing a mathematical model of the yeast cell cycle. *Mol Biol Cell* 13: 52–70.
- Bagci EZ, Vodovotz Y, Billiar TR, Ermentrout GB, Bahar I (2006) Bistability in apoptosis: roles of Bax, Bcl-2, and mitochondrial permeability transition pores. *Biophys J* 90: 1546–59.
- Eissing T, Conzelmann H, Gilles ED, Allgöwer F, Bullinger E, et al. (2004) Bistability analyses of a caspase activation model for receptor-induced apoptosis. *J Biol Chem* 279: 36892–7.
- Legewie S, Blüthgen N, Herzog H (2006) Mathematical modeling identifies inhibitors of apoptosis as mediators of positive feedback and bistability. *PLoS Comput Biol* 2: e120.
- Avery SV (2005) Cell individuality: the bistability of competence development. *Trends Microbiol* 13: 459–62.

### Identification of bistable networks in *S. cerevisiae*

A set of 228 yeast genes previously established as coding for transcriptional regulators [79,80] was used as the primary source for candidate network TF genes (Supplementary Table S1). Protein-protein interactions were retrieved from the BioGRID database [81] (Supplementary Table S2) and protein-DNA interactions were retrieved from the YeastRACT database [82] (Supplementary Table S3). The effect of the protein-DNA interactions on target gene expression (activation or repression) is usually unknown, and any information suggestive of a particular effect was used supplementarily in the network discovery process (Supplementary Table S4).

### Supporting Information

**Figure S1** Large-scale GRN in *S. cerevisiae*. GRN was generated through the combination of protein-protein interaction, protein-DNA interaction, and gene expression data. (EPS)

**Table S1** List of genes/proteins considered as transcriptional regulators in yeast. Data taken from [79,80]. (XLS)

**Table S2** List of protein-protein interactions. Physical protein-protein interactions between yeast transcriptional regulators extracted from BioGRID database [81]. (XLS)

**Table S3** List of protein-DNA interactions. Physical protein-DNA interactions were extracted from YeastRACT database [82]. (XLS)

**Table S4** Transcriptional effect of protein-DNA interactions. (XLS)

**Text S1** ODEs and parameter values for Fig. 6, and the method used in translating bistable network models into the experimental data mining format. (PDF)

### Acknowledgments

We are grateful to M. Feinberg for the preliminary version of the Chemical Reaction Network Toolbox v2.0 and for many useful discussions, and to G. Craciun, J. Stelling, A. P. Arkin, J. Paulsson, and J. J. Collins for their comments as well. DSG was jointly mentored by GDS and EG.

### Author Contributions

Conceived and designed the experiments: DSG GDS EG. Performed the experiments: DSG MKMG GDS. Analyzed the data: DSG MKMG. Contributed reagents/materials/analysis tools: DSG GDS. Wrote the paper: DSG MKMG GDS EG.

10. Ninfa AJ, Mayo AE (2004) Hysteresis vs. graded responses: the connections make all the difference. *Sci STKE* 2004: pe20.
11. Guidi G, Goldbeter A (1997) Bistability without hysteresis in chemical reaction systems: a theoretical analysis of irreversible transitions between multiple steady states. *J Phys Chem A* 101: 9367–9376.
12. Pomerening JR, Sontag ED, Ferrell JE (2003) Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2. *Nat Cell Biol* 5: 346–51.
13. ShaW, Moore J, Chen K, Lassaletta AD, Yi CS, et al. (2003) Hysteresis drives cell-cycle transitions in *Xenopus laevis* egg extracts. *Proc Natl Acad Sci USA* 100: 975–80.
14. Ozbudak EM, Thattai M, Lim HN, Shraiman BI, van Oudenaarden A (2004) Multistability in the lactose utilization network of *Escherichia coli*. *Nature* 427: 737–40.
15. Tyson JJ, Chen KC, Novák B (2001) Network dynamics and cell physiology. *Nat Rev Mol Cell Biol* 2: 908–16.
16. Feinberg M (1987) Chemical reaction network structure and the stability of complex isothermal reactors—I. The Deficiency Zero and Deficiency One Theorems. *Chem Eng Sci* 42: 2229–2268.
17. Feinberg M (1988) Chemical reaction network structure and the stability of complex isothermal reactors. II: Multiple steady states for networks of deficiency one. *Chem Eng Sci* 43: 1–25.
18. Ellison P, Feinberg M (2000) How catalytic mechanisms reveal themselves in multiple steady-state data: I. Basic principles. *J Mol Catal A-Chem* 154: 155–167.
19. Kaltenbach HM, Dimopoulos S, Stelling J (2009) Systems analysis of cellular networks under uncertainty. *FEBS Lett* 583: 3923–30.
20. Conradi C, Saez-Rodriguez J, Gilles ED, Raisch J (2005) Using chemical reaction network theory to discard a kinetic mechanism hypothesis. *Syst Biol* 152: 243–8.
21. Otero-Muras I, Banga JR, Alonso AA (2009) Exploring multiplicity conditions in enzymatic reaction networks. *Biotechnol Prog* 25: 619–31.
22. Craciun G, Tang Y, Feinberg M (2006) Understanding bistability in complex enzyme-driven reaction networks. *Proc Natl Acad Sci USA* 103: 8697–702.
23. Siegal-Gaskins D, Grotewold E, Smith GD (2009) The capacity for multistability in small gene regulatory networks. *BMC Syst Biol* 3: 96.
24. Ellison P (1998) The advanced deficiency algorithm and its applications to mechanism discrimination [PhD thesis]. Rochester (New York): Department of Chemical Engineering, University of Rochester.
25. Conradi C, Flockerzi D, Raisch J, Stelling J (2007) Subnetwork analysis reveals dynamic features of complex (bio)chemical networks. *Proc Natl Acad Sci USA* 104: 19175–80.
26. Flockerzi D, Conradi C (2008) Subnetwork analysis for multistationarity in mass action kinetics. *J Phys: Conf Ser* 138: 012006.
27. Miller CA, Beard DA (2008) The effects of reversibility and noise on stochastic phosphorylation cycles and cascades. *Biophys J* 95: 2183–92.
28. Saez-Rodriguez J, Hammerle-Fickinger A, Dalal O, Klamt S, Gilles ED, et al. (2008) Multistability of signal transduction motifs. *IET Syst Biol* 2: 80–93.
29. Keller AD (1995) Model genetic circuits encoding autoregulatory transcription factors. *J Theor Biol* 172: 169–85.
30. Hasty J, Isaacs F, Dolnik M, McMillen D, Collins JJ (2001) Designer gene networks: Towards fundamental cellular control. *Chaos* 11: 207–220.
31. Isaacs FJ, Hasty J, Cantor CR, Collins JJ (2003) Prediction and measurement of an autoregulatory genetic module. *Proc Natl Acad Sci USA* 100: 7714–9.
32. Becskei A, S  raphin B, Serrano L (2001) Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J* 20: 2528–35.
33. Cinquin O, Demongeot J (2002) Positive and negative feedback: striking a balance between necessary antagonists. *J Theor Biol* 216: 229–41.
34. Cherry JL, Adler FR (2000) How to make a biological switch. *J Theor Biol* 203: 117–33.
35. Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403: 339–42.
36. Fran  ois P, Hakim V (2004) Design of genetic networks with specified functions by evolution in silico. *Proc Natl Acad Sci USA* 101: 580–5.
37. Lipshtat A, Loinger A, Balaban NQ, Biham O (2006) Genetic toggle switch without cooperative binding. *Phys Rev Lett* 96: 188101.
38. Buchler NE, Louis M (2008) Molecular titration and ultrasensitivity in regulatory networks. *J Mol Biol* 384: 1106–19.
39. Sedlak TW, Oltvai ZN, Yang E, Wang K, Boise LH, et al. (1995) Multiple Bcl-2 family members demonstrate selective dimerizations with Bax. *Proc Natl Acad Sci USA* 92: 7834–8.
40. Basak S, Shih VFS, Hoffmann A (2008) Generation and activation of multiple dimeric transcription factors within the NF-  B signaling system. *Mol Cell Biol* 28: 3139–50.
41. Monod J, Jacob F (1961) Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb Symp Quant Biol* 26: 389–401.
42. Ramakrishnan N, Bhalla US (2008) Memory switches in chemical reaction space. *PLoS Comput Biol* 4: e1000122.
43. Ma W, Trusina A, El-Samad H, Lim WA, Tang C (2009) Defining network topologies that can achieve biochemical adaptation. *Cell* 138: 760–73.
44. Sabouri-Ghomi M, Ciliberto A, Kar S, Nov  k B, Tyson JJ (2008) Antagonism and bistability in protein interaction networks. *J Theor Biol* 250: 209–18.
45. Bundschuh R, Hayot F, Jayaprakash C (2003) Fluctuations and slow variables in genetic networks. *Biophys J* 84: 1606–15.
46. Stamatakis M, Mantzaris NV (2009) Comparison of deterministic and stochastic models of the lac operon genetic network. *Biophys J* 96: 887–906.
47. Kepler TB, Elston TC (2001) Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys J* 81: 3116–36.
48. Blake WJ, Kaern M, Cantor CR, Collins JJ (2003) Noise in eukaryotic gene expression. *Nature* 422: 633–7.
49. Samoilov M, Plyasunov S, Arkin AP (2005) Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proc Natl Acad Sci USA* 102: 2310–5.
50. Arkin AP, Ross J, McAdams HH (1998) Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics* 149: 1633–48.
51. Gillespie DT (2007) Stochastic simulation of chemical kinetics. *Annu Rev Phys Chem* 58: 35–55.
52. Jenal U, Hengge-Aronis R (2003) Regulation by proteolysis in bacterial cells. *Curr Opin Microbiol* 6: 163–72.
53. Johnson PR, Swanson R, Rakhilina L, Hochstrasser M (1998) Degradation signal masking by heterodimerization of MAT  2 and MATA1 blocks their mutual destruction by the ubiquitin-proteasome pathway. *Cell* 94: 217–27.
54. Kelemen JZ, Ratna P, Scherrer S, Becskei A (2010) Spatial epigenetic control of mono- and bistable gene expression. *PLoS Biol* 8: e1000332.
55. Soranzo N, Altafini C (2009) ERNEST: a toolbox for chemical reaction network theory. *Bioinformatics* 25: 2853–4.
56. Chothia C, Gough J, Vogel C, Teichmann SA (2003) Evolution of the protein repertoire. *Science* 300: 1701–3.
57. Force A, Cresko WA, Pickett FB, Proulx SR, Amemiya C, et al. (2005) The origin of subfunctions and modular gene regulation. *Genetics* 170: 433–46.
58. Lynch M, Conery JS (2003) The origins of genome complexity. *Science* 302: 1401–4.
59. Buljan M, Frankish A, Bateman A (2010) Quantifying the mechanisms of domain gain in animal proteins. *Genome Biol* 11: R74.
60. Babushok DV, Ostertag EM, Kazazian HH (2007) Current topics in genome evolution: molecular mechanisms of new gene formation. *Cell Mol Life Sci* 64: 542–54.
61. Koonin EV, Aravind L, Kondrashov AS (2000) The impact of comparative genomics on our understanding of evolution. *Cell* 101: 573–6.
62. Veening JW, Smits WK, Kuipers OP (2008) Bistability, epigenetics, and bet-hedging in bacteria. *Annu Rev Microbiol* 62: 193–210.
63. Jablonka E, Lamb M (1998) Epigenetic inheritance in evolution. *J Evol Biol* 11: 159–183.
64. Bishop AL, Rab FA, Sumner ER, Avery SV (2007) Phenotypic heterogeneity can enhance rare-cell survival in ‘stress-sensitive’ yeast populations. *Mol Microbiol* 63: 507–20.
65. Kussell E, Leibler S (2005) Phenotypic diversity, population growth, and information in fluctuating environments. *Science* 309: 2075–8.
66. Dubnau D, Losick R (2006) Bistability in bacteria. *Mol Microbiol* 61: 564–72.
67. Hollenhorst PC, Pietz G, Fox CA (2001) Mechanisms controlling differential promoter-occupancy by the yeast forkhead proteins Fkh1p and Fkh2p: implications for regulating the cell cycle and differentiation. *Genes Dev* 15: 2445–56.
68. Zhu G, Spellman PT, Volpe T, Brown PO, Botstein D, et al. (2000) Two yeast forkhead genes regulate the cell cycle and pseudohyphal growth. *Nature* 406: 90–4.
69. Hannenhalli S, Kaestner KH (2009) The evolution of Fox genes and their role in development and disease. *Nat Rev Genet* 10: 233–40.
70. Lamb TM, Mitchell AP (2003) The transcription factor Rim101p governs ion tolerance and cell differentiation by direct repression of the regulatory genes NRG1 and SMP1 in *Saccharomyces cerevisiae*. *Mol Cell Biol* 23: 677–86.
71. Bensen ES, Martin SJ, Li M, Berman J, Davis DA (2004) Transcriptional profiling in *Candida albicans* reveals new adaptive responses to extracellular pH and functions for Rim101p. *Mol Microbiol* 54: 1335–51.
72. Kumamoto CA, Vines MD (2005) Contributions of hyphae and hypha-co-regulated genes to *Candida albicans* virulence. *Cell Microbiol* 7: 1546–54.
73. Baumgartner U, Hamilton B, Piskacek M, Ruis H, Rottensteiner H (1999) Functional analysis of the zn(2)cys(6) transcription factors oaf1p and pip2p. different roles in fatty acid induction of beta-oxidation in *saccharomyces cerevisiae*. *J Biol Chem* 274: 22208–16.
74. Smith JJ, Ramsey SA, Marelli M, Marzolf B, Hwang D, et al. (2007) Transcriptional responses to fatty acid are coordinated by combinatorial control. *Mol Syst Biol* 3: 115.
75. Andrianantoandro E, Basu S, Karig DK, Weiss R (2006) Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol* 2: 2006.0028.
76. Ellis T, Wang X, Collins JJ (2009) Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat Biotechnol* 27: 465–71.
77. Shinar G, Feinberg M (2010) Structural sources of robustness in biochemical reaction networks. *Science* 327: 1389–91.
78. Bailey JE (2001) Complex biology with no parameters. *Nat Biotechnol* 19: 503–4.
79. Harbison C, Gordon D, Lee T, Rinaldi N, Macisaac K, et al. (2004) Transcriptional regulatory code of a eukaryotic genome. *Nature* 431: 99–104.
80. Drobn   E, Bialkova A, Subik J (2008) Transcriptional regulators of seven yeast species: Comparative genome analysis - review. *Folia Microbiol* 53: 275–287.

81. Breitkreutz BJ, Stark C, Reguly T, Boucher L, Breitkreutz A, et al. (2008) The BioGRID interaction database: 2008 update. *Nucleic Acids Res* 36: D637–D640.
82. Teixeira MC, Monteiro P, Jain P, Tenreiro S, Fernandes AR, et al. (2006) The YEASTRACT database: a tool for the analysis of transcription regulatory associations in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 34: D446–D451.