The evolution of central dogma of molecular biology: a logic-based dynamic approach

Mohieddin Jafari[1], Mehdi Mirzaie[2]

[1]Drug Design and Bioinformatics Unit, Medical Biotechnology Department, Biotechnology Research Center, Pasteur Institute of Iran, 69 Pasteur St, PO Box 13164, Tehran, Iran

[2]Department of Applied Mathematics, Faculty of Mathematical Sciences, Tarbiat Modares University, Jalal Ale Ahmad Highway, PO Box 14115-134, Tehran, Iran

*Corresponding Author:

Mohieddin Jafari,

Faculty Member, Pasteur Institute of Iran,

69, Pasteur St., 13164 Tehran, Iran.

www.jafarilab-pasteur.com

m_jafari@pasteur.ac.ir

mjafari@ipm.ir

Tel: +98 21 6411 2466

Fax: +98 21 66480780

**Co-Corresponding authors:

Mirzaie.mhd@gmail.com

Biological notes

Mohieddin Jafari is an assistant professor in the Biotechnology Research Center of Pasteur Institute of Iran. The subject of his research includes systems and network biology, biological data mining and proteomics.

Mehdi Mirzaie, is an assistant professor in the Department of Applied Mathematics at Tarbiat Modares University. His research field ranges broadly from structural bioinformatics and biostatistics to network science.

Pasteur Institute of Iran (IPI) was founded in the heart of Iran's capital, Tehran in the 1920s to pave the way for advanced research and to provide innovative programs in basic and applied medical sciences. Since 1983, Pasteur Institute of Iran offers a broad range of teaching activities in M.Sc. and Ph.D in different areas of medical and pharmaceutical biotechnology. It meets the specialized and scientific health demands of the local community.

Tarbiat Modares University (TMU) was founded in 1982 in Tehran, Iran. The main mission of this university is to train academic staff as well as researchers for universities and higher education centers in Iran.

Abstract

The age of the central dogma of molecular biology exceeded a half of a century. This biological axiom is grown up and now is more complicated. In this review, we renewed and re-expressed the central dogma as a mathematical model. It showed that the enhancement of the complexity kept pace with the gaining robustness.

Keywords

Central Dogma, Molecular Biology, Boolean Modeling, Robustness Analysis, Attractor State, State Transition Graph

Keynotes

- The central dogma of information flow is one of two main axiom in molecular biology.
- Boolean Modeling is useful to shed light on the conceptual notions in Biology.
- Although the molecular biology gets more complicated but it maintains its logical robustness.

Introduction

Biology, in this contemporary era, is an edifice consist of two main axioms as building materials: Darwin's theory of evolution and the central dogma of information flow in molecular biology [1]. Although, the central dogma skeleton i.e. DNA-> RNA-> Protein remain unchanged during the past half century, the complexity of this information flow is surprisingly increased. From discovering diverse species of RNA with distinguishing roles to dynamic features of DNA, mRNA and Protein cell to cell and over time, all take part in this complexity [2-4].

In 1965, the pioneering work of Jacob and Monod had shown that DNA was transcribed to RNA, which was translated into protein, and that the rate of transcription was controlled by a feedback loop in which protein levels regulated the activity of the transcriptional complex [5]. This vignette was illustrated in details in three kingdom of life by Francis Crick [6].

By the now, the flow of information was understood to be much more complex. DNA itself had been discovered not to be constant, and that genes could be permanently silenced. Proteins also carry information. Post-translational chemically modification can change a protein from active to inactive or vice versa. Ubiquitin regulate protein levels by proteasome complex, the cellular recycling bin. Genes can also be transcribed in alternate forms, or splice variants, in which entire exons may be omitted or alternate exons used instead. This process can produce proteins lacking or adding particular functions. The latest discovery is that processing of mRNA which can silence entire sets of genes by degrading the message at mRNA level and facilitating gene silencing at DNA level [7].

In this study we want to compare these two different abstraction of information flow in 1965 and present-day by a dynamic modeling approach. We would like to know what is could be expected when this conceptual wiring diagram changes over time? Which dynamic behavior is known right now and which one has predictive value? At the same time with addressing these kinds of questions, we also try to clear logic-based modeling potentials to map out perceptual design of scientific notions. Although, the simplest species of these modeling i.e. Boolean network is known as qualitative modelling approach [8], it is previously showed that complex dynamic behaviors could be extracted such as bistability [9]. This could be originated from the capability of step-function to approximate sigmoidal kinetic function of molecular binding e.g. enzymes and exist of thresholds in the most of biological processes [10-12]. We think that this approach would be helpful to understand potential dynamic behavior across the genome.

Let us to describe central dogma models which used in this study. It should be mentioned that we simulate all the possible behavior of the single genes and gene products and the combinatory behavior of the multiple genes are not examined in this study.

Boolean dynamic models

The initial representation of central dogma contains four different components i.e. DNA, mRNA, Protein and Activator (Fig. 1). Based on activation or inhibition effect of Protein on the mRNA transcription, two possible versions of this model could be imagine (Table 1). In this model, the long half-life components i.e. DNA, Protein and Activator are distinguished by a self-loop in Boolean rules used for the model. It is crucial to accent that the names used in the model do not fully comply the molecular biology meaning. As described in the second column of the table 1, for example, the DNA mean a gene which could be transcribed to the coding RNA not more. These definitions are based on the concept proposed by Francis Crick [6] when the splicing, post-transcriptional and -translational modifications are not understood as well as now. For instance, when Jacob and Monod (R was described the genetic regulatory mechanism, they used Allolactose or cAMP just as activating factor of LacI and CRP, therefore we define this component generally in this model. Also, the Protein here mean the active and functional protein not any sequence of amino acids translated by means of ribosome.

Table 1: Boolean rules governing the state of the 4-node network of the 1965 proposed central dogma depicted in Fig. 1.

		Logical rule	
Node	***Definition***	V1: Activatory gene expression	V2: Inhibitory gene expression
Activator	The chemical molecules which facilitate active protein generation before or after translation	Activator	
DNA	The genes transcribed to coding RNA	DNA	
mRNA	The coding RNA	DNA & Protein	DNA & !Protein
Protein	The functional active protein	mRNA & Activator | Protein	

Obviously, in the current view, the central dogma is more complicated in respect of new components and relationships (Figure 2 and Table 2). The main difference is considering the concept of "turnover" of the dynamic molecules. The DegRNA and Ubiquitin nodes are used to represent degradation of mRNA and Protein in this model. The other modification is adding the miRNA into the model as a representative of all RNA species which affect protein production or transferring information to protein level negatively. The meaning of DNA and Activator are also modified in terms of emergence of epigenetics and post-translational modification concepts. The DNA label is used in this model to explain the non-silenced genes which is detectable by RNA polymerase and the Activator is connected to Protein in a feedback loop to delineate the main role of Protein in the Activator turnover. The activity of exogenous and endogenous Activator usually affected by Proteind as an enzymes, transporters and channels. Based on three edges in this model with two opposite direction i.e. activation and inhibition between Protein-mRNA, Protein-miRNA and miRNA-DNA, the eight different versions of Boolean rules assumed to govern the nodes' states (Table 2).

Table 2: Boolean rules governing the state of the 7-node network of the present-day proposed central dogma depicted in Fig. 2.

		Logical rule							
Node	***Definition***	V1: Activatory	V2: mRNA expression inhibition	V3: miRNA expression Inhibition	V4: Gene silencing	V6: Gene silencing & miRNA expression inhibition	V5: Gene silencing & mRNA expression inhibition	V7: miRNA & mRNA expression inhibition	V8: Inhibitory
Activator	The chemical molecules or residue which facilitate active protein generation before or after translation	Protein | Activator							
DegRNA	The protein complex which degrades RNA species	Protein & !DegRNA							
DNA	The non-silenced genes which is detectable by RNA polymerase	miRNA & Protein | DNA			**!miRNA** & Protein | DNA			miRNA & Protein | DNA	**!miRNA** & Protein | DNA
miRNA	A RNA species which is involved in regulation of gene silencing (for simplification, all other RNA species which has negative effect on protein production is shown with the same label)	DNA & Protein | mRNA & !DegRNA		DNA & **!Protein** | mRNA & !DegRNA	DNA & Protein | mRNA & !DegRNA	DNA & **!Protein** | mRNA & !DegRNA	DNA & Protein | mRNA & !DegRNA	DNA & **!Protein** | mRNA & !DegRNA	
mRNA	The mature messenger RNA that is ready for translation (for simplification, all other RNA species which has positive effect on protein production is shown with the same label)	DNA & Protein & !DegRNA	DNA & **!Protein** & !DegRNA	DNA & Protein & !DegRNA	DNA & Protein & !DegRNA		DNA & **!Protein** & !DegRNA		
Protein	The functional active protein	!miRNA & mRNA & Activator & !Ubiquitin | Protein							
Ubiquitin	A post-translational modification that is involved in negative regulation of protein amount	Protein & !Ubiquitin							

Model simulation and identification of attractors

In this study, we performed R package called BoolNet [13] to identify synchronous attractors by exhaustive search and complex asynchronous attractors. The synchronous analysis was performed on all ten proposed versions separately whereas the asynchronous analysis was done just on present-day model versions. In the synchronous simulation, the state-transition graphs are studied and the attractors along with the size of the basin of attraction abstracted in the tables. For the asynchronous simulation the complex attractor are represented in a figure which illustrate the closeness centrality of each state (node) by the node size. This figure and the centrality analysis is performed using Gephi software (v.0.8.2) [14]. Also, we calculated the probability of reaching states using Markov chain simulations. In addition to all eight versions of new central dogma model, this analysis was also performed on the following probabilistic Boolean networks shown in Table 3.

Table 3: Boolean rules governing the state of the present-day proposed central dogma model in a Boolean model included probabilities.

Targets	*Factors*	*Probabilities*
Activator	Protein | Activator	1
DegRNA	Protein & !DegRNA	1
DNA	miRNA & Protein	0.1
	!miRNA & Protein	0.9
miRNA	DNA & Protein | mRNA & !DegRNA	0.8
	DNA & !Protein | mRNA & !DegRNA	0.2
mRNA	DNA & Protein & !DegRNA	0.8
	DNA & !Protein & !DegRNA	0.2
Protein	!miRNA & mRNA & Activator & !Ubiquitin | Protein	1
Ubiquitin	Protein & !Ubiquitin	1

Finally, the plausibility of our constructed models are examined using three different measures of robustness to noise and mismeasurements. This analysis was performed only on the present-day model of central dogma. The normalized Hamming distance which indicates the fraction of different bits between each state and the corresponding perturbed copies was obtained after 100 randomly generated copies [15]. The Gini index which is an index for in homogeneity in the in-degree property of the nodes was calculated in the state-transition graph [16]. The long-term behavior also was evaluated to compare the attractor profile in original and perturbed copies of the models [17]. All of these measures were assessed based on the statistical test (z-test) and randomized version of the networks to demonstrate the statistical significance of them by p-value.

Results and discussion

We aim to determine whether the sustainable oscillatory behavior or single state attractor are observed in the synchronous and asynchronous simulation of the 10 mentioned versions of central dogma models. Also, different measures such as Hamming distance and Gini index were used to illustrate robustness of these dynamic modeling.

Synchronous modeling

We found by simulation that with synchronous update, the old conceptual model possesses 11 simple attractors or fixed points out of 16 possible initial conditions (Table 4). It means 69% out of all states are steady state in which the eight and seven fixed points are observed in activatory and inhibitory model respectively. Four out of these fixed points occurs in both activatory and inhibitory gene expression versions (A, B, C and D). A null fixed point (state A) in which all the 4 nodes are in the OFF state. This SS is reachable from 12.5% (2 out of 16). All of three other SSs (state B, C and D) are not included any transcribable genes, therefore mRNA, Protein and Activator remain unchanged. Among four other fixed points namely E, F, G and H which are only reached in the activatory model, the G state has larger basin of attraction (1110→1101→1111) and it is a formal statement of the provider model at 1965. The G and H states are fixed points in this model because mRNA and Protein turnover does not considered in this model. The J state is another formal statement of the provider model when the protein negatively regulate mRNA transcription. This state has the largest basin of attraction in the inhibitory model (1100→1110→1111→1101). Similar to state G and H, the I state is also the artifact of ignoring mRNA turnover. There is no any limit cycle observed in these models based on assumptions that have been available in the past. Also, three states namely D, H and K are the specific states to the aspect of conventional central dogma modeling and obviously not a perfect match with today's knowledge of molecular biology. For instance, we could not reach any steady state in the attractor analysis of present-day model which the state of Activator/DNA/mRNA are zero and Protein is one same as D state.

The steady states of the present-day model of central dogma obtained by the synchronous updating method is more complicated and consistent with the current knowledge of molecular biology. Overally, 14% i.e. 18 out of 128 possible initial states fall in the attractor set (Table 4). Similar to pervious model, the null fixed point (state A) and only Activator ON state (state B) are represented in all eight versions with the same basin of attraction. There are six other fixed points that cluster our version into three class. The C and D states are observed in the mRNA expression inhibition (V2), Gene silencing & mRNA expression inhibition (V5), miRNA & mRNA expression inhibition (V7) and Inhibitory (V8) models. Although RNA species turnover is considered in this model, we see remaining of these rapidly degraded molecules in the attractor. This situation could be matched to amino acid poverty that we did not include into model. On the other hand, in this model the Protein used to explain “The functional active proteins” (See table 1), therefore this attractor could explain the situation which the translation occurs but the active proteins finally not achieved. The C state appropriated 12.5% (16 out of 128) of all states in all four mentioned version, whereas the D state is not the same. The D state is rarely found in the simulation of mRNA expression inhibition (V2) and Gene silencing & mRNA expression inhibition (V5) models i.e. 3 out of 128, but it contains larger basin of attraction in miRNA & mRNA expression inhibition (V7) and Inhibitory (V8) models i.e. 14 out of 128. It suggested that negatively regulated mRNA together with positively regulated miRNA acts like Activator for Protein production. The D state is the only state that is specific to present-day model of central dogma and previously it could not be predicted. There is not an attractor state which Activator, DNA and mRNA are presented but the

Protein is not produced (See Table 4). Based on the E and F states, it seems that without basal expression of Protein the Activatory (V1) and Gene silencing system (V4) have similar behavior and the system stops. The G and H states are the analogous of C and D states but in the miRNA expression inhibition (V3) and Gene silencing and miRNA expression inhibition (V6) versions of the model. Unlike to pervious model, there are 10 different limit cycles play the attractor role in this simulation. The I and J states are the common cycles between Activatory (V1), mRNA expression inhibition (V2), miRNA expression inhibition (V3) and mRNA & miRNA expression inhibition (V7). In these attractors, the long half-life components of model i.e. Activator, DegRNA, Protein and Ubiquitin oscillate concurrently. These rare oscillations occur at the end of information flow of molecular biology and they are independent of mRNA and miRNA alterations. This situation can be happened in the stationary phase of growth. The K and L states are the typical expected dynamic behavior that usually observed in the system with activatory regulated mRNA (V1 & V4). The most of genes usually follow this kind of behavior especially in the log phase of growth. The M,N,O and P states also are similar limit cycles to the K and L with the exception that the expression of miRNA are inhibited in M and N (V3 & V6) and mRNA expression supposed to inhibited in the O and P states (V2 & V5). The final two limit cycles i.e. Q and R are special states occur in simultaneous inhibition of mRNA and miRNA expression with different effects on Protein production (V7 & V8). A common trait in all of the limit cycles is constant presence of Protein and Activator. It is expected that steady states are related to long half-life components of the biological systems.

Table 4: The attractors of the Boolean model depicted in Fig. 1 obtained from the synchronous update method. In each version of the model, the number of single SSs and corresponding normalized basin of attractions are represented. The color gradient from red, yellow to green illustrate lowest to highest value. The attractors namely A to K are encoded in the following order: Activator/DNA/mRNA/Protein. The starred states i.e. D*, H* and K* denote those states which are specific to this central dogma modeling.

Central Dogma	Description	No. SSs with 1	A 0000	B 1000	C 1001	D* 0001	E 0100	F 1100	G 1111	H* 0111	I 0110	J 1101	K* 0101
V1	Activatory	8	12.5	6.25	18.75	12.5	12.5	6.25	18.75	12.5	-	-	-
V2	Inhibitory	7	12.5	6.25	18.75	12.5	-	-	-	-	12.5	25	12.5
Activator/DNA/mRNA/Protein													

Asynchronous modeling

Using the asynchronous Boolean dynamic approach, we analyzed the present-day model of the central dogma to explore different attractors. Similar to fixed points found in the synchronous updating model, there are eight simple attractors (states A to H) with the same profile (Table 6). But the complex/loose attractors are different from limit cycles of synchronous approach and get more complicated. There are five complex attractors with different size of nodes and edges (states I to M) (See Table 6 and Fig. 2). The I, J and K states are the small loose attractors with 4 nodes and 4 bi-directional edges, and homogenous node properties. The I state is the combination of I and J states of the Table 5 and it happens when there is not any gene silencing events. The J state is the common attractor between inhibited mRNA expression versions i.e. 2, 6 and 7 and the K state occurs in the mRNA and miRNA expression inhibition versions of model. In the larger complex attractors i.e. L and M, the nodes play different role in this graph as shown in Fig. 2. The node size is depicted based on closeness centrality. The larger nodes means the nodes that are more accessible or close to other nodes. In both L and M attractors, the state nodes 1111111, 1111110, 1011110 and 1011011 explain more the dynamic behavior of the system whose mRNA expression is not inhibited by Protein.

Table 5: The attractors of the Boolean model depicted in Fig. 1 obtained from the synchronous update method. In each version of the model, the number of single SSs, SSs with multiple states (limit cycle) and corresponding normalized basin of attractions are represented. The color gradient from red, yellow to green illustrate lowest to highest value. The attractors namely A to R are encoded in the following order: Activator/DegRNA/DNA/miRNA/mRNA/Protein/Ubiquitin. The starred state i.e. D* denotes the state which is specific to this central dogma modeling.

Central Dogma 2009	System name	No. SSs with 1 state	No. SSs with multiple states	A	B	C	D*	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
				0000000	1000000	0011100	1011100	0010000	1010000	0011000	1011000	1000010 > 1100011	1100010 > 1000011	1011010 > 1111111	1111110 > 1011011	1010010 > 1110111	1110110 > 1010011	1011010 > 1111011	1111010 > 1011011	1010010 > 1110011	1110010 > 1010011
V1	Activator y	4	4	12.5	10.94	-	-	12.5	10.94	-	-	5.47	4.69	22.66	20.31	-	-	-			
V2	mRNA expressi on inhibitio n	4	4	12.5	10.94	12.5	2.34	-	-	-	-	5.47	4.96	-	-	-	-	31.25	20.31		
V3	miRNA expressi on inhibitio n	4	4	12.5	10.94	-	-	-	-	12.5	10.94	5.47	4.69	-	-	22.66	20.31	-			
V4	Gene silencing	4	2	12.5	10.94	-	-	12.5	10.94	-	-	-	-	28.12	25	-	-	-			
V5	Gene siler	4	2	12.5	10.94	12.5	2.34	-	-	-	-	-	-	-	-	-	-	36.72	25		
V6	Gene siler	4	2	12.5	10.94	-	-	-	-	12.5	10.94	-	-	-	-	28.12	25	-			
V7	miRNA & i	4	4	12.5	10.94	12.5	10.94	-	-	-	-	5.47	4.69	-	-	-	-	-		22.66	20.31
V8	Inhibitor y	4	2	12.5	10.94	12.5	10.94	-	-	-	-	-	-	-	-	-	-	-		28.12	25
Activator/DegRNA/DNA/miRNA/mRNA/Protein/Ubiquitin																					

Table 6: The attractors of the Boolean model depicted in Fig. 1 obtained from the asynchronous update method. In each version of the model, the number of simple attractor, complex/loose attractor and corresponding edge numbers in complex/loose attractor are represented. The binary digits of the attractors (namely A to M) from left to right represent the state of the nodes Activator, DegRNA, DNA, miRNA, mRNA, Protein and Ubiquitin respectively.

Central Dogma 2009	System name	Simple attractor								Complex/loose attractor				
		A	B	C	D	E	F	G	H	I	J	K	L	M
		0000000	1000000	0011100	1011100	0010000	1010000	0011000	1011000	1100011, 1000011, 1100010, 1000010	1111011, 1011011, 1111010, 1011010	1110011, 1010011, 1110010, 1010010	1111111, 1011111, 1111011, 1011011, 1111110, 1011110, 1111010, 1011010	1111111, 1011111, 1110111, 1010111, 1111011, 1011011, 1110011, 1010011, 1111110, 1011110, 1110110, 1010110, 1111010, 1011010, 1110010, 1010010
V1	Activatory	✓	✓			✓	✓			8			20	
V2	mRNA expression inhibition	✓	✓	✓	✓					8	8			
V3	miRNA expression inhibition	✓	✓					✓	✓	8				48
V4	Gene silencing	✓	✓			✓	✓						20	
V5	Gene sile	✓	✓	✓	✓						8			
V6	Gene sile	✓	✓					✓	✓					48
V7	miRNA &	✓	✓	✓	✓					8	8			
V8	Inhibitory	✓	✓	✓	✓							8		
Activator/DegRNA/DNA/miRNA/mRNA/Protein/Ubiquitin														

In order to identify the probability of the reaching attractor after a large number of iterations, we performed Markov chain simulation for present-day central dogma model. As shown in table 7, the states reached at the end of simulation i.e. attractors are listed for all versions of this model. In addition, the new version called V1_V8 is added to this analysis to allow the systematic study of global network dynamics which is robust in the face of uncertainty. In all versions, reaching null fixed point and fixed point with only ON Activator are equally probable. Based on the activation/inhibition functions layout the other states happen meaningfully different. Except these two states (0000000, 1000000), as expected, the most likely probability in Activatory version (V1), Gene silencing (V4) and V1_V8 versions is full fixed point (1111111). The state 1111011 and

1110111 are the most probable states in the inhibited mRNA expression versions (V2 & V5) and the inhibited miRNA expression versions (V3 & V6) respectively. The 1110011 is the expected probable state for the V7 and V8 versions of the model due to inhibited expression of mRNA and miRNA together.

Table 7: The absorption probabilities for the Markov chain corresponding to the Boolean model depicted in Fig. 1 in all versions. These probabilities reached after 1000 iterations. The last column represent the absorption probabilities of the probabilistic Boolean network of the present-day central dogma (See table 2).

Activator/DegRNA/DNA/miRNA/mRNA/Protein/Ubiquitin	V1	V2	V3	V4	V5	V6	V7	V8	V1_V8
0000000	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13
0010000	0.13			0.13					0.06
0010100									0.02
0011000			0.13			0.13			0.04
0011100		0.13			0.13		0.13	0.13	0.01
1000000	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11
1000010	0.02	0.02	0.02				0.02		
1000011	0.02	0.02	0.02				0.02		
1010000	0.11			0.11					
1010010			0.11			0.13	0.11	0.13	0.04
1010011			0.11			0.13	0.11	0.13	0.03
1011000			0.11			0.11			
1011010	0.11	0.14		0.13	0.16				0.14
1011011	0.11	0.11		0.13	0.13				0.10
1011100		0.02			0.02		0.11	0.11	
1100010	0.03	0.03	0.03				0.03		
1100011	0.04	0.04	0.04				0.04		
1110010							0.09	0.13	0.01
1110011							0.12	0.16	0.01
1110110			0.09			0.13			0.02
1110111			0.12			0.16			0.03
1111010		0.09			0.13				0.02
1111011		0.17			0.21				0.03
1111110	0.09			0.13					0.08
1111111	0.12			0.16					0.14

Robustness of the reconstructed model

To assess plausibility of the central dogma models, we tested the robustness to noise and mismeasurements of all versions of initial and up-to-date model of central dogma. As shown in Table 8, the present-day model of central dogma is more robust than the previous model. The

Hamming distance is significantly lower in all eight versions of new model. It means that by applying noise on the network states in these models, the more similar state to unperturbed versions reached than the 1965 model and the new conceptual model of central dogma is considerably more robust to noise in the states than the randomly generated models. To test long-term dynamic behavior, the fraction of pairs of states and perturbed copies that yield the same attractor was assessed in both models. In general, these fractions are higher in the new model in comparison with the traditional model. In the versions which the expression of mRNA inhibited and/or the gene silencing is happened i.e. V2, V4 and V6, the models are slightly more robust to perturbations based on this measure. Although the long-term dynamic behavior does not generally show the large fractions, this measure has been greater value and its p-value smaller by extending information flow of central dogma. It seems that including more details at the molecular level specific to each gene increase this value more. Finally a measure of inhomogeneity i.e. Gini index was assess in the state transition graph of all versions of both model and it was compared to randomized graph to compute p-value. As expected in biological networks, the present-day model significantly has many states with a low in-degree and a few states with a high in-degree in state transition graph. It means that the updated model is more consistent with reality in biology.

Table 8: The robustness analysis of the reconstructed models of central dogma.

	Hamming distance	P-value	Same Attractors	P-value	P-value of **the Gini Index**
1956					
V1	0.26	0.16	0.21	0.13	0.59
V2	0.26	0.21	0.31	0.41	0.61
2009					
V1	0.16	0	0.33	0.14	0.01
V2	0.16	0	0.51	0.09	0
V3	0.15	0	0.33	0.16	0
V4	0.15	0	0.57	0.03	0
V6	0.16	0	0.57	0.08	0
V5	0.16	0.01	0.44	0.13	0
V7	0.16	0.02	0.25	0.21	0
V8	0.16	0	0.35	0.07	0

Conclusion

Boolean network modeling helps us to depict the development of the central dogma models and its potential remarks. It is showed that in addition to predict reliable steady states, the present-day model of central dogma is more robust and insensitive to noise than previous version. Using this approach which is a more precise explanations of our thinking [18] is useful to improve and validate our imagination and knowledge of conceptual structure in biology.

References

1. Koonin EV. Does the central dogma still stand?, Biology direct 2012;7:27-27.
2. Lin Y, Elowitz MB. Central Dogma Goes Digital, Molecular cell 2016;61:791-792.
3. Prabakaran S, Hemberg M, Chauhan R et al. Quantitative profiling of peptides from RNAs classified as noncoding, Nature communications 2014;5:5429-5429.
4. Gerstein MB, Bruce C, Rozowsky JS et al. What is a gene, post-ENCODE? History and updated definition, Genome Research 2007;17:669-681.
5. Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins, Journal of molecular biology 1961;3:318-356.
6. Crick F. Central dogma of molecular biology, Nature 1970;227:561-563.
7. Systems Biology for Signaling Networks. New York, NY: Springer New York, 2010.
8. Saadatpour A, Albert R. A comparative study of qualitative and quantitative dynamic models of biological regulatory networks, EPJ Nonlinear Biomedical Physics 2016;4:5-5.
9. Robeva R, Murrugarra D. The spruce budworm and forest: a qualitative comparison of ODE and Boolean models, Letters in Biomathematics 2016;3:75-92.
10. Albert R, Saadatpour A. Boolean modeling of biological regulatory networks: a methodology tutorial, Methods (San Diego, Calif.) 2013;62:3-12.
11. Saadatpour A, Guo G, Orkin SH et al. Characterizing heterogeneity in leukemic cells using single-cell gene expression analysis, Genome Biology 2014;15:525-525.
12. Saadatpour A, Wang R-S, Liao A et al. Dynamical and Structural Analysis of a T Cell Survival Network Identifies Novel Candidate Therapeutic Targets for Large Granular Lymphocyte Leukemia, PLoS Computational Biology 2011;7:e1002267-e1002267.
13. Müssel C, Hopfensitz M, Kestler HA. BoolNet--an R package for generation, reconstruction and analysis of Boolean networks, Bioinformatics (Oxford, England) 2010;26:1378-1380.
14. Bastian M, Heymann S, M J. Gephi : An Open Source Software for Exploring and Manipulating Networks.
15. Shmulevich I, Kauffman SA. Activities and sensitivities in boolean network models, Physical review letters 2004;93:048701-048701.
16. Xiao Y, Dougherty ER. The impact of function perturbations in Boolean networks, Bioinformatics (Oxford, England) 2007;23:1265-1273.
17. Saadatpour A, Albert In, Albert Rk. Attractor analysis of asynchronous Boolean models of signal transduction networks, Journal of Theoretical Biology 2010;266:641-656.
18. Gunawardena J. Models in biology: 'accurate descriptions of our pathetic thinking', BMC biology 2014;12:29-29.

Figure legends

Fig. 1. (A) The 1965 and the present-day central dogma of molecular biology (B) The initial 4-node central dogma wiring diagram formed by the DNA, mRNA, Protein and Activator. (C) The present-day 7-node central dogma wiring diagram formed by the DNA, mRNA, Protein, Activator, miRNA, Ubiquitin and DegRNA. The directed edge —o denotes activation and inhibition in different versions of the models. (D) Network representation of the Boolean rules corresponding to this simple proposed model. (E) Network representation of the Boolean rules corresponding to this proposed model. For simplicity the first versions are only shown.

Fig. 2. The complex/loose attractors indicated in Table 6 namely (A) I, (B) J, (C) K, (D) L and (E) M with sorted node size based on closeness centrality.