The Third International Symposium on Optimization and Systems Biology (OSB'09)
Zhangjiajie, China, September 20–22, 2009
Copyright © 2009 ORSC & APORC, pp. 435–447

Analysis for Stimuli to Shoot Genes of Arabidopsis Thaliana Based on Logical Relationships*

Shudong Wang[1,†] Yan Chen[2] Qingyun Wang[3] Eryan Li[1]
Yansen Su[1] Dazhi Meng[2,‡]

[1] College of Information Science and Engineering, Shandong University of Science and Technology, Qingdao, Shandong 266510
[2] College of Applied Sciences, Beijing University of Technology, Beijing 100124
[3] College of Mathematics and Computer Science, GanNan Normal University, GanZhou, JiangXi 341000

Abstract The reverse construction and analysis of the networks of molecular interactions are essential for understanding their functions within cells. In this paper, a logic network model is constructed to investigate the complicated regulation mechanism of shoot genes of Arabidopsis Thaliana in response to stimuli. The dynamics of the complicated logic network is analyzed, discussed and simulated. The simulating results show that the logic network of the active genes of shoot eventually evolves into eleven attractors under the stimuli, where including five 1- periodic and six 2-periodic attractors. Our work provides valuable reference and guidance for biologists to understand and explain Arabidopsis' response to outside stimuli by experiments.

Keywords Systems Biology, Gene Network, Logic Analysis of Phylogenetic Profiles, Dynamics

1 Introduction

Reverse construction and analysis of gene or protein interactive networks based on massive genome data are essential to discover and understand the functions of various cell components. Traditional computational methods have just been developed to detect functional linkages between proteins or genes from the set of full sequenced genomes, such as Pearson Correlation coefficient, Euclidean and Hamming distances, mutual information, the hypergeometric distribution and shortest-path analysis[1-3]. However, traditional pairwise relationships cann't adequately illustrate the complexities that arise in cellular networks because of branching and alternate pathways. For instance, Shikimate 5-dehydrogenase

* This work is supported by Natural Science Foundation of China (Nos. 60874036, 60503002)
† wangshd2008@yahoo.com.cn
‡ dzhmeng07@yahoo.com.cn

(COG0169) is present if and only if 3-dehydroquinate dehydratase Ⅱ(COG0757) or 3-dehydroquinate dehydratase (COG0710) is present. An archaeal DNA-binding protein (COG1581) is present if and only if ATPase involved in DNA repair (COG0419) is present and a mismatch repair ATPase (MutS family, COG0249) is absent[4]. In order to further address the ubiquitous logical relationships mentioned above in biology, Bowers *et al*[4] proposed a computational approach for identifying detailed relationships among proteins on the basis of genomic data, which was applied into 4873 distinct orthologous protein families of 67 fully sequenced organisms, and identified 750,000 triplets from analysis of the original, unshuffled biological protein profiles in 2004. As an application of triplet logic analysis to 85 diffuse infiltrating gliomas quantified using oligonucleotide arrays. Bowers *et al*[5] explored the meaningful sets of genes that matched clinical outcomes in 2005. Zhang X *et al*[6] described a Bayesian modeling framework for combining phylogenetic profile data via a likelihood with Rosetta Stone data via a prior probability in 2006. The method made up the defect of inferring protein logical relationships using pairwise and triplet logic analysis on phylogenetic profiles solely. Zhang J X *et al*[7] described a three-way gene interaction model that captured the dynamic of co-expression relationships between two genes in 2007. However, the dynamical analysis of logical network wasn't mentioned in the above papers. In fact, the dynamical research of the network plays a very important part in the entire process of analyzing and understanding biological meanings. Kauffman *et al*[8] developed to describe dynamical properties of genetic regulatory networks using Boolean dynamic model. And in 2005, the biological experiments confirmed the method of Boolean dynamics proposed by Kauffman[9]. Li *et al*[10] introduced simple Boolean rules and discovered that the yeast cell-cycle network is extremly stable and there is one big attractor. Luscombe *et al*[11] systematically studied the topological properties of these genetic regulatory networks under different inside and outside conditions and showed that there are big differences between the actual and man-made gene networks. To investigate the reaction mechanism of shoot active genes of Arabidopsis thaliana, we construct the logical regulatory networks under different stimuli by utilizing LAPP (logic analysis of phylogenetic profiles) method and do randomized test to the gene logical regulatory networks. The results show that LAPP method is of the specificity to our research data. What's more, we discover that the active genes in shoot mostly are inactive in root under outside different stimuli from the data. Based on the above obtained gene logical network, the dynamics of the network is characterized and simulated. The simulating results show that the initial states eventually would evolve into eleven attractors. In the eleven attractors, there are five 1- periodic and six 2-periodic attractors. Our work provides valuable reference and guidance for biologists to understand and explain Arabidopsis response to outside stimuli by experiments.

2 Construction and Analysis of Logical network

With the example of a public set of Arabidopsis thaliana expression data provided by Department of Plant Biology, Southern Illinois University Carbondale, United States, we give a complete way of constructing gene expression regulatory networks

under different stimuli. Within the constraints of actual gene network, average degree of biological network, mutual information network and support of each combination, we can determine the threshold value and logic types of the logical relationships, and then filter the logical combinations of genes whose uncertain coefficients are less than the threshold value. Finally, by comparing with the results obtained from the random data, we make a difference analysis for the gene logical combinations from the original data.

2.1 Data Source and Processing

We use 47 samples of mRNA abundance from 25,090 genes of Arabidopsis detected by Affymetric HG133A chip. These original data scanning by laser chip are processed to expression profile data using Dchip6.0. These gene-expression data are detected in the special part (shoot) of Arabidopsis under different stimuli, including 47 local (shoot) acquired defense and resistance samples as follows (where hr denotes processing time unit):

6 shoot samples processed by Methyl Viologen: 0.5hr, 1hr, 3hr, 6hr, 12hr, 24hr;

7 shoot samples processed by Drought: 0.25hr, 0.5hr, 1hr, 3hr, 6hr, 12hr, 24hr;

6 shoot samples processed by Salt: 0.5hr, 1hr, 3hr, 6hr, 12hr, 24hr;

6 shoot samples processed by Cold: 0.5hr, 1hr, 3hr, 6hr, 12hr, 24hr;

4 shoot samples processed by Heat: 0.25hr, 0.5hr, 1hr, 3hr;

4 shoot samples processed by Heat Recovery: 1hr, 3hr, 9hr, 21hr;

7 shoot samples processed by Wounding: 0.25hr, 0.5hr, 1hr, 3hr, 6hr, 12hr, 24hr;

7 shoot samples processed by Mannitol300: 0.5hr, 1hr, 3hr, 6hr, 12hr, 24hr.

The 47 samples are the ratio of the expression profiles under the above different stimuli and no stimuli(see Appendix-1). In order to analyse the logical relationships among shoot genes of Arabidopsis under different stimuli, we discretize the actual ratio data into 0-1 values as follows. For any expression profile data, the values belonging to [0.51, 1.9] are denoted by 0 and the values belonging to (1.9, 1000)$\cup$[(0.000001, 0.51) by 1. The number 0 and 1 represent absence or presence of genes under outside stimuli respectively. In these 47 0-1 samples, we think the shoot genes including more Zeros are insensitive to stimuli responses, while those including more Ones are insensitive to logical relationships. Sixteen Arabidopsis shoot genes are selected from the above 25090 genes as active genes by data deleting, merging and active genes choosing. For the convenience of describing and studying, each of these 16 genes is assigned a sequence number according to the order, which is from 1 to16, see Table 1.

2.2 Determination of Thresholds

Bowers and Xin Zhang[4-6] utilized uncertainty coefficient (abbreviated as U Value) to define the logical relationships and logic types of the 1-, 2- and 3-order among gene A, B, C, D,

$$U(B|A)=\frac{H(B)+H(A)-H(A,B)}{H(B)} \tag{3}$$

$$U\left(C\middle|f\left(A,B\right)\right)=\frac{H\left(C\right)+H\left(f\left(A,B\right)\right)-H\left(C,f\left(A,B\right)\right)}{H\left(C\right)} \quad (4)$$

Table 1: List and Description of Arabidopsis Thaliana Shoot Genes' Responses to Stimuli

Gene No.	Gene	Function	Physical Process	Composition
1	AT1G21100	Oxygen Position - Methyltransferase	—	Cytosol
2	AT1G51850	Leucine-Rich Repeat Protein Kinase	phosphate-bearing amino acids on the protein	Endomembrane System
3	AT1G51860	Leucine-Rich Repeat Protein Kinase	phosphate-bearing amino acids on the protein	Endomembrane System
4	AT1G59590	ZCF37(Function unknown)	Unknown	Unknown
5	AT1G67810	Fe-S Metabolizing structural Domain Protein	—	Chloroplast
6	AT2G35910	Zinc Fingers Protein	Protein-Binding Combination Zinc ion Binding	—
7	AT2G36690	Oxidoreductase	Biosynthesis	—
8	AT2G39530	transmembrane protein	—	Velum
9	AT3G02140	Arabidopsis thalianaTMAC2 Gene（Unknown）	—	—
10	AT3G11340	UDP-Glucosyltransferase	Metabolic Pathway	Endomembrane System
11	AT3G46690	UDP-Glucosyltransferase	Metabolic Pathway	—
12	AT4G19370	Unknown	—	Endomembrane System
13	AT4G23030	Multidrug Efflux	Multidrug Transportation	Membrane
14	AT4G37370	Cytochorome P450	Metabolism、Biosynthesis、Degradation	—
15	AT5G01830	Ubiquitin-protein ligase	Ubiquitination	Ubiquitin Conjugation Complex
16	AT5G44990	Unknown	Unknown	Unknown

$$U\left(D\middle|g\left(A,B,C\right)\right)=\frac{H\left(D\right)+H\left(f\left(A,B,C\right)\right)-H\left(D,g\left(A,B,C\right)\right)}{H\left(D\right)} \quad (5)$$

where $H\left(X\right)$ is the entropy of X, $H\left(X,Y\right)$ is the joint entropy of X and Y. $U\left(B\middle|A\right)$ is the impact extent of some event B, given the occurrence of some other event A, which denotes the measure of statistic existence of the uncertainty logical relationship "$A\to B$" between gene A and B. It is similar for the 2-order logical relationships. We just need calculate $U\left(C\middle|f\left(A,B\right)\right)$ for existent probability of certain 2-order logical relationship among gene A, B, C, where function f denotes the 2-order logical type corresponding to the maximum uncertainty coefficient and Support of A, B to C under the effect of f [4,5,12]. Function g in the 3-order logical relationship is the 3-order logical type corresponding to the maximum uncertainty coefficient and Support of A, B, C to D under the effect of g [6,12]. Calculate value $U\left(B|A\right)$, $U\left(C\middle|f\left(A,B\right)\right)$ and $U\left(D\middle|g\left(A,B,C\right)\right)$ by formula (3), (4) and (5), respectively, based on the 0-1 discretized data expression profile of the 16 processed shoot active genes of Arabidopsis thaliana in section 2.1. And thus, all 1-, 2- and 3- order logical relationships and types among the 16 genes can be got. Obvious, it is a complete connected and direct logic network if we denote the above relationships by a graph.

Let u_i and n_i be the threshold value and the number of i-order logical relationships respectively, $i=1,2,3$, $\langle k\rangle$ be the average degree of the whole logical network.

Then, the threshold u_1 of 1-order logical relationships is restricted by the average degree of biological network. The threshold value u_2 of 2-order logical relationships is determined by the following formula (6):

$$\begin{cases} U\left(C\middle|f\left(A,B\right)\right)>U\left(C|A\right)\&\,U\left(C|B\right) \\ U\left(C|A\right)\ll u_1 \\ U\left(C|B\right)\ll u_1 \\ u_2>u_1 \\ n_2<n_1 \end{cases} \quad (6)$$

The threshold value u_3 of 3-order logical relationships is determined by the following formula (7):

$$\begin{cases} U\left(D\middle|g\left(A,B,C\right)\right)>U\left(D\middle|f\left(A,B\right)\right)\&U\left(D\middle|f\left(A,C\right)\right)\&U\left(D\middle|f\left(B,C\right)\right) \\ U\left(D\middle|f\left(A,B\right)\right)\ll u_2\ \&U\left(D\middle|f\left(A,C\right)\right)\ll u_2\ \&U\left(D\middle|f\left(B,C\right)\right)\ll u_2 \\ U\left(D\middle|A\right)\ll u_1\ \&U\left(D\middle|B\right)\ll u_1\ \&U\left(D\middle|C\right)\ll u_1 \\ u_3>u_2 \\ n_3<n_2 \\ \left\langle k\right\rangle \cong 2.12 \end{cases} \tag{7}$$

Through numerical calculation of (6), (7), we obtain that the 1-, 2- and 3-order threshold value is 0.0346, 0.106 and 0.39101 respectively.

2.3 Construction of Logical Network

We'll further construct gene logic network based on the threshold value determined in the section 2.2. For any genes whose U value is larger than the threshold of 1-order (2-order, 3-order), we think that there exist obvious 1-order (2-order, 3-order) logical relationships between (among) them, otherwise, there doesn't exist obvious logical relationships. According to the above rules, we screen out 15 1-order logics: 1→4，4→1，1→5，5→1，14→1，5→14，14→5，5→2，2→5，10→14，14→10，14→12，12→14，8→12，12→9; 6 2-order logics: 3,12→6 (logical function: C=~A∧B), 7,12→6 (logical function: C=A∨B), 3,12→7 (logical function: C=A∧~B), 6,12→7 (logical function: C=~A∨B), 6,7→12 (logical function: C=~A∨B), 6,13→12 (logical function: C=~(A∧B)); 1 3-order logic: 1,9,13→6 (logical function: D=(A∧¬B∧¬C)∨(¬A∧B∧¬C)∨(A∧B∧C)). More details are shown in Table 2.

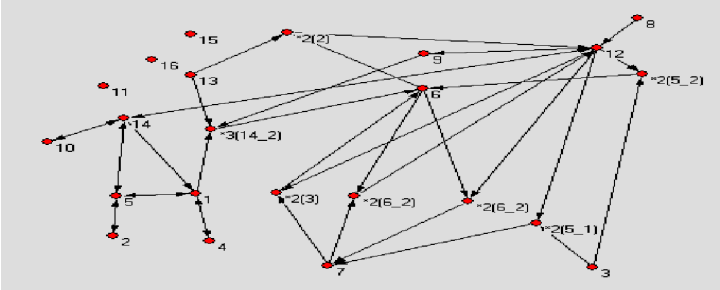

Figure 2: Logical Regulatory Network of 16 Shoot Active Genes of Arabidopsis

The above all logical relationships are integrated into a logic regulatory network of 16 shoot active genes of Arabidopsis thaliana, where three gene 11 (AT3G46690), 15 (AT5G01830), 16 (AT5G44990) are isolated vertices. The vertices in the network are divided into two types: the vertex of labeling "i" denotes gene "i" (i=1,2,…,16), the vertex of labeling "*j(…)" (j=2,3) denotes intermediate vertex of j-order logical relationships. For example, "*2(5_1)" denotes the type 5_1 of 2-order logic. The out-degree and in-degree of the intermediate vertices are fixed, i.e. when j=2, the in-degree and out-degree of the vertex is 2 and 1, respectively; when j=3, the

in-degree and out-degree of the vertex is 3 and 1, respectively. For instance, the vertex "*2(3)" in Fig.2 denotes that gene 7 and 12 regulate gene 6 through the type 3 of 2-order logic.

Table 2: List of All Logical Relationships and Types

	Arc	Logical Function	Corresponding Logical Type
1-order	1→4, 4→1, 1→5, 5→1, 14→1, 5→14, 14→5, 5→2, 2→5, 10→14, 14→10, 14→12, 12→14, 8→12, 12→9		
2-order	3,12→ 6	C=~A∧B	5_2
	7,12 →6	C=A∨B	3
	3,12 →7	C=A∧~B	5_1
	6,12 →7	C=~A∨B	6_2
	6,7→12	C=~A∨B	6_2
	6,13→12	C=~(A∧B)	2
3-order	1,9,13→6	D=(A∧¬B∧¬C)∨(¬A∧B∧¬C)∨(A∧B∧C)	14_2

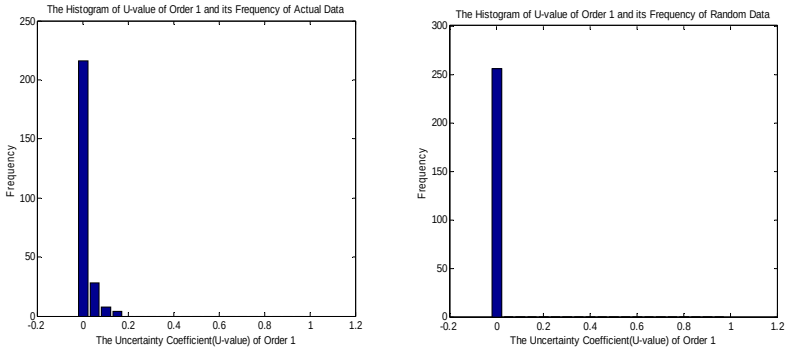

Figure 3: Histogram of 1-order U Value and Frequency in Actual and Random Data

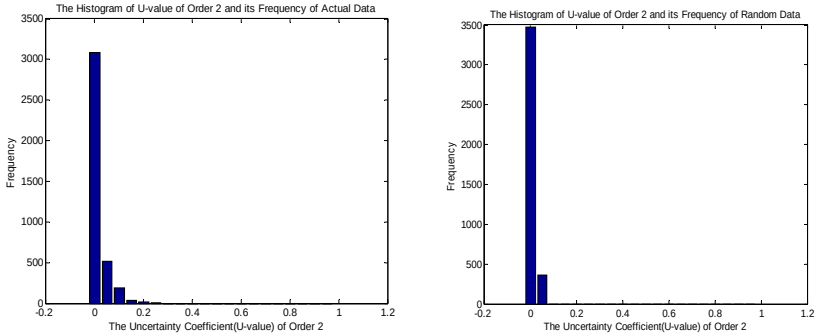

Figure 4: Histogram of 2-order U Value and Frequency in Actual and Random Data

2.4 Significance Test of Logical Relationships

Through the method of LAPP, logical relationships can be obtained by calculating U value. Are these logical relationships reliable, and are these data specific? Therefore, we need to do significance test. In this paper, we analyze the parameter significance test by comparing the actual data and random data. The method is as follows: 1. Generate independent identically distributed random variables with invariable 0,1 distribution of each gene expression profile data; 2. Calculate the U value in the every order logic combination of random variables; 3. Repeat the above steps 100 times, and calculate the average U value of each logic combination; 4. Divide U value and count the number of logical combinations of each segment separately; 5. The statistical significance test is done by calculating the numbers of all order logic gene combinations of random data and actual data. Finally, Compare the distribution of U value of random data with that of actual data, see Fig.3, Fig.4 and Fig.5. And we can see the significant difference of the number of logic combinations between the actual and random data from Fig. 6.

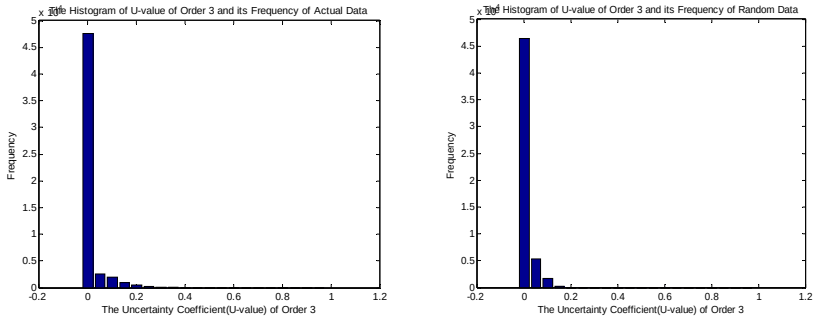

Figure 5: Histogram of 3-order U Value and Frequency in Actual and Random Data

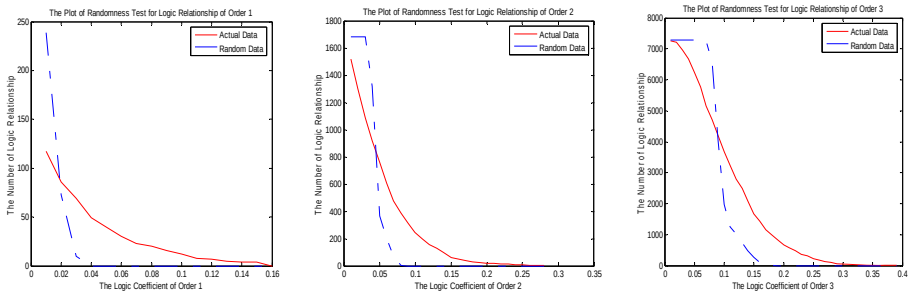

Figure 6: The Curve of Random Test for 1-, 2- and 3-order Logical Relationships

3 Dynamical Analysis of Logical Network

For system biology, domestic and overseas research interests mostly focus on the topological structure of gene regulatory networks. That is because it is well-known that the topological structure can reflect important biological meanings. Although the study of topological structure can obtain some biological principles and information,

the research method has one-sidedness. Because biological functions directly affect the dynamical behaviors, whereas, the dynamical behaviors can't be reflected by topological structure in most situations. Therefore, we'll study the evolution law of gene logic regulatory network from dynamics point of view.

3.1 Dynamical transition rules

Let $s(t)=(x_1(t),x_2(t),\cdots,x_{16}(t))$ denote the state-vector of logic regulatory network at moment t, where $x_i(t)\in\{0,1\}$ is the state of gene i at any time t, $i=1,2,\cdots,16$. Let the triple set comprised of the gene affecting gene i, U value and 1-order logic at any time t be $Z_1=\{(k,U_{k\to i},L_k^1)|k\in\{1,2,\cdots,16\},k\neq i\}$, which denotes that gene k acts on gene i through the 1-order logic L_k^1, and the uncertainty coefficient is $U_{k\to i}$; Let the triple set comprised of the genes affecting gene i, U value and 2-order logic at any time t be $Z_2=\{((k_1,k_2),U_{k_1,k_2\to i},L_{k_1,k2}^2)|k_1,k_2=1,2,\cdots,16,k_1\neq i,k_2\neq i,k_1\neq k_2\}$, which denotes that genes k_1,k_2 act on gene i through the 2-order logic L_{k_1,k_2}^2, and the uncertainty coefficient is $U_{k_1,k_2\to i}$; Let the triple set comprised of the genes affecting gene i, U value and 3-order logic at any time t be

$$Z_3=\left\{((k_1,k_2,k_3),U_{k_1,k_2,k_3\to i},L_{k_1,k_2,k_3}^3)\left|\begin{array}{l}k_1,k_2,k_3=1,2,\cdots,16,k_1\neq i,k_2\neq i,k_3\neq i,\\ k_1,k_2,k_3 \text{ are different from each other}\end{array}\right.\right\},$$

which denotes that genes k_1,k_2,k_3 act on gene i through the 3-order logic L_{k_1,k_2,k_3}^3, and the uncertainty coefficient is $U_{k_1,k_2,k_3\to i}$. For each gene i at time t, take

$$P_i(t)=\left(\sum_{Z_1}U_{k\to i}\bullet L_k^1+\sum_{Z_2}U_{k_1,k_2\to i}\bullet L_{k_1,k_2}^2+\sum_{Z_3}U_{k_1,k_2,k_3\to i}\bullet L_{k_1,k_2,k_3}^3\right)\Big/\left(\sum_{Z_1}U_{k\to i}+\sum_{Z_2}U_{k_1,k_2\to i}+\sum_{Z_3}U_{k_1,k_2,k_3\to i}\right)$$

Then, the transition function is denoted by $x_i(t+1)=\begin{cases}1, & P_i(t)\geq 0.5,\\ 0, & P_i(t)<0.5.\end{cases}$, $i=1,2,\cdots,16$.

3.2 Analysis of Attractors

The whole state-space of the system includes 2^{16}(=65536) states. For the convenience of writing and computing, we put all the 65536 state-vectors into serial number, each binary state-vector is responding to the corresponding decimal numbers. So each state-vector in our state configuration network appears as its serial number.

According to the state transition rules in the section 3.1, there are 11 domains of

attraction altogether in the above transition configuration network, which includes five 1-periodic domains of attraction and six 2-periodic domains of attraction. Fig.3 and Fig.4 are the attraction domain of attractor $852 \to 852$ and $56384 \to 56324 \to 56384$ respectively. And the detailed information about the above domains of attraction is as Table 3 and Table 4. In the 17 state-vectors responding to the 11 attractors, the number of state 1 from 1^{st} to 16^{th} subvectors is 11, 11, 0, 11, 11, 11, 11, 11, 0, 11, 0, 11, 0, 11, 0, 0, respectively. Therein, the occurrence numbers of 1^{st}, 2^{nd}, 4^{th}, 5^{th}, 6^{th}, 7^{th}, 8^{th}, 10^{th}, 12^{th} and 14^{th} subvector are all 11, accounting for 64.71% (=11/17); and the occurrence numbers of the 3^{rd}, 9^{th}, 11^{th}, 13^{th}, 15^{th} and 16^{th} subvector are all 0. This implies that these ten genes 1, 2, 4, 5, 6, 7, 8, 10, 12, 14 will tend to get to stable state 1 in most cases. Whereas, the occurrence numbers of gene 3, 9, 11, 13, 15, 16 are all 0, which suggests that these genes will tend to get to stable state 0 in the dynamical evolution process of the gene network (Fig.2). Judging from Table 3 and Table 4, the number of states of the attraction domain of the attractors $57172 \to 57172$, $49940 \to 8016 \to 49940$ and $50004 \to 8020 \to 50004$ are much more than that of other attractors, They are 13312, accounting for 20.3% (=13312/65536), 22016, accounting for 33.6% (=22016/65536), and 10752, accounting for 16.4% (=10752/65536), respectively. The above analysis suggests that most states of the system can transfer to the three attractors, and the transient time of the three attraction domains is between 3 and 4. Therefore, Arabidopsis thaliana shoot system can revert into stable states in shorter time under outside different stimuli. Furthermore, we have noticed that the patterns of both attractor $49940(1100001100010100) \to 8016(0001111101010000) \to 49940(1100001100010100)$ and attractor $50004(1100001101010100) \to 8020(0001111101010100) \to 50004(1100001101010100)$ are very similar, both from 110000110*010100 to 0001111101010*00. This shows that gene 10 (AT3G11340) and gene 14 (AT4G37370) play an important role in self-regulation of Arabidopsis shoot system under outside stimuli.

Table 3: Details about Six 2-Periodic Attractors

Attractor (Serial Number)	Number of States in the States Transition Configuration Network of Attractor
0000001100010100（788） 0000001101010000（848）	7424
1100001100010100（49940） 0001111101010000（8016）	22016
1100001101010100（50004） 0001111101010100（8020）	10752
1101110001000000（56384） 1101110000000100（56324）	512
1101111100000100（57092） 1101110001010000（56400）	1152
1101111101000100（57156） 1101110001010100（56404）	896

Table 4: Details about Five 1-Periodic Attractors

Attractor (Serial Number)	Number of States in the States Transition Configuration Network of Attractor
0000001100010000（784）	7296
0000001101010100（852）	1664
1101110000000000（56320）	320
1101110001000100（56388）	192
1101111101010100（57172）	13312

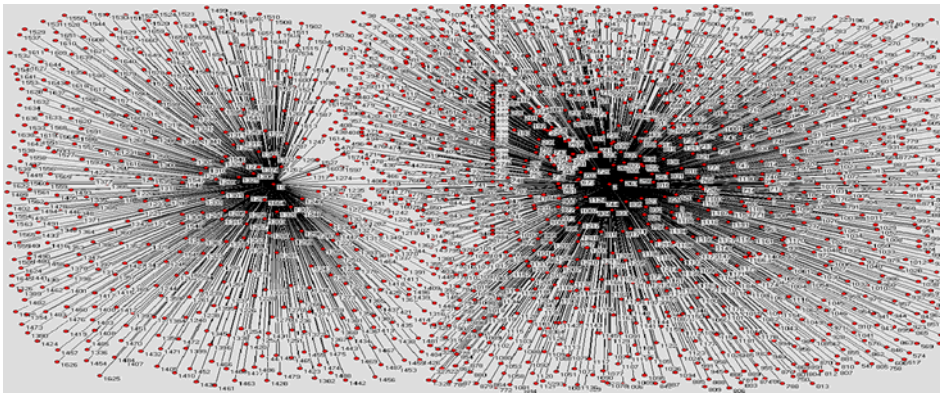

Figure 3: State Transition Configuration Network of 1-Periodic Attractor $852 \rightarrow 852$

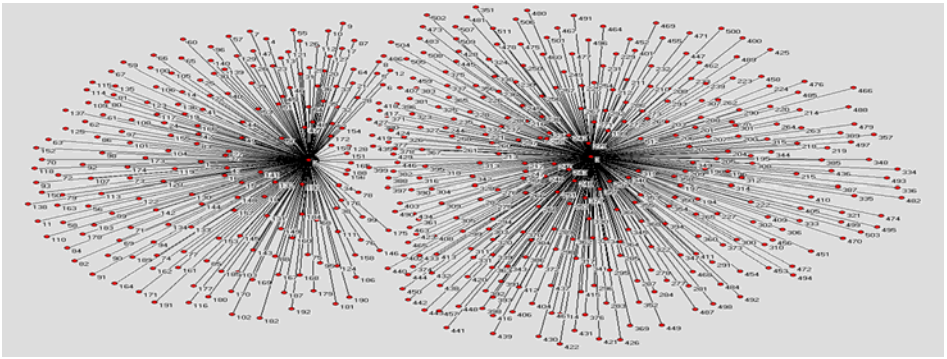

Figure 4: State Transition Configuration Network of Attractor $56384 \rightarrow 56324 \rightarrow 56384$

4 Discussions

This paper has utilized LAPP to search 1-, 2- and 3-order logical relationships and logic types among the 16 shoot active genes of Arabidopsis under different stimuli. Moreover, we have simulated and analyzed the dynamical evolution process of these logical relationships. However, Using LAPP to determine the logical relationships of genes and proteins, there may appear “false-positive” and “false-negative”, that is to

say there are different errors during determining logical relationships. But we don't analyze the error creared by LAPP in this paper.

In the other hand, we have represented the larger ratio (1.9, 1000) and the smaller (0.000001, 0.51) with 1 when discretizing the actual data to 0-1 value. The way of discretization is mainly used in the Boolean network model we have studied. In fact, more reasonable way of discretization is to use 1 and -1 to denote the larger ratio and the smaller ones respectively. This drives us extend the two-value {0, 1} logic to three-value {-1, 0, 1} logic. At present, the proper functions and logical types of the three-value logic have been still studying. The three-value logic network model can depict, describe and simulate biological models more elaborately. In addition, through analyzing the root ratio data (see Appendix-2) of the 16 same genes in the shoot, we find from the data: the genes that are active in shoot mostly are inactive in root and some genes AT1G51860, AT2G35910 and AT3G11340 nearly don't express. Indeed, this fact supports the function of the same genes (genome) in different Arabidopsis parts (shoot and root part) is distinct. In another paper, we will further compare the respective active genes in root and shoot, and find the functional gene clusters in different parts under outside stimuli. In a word, our work provides reference and guidance value for biologists to understand and explain Arabidopsis thaliana reaction to outside stimuli by experiments.

References

[1] Marcotte EM, Pellegrini M, Ng HL, Rice DW, Yeates TO & Eisenberg D. Detecting protein function and protein–protein interactions from genome sequences. Science, 1999, 285: 751–753.

[2] Strong M, Graeber TG, Beeby M, Pellegrini M, Thompson MJ, Yeates TO & Eisenberg D. Visualization and interpretation of protein networks in Mycobacterium tuberculosis based on hierarchical clustering of genome-wide functional linkage maps. Nucleic Acids Res., 2003, 31: 7099–7109.

[3] Zhou X, Kao MC & Wong WH. Transitive functional annotation by shortest-path analysis of gene expression data. Proc. Natl. Acad. Sci., 2002, 102(38):12783–12788.

[4] Bowers P M, Cokus S J, Eisenberg D, et a1. Use of logic relationships to decipher protein network organization. Science, 2004, 306(5706):2246–2249.

[5] Bowers P M, O'Connor B D, Cokus S J, et al. Utilizing logical relationships in genomic data to decipher cellular processes. The FEBS Journal, 2005, 272 (20): 5110–5118.

[6] Zhang Xin, Kim Seungchan, Wang Tie and Baral Chitta. Joint learning of logic relationships for studying protein function using phylogenetic profiles and the Rosetta Stone method. IEEE Transactions on Signal Processing, 2006, 54 (6):2427–2435.

[7] Zhang J X, Ji Y, Zhang L. Extracting three-way gene interactions from microarray data. Bioi nformatics, 2007, 23(21):2903–2909.

[8] Kauffman S.A.. The Origins of Order: Self-organization and Selection in Evolution, Oxford University Press, Oxford, 1993.

[9] Shmulevich I., Kauffman S.A., Aldana M.. Eukaryotic cells are dynamically ordered or critical but not chaotic. Proc. Natl. Acad. Sci., 2005, 102(38): 13439-13444.

[10] Li F., Long T., Lu Y., Ouyang Q., Tang C.. The yeast cell-cycle network is robustly designed. Proc. Natl. Acad. Sci., 2004, 101(14):4781–4786.

[11] Luscombe N.M., Badu M.M., Yu H., Snyder M., Teichmann S.A., Gerstein M.N. Genomic analysis of regulatory network dynamics reveals large topological changes.

Nature, 2004, 431: 308–312.
[12] Wang QingYun, Yi Hui, Liu Laifu, Meng Dazhi. Progress in gene logic networks. Progress In Biochemistry and Biophysics, 2008, 35(11): 1239–1246.

Appendix-1 (Shoot Ratio Data):

http://cise.sdkd.net.cn/institute/ISBBC/data.htm/Appendix-1.xls

Appendix-2 (Root Ratio Data):

http://cise.sdkd.net.cn/institute/ ISBBC/data.htm/Appendix-2.xls