数理解析研究所講究録
第827巻 1993年 57-74

Development of Developmental System Analysis System

Laboratory of Theoretical Biophysics, Department of Biophysics,
Faculty of Science, Kyoto University

Takeda Yasuhiko 武田裕彦

§1 Introduction

In a sense, the problems of biology can be considered as elucidation of the contents of black box. A biological system determines its behavior responding to the information from outside or inside. The phenomena of development or differentiation proceed by iteration of that kind of processes. Our purpose is to infer what kind of system is at work in each step from experimental data. (Fig. 1)

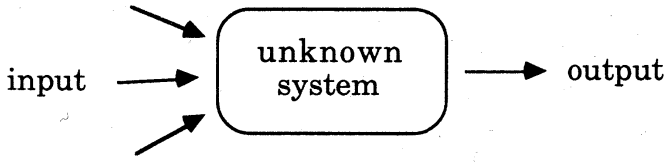


Fig. 1

Thinking about the problem of development, there are several levels of systems that receive input and return output, tissue, cell, gene, and so on. (Fig. 2) How do the signals interact? What kind of influence do they give output? We want to solve these problems.

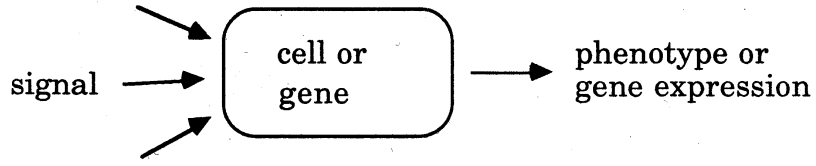


Fig. 2

The ultimate purpose of this approach is to simulate total system by combining subsystems of known function. To realize it, it is necessary to simulate each subsystem correctly. I developed a system that carry out such procedure.

§ 2 Development of Developmental system analysis system

§ 2 - 1 A modeling framework for development

I take the view point that regards developmental phenomena as "From positional information to pattern formation". First, "positional information", namely, fate of development is determined by expression pattern of genes. Afterward, "pattern formation" proceeds by proliferation, movement, and death of each cell obeying "positional information". Sometimes two phenomena proceed together like limb bud formation. However, in practical, many processes are separated into two phases, such as "determination of positional information" and "progress of pattern formation". Here I concentrate on the former phase, and consider the problem to determine the interaction among informational genes from their expression pattern.

§ 2 - 2 The level of analysis

There are several ways to molecular implementation of interaction, but elucidation of it is the job for molecular biologist. Here, I try to contribute to elucidation of the system by inferring qualitative aspect of interaction, such as specification of morphogen, or negative or positive correlation between the genes.

As the system to discuss, I take

1) genes that take spatio-temporal expression pattern,
2) genes that make effect on an interested phenomenon.

The I/O for them are, respectively,

1) the expression pattern of the genes,
2) the mutant and its phenotype of the genes.

§ 2 - 3 Instrument for analysis

§ 2 - 3 - 1 Back propagation learning by feed forward network

I use two layered feedforward network of neural units to describe the system that receive input and return output. (Fig. 3)

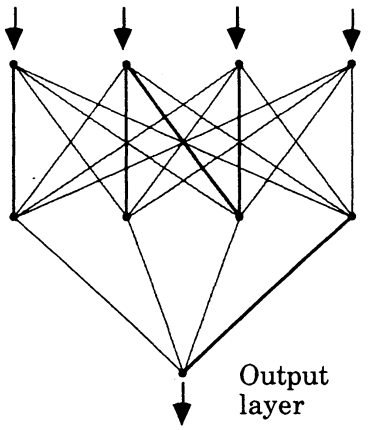


Fig. 3

Each unit determines its state by the sum of all the strength of all connections to it. Its output is a sigmoid function of its state.

$$f_i^k \equiv f(v_i^k) = \frac{1}{1 + \exp(-v_i^k)}$$

$$v_i^k = \sum_j w_{ij}^k f_j^{k-1} - \theta_i^k$$

v_j^k : state of the j-th unit in the k-th layer

f_j^k : output of the j-th unit in the k-th layer

w_{ij}^k : strength of connection from the j-th unit in the (k-1)-th layer to the j-th unit in the k-th layer

θ_i^k : threshold of the j-th unit in the k-th layer

It's proved that any continuous mapping can be approximately realized by two layered feedforward network whose output functions are sigmoid functions[1]. Therefore the network is enough to have ability to describe complex interaction of biological system.

Pairs of input pattern and teaching signal are learned by the network using back propagation algorithm[2]. Here, I don't use the usual algorithm that minimize the mean square error between actual output and teaching signal. Using it, the learned network shows distributed expression so that it is difficult to recognize the function of a hidden unit, or the effect of an input unit for the function of the network. Accordingly, I use an algorithm that minimize the objective function R, which is the summation of mean square error, and a term whose effect is to make the network smaller, and a term that clarifies each hidden unit's state fully activated or rested for each input pattern[3]. As a result, the learned network is as near as the most compact network that has desirable I/O relation.

o_i : output for the i-th input pattern

t_i : teaching signal for the i-th input pattern

h_{il}^k : activation of the i-th unit of the k-th layer for the l-th input pattern

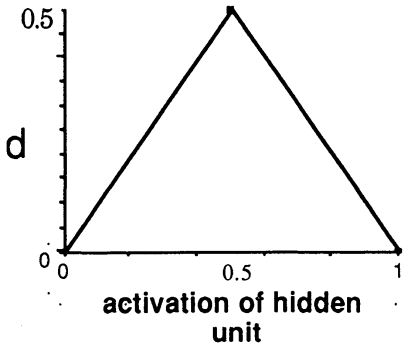
$$R = \sum_i (o_i - t_i)^2 + \varepsilon \sum_{k\,i\,j} |w_{ij}^k| + \varepsilon \sum_{k\,j} |\theta_j^k| + \gamma \sum_{i\,k\,l} d(h_{il}^k)$$

Correction for the t-th step

$$W_{ij}^{k}(t) = W_{ij}^{k}(t-1) - \alpha \frac{\partial R(t-1)}{\partial W_{ij}^{k}(t-1)} - m \frac{\partial R(t-2)}{\partial W_{ij}^{k}(t-2)}$$

$$\theta_{i}^{k}(t) = \theta_{i}^{k}(t-1) - \beta \frac{\partial R(t-1)}{\partial \theta_{i}^{k}(t-1)} - m \frac{\partial R(t-2)}{\partial \theta_{i}^{k}(t-2)}$$

output of function d



§ 2 - 4 Objects for application

It is no need to say that known interaction should be reproduced when this instrument is applied to the system whose units' interaction has already understood. If it happens, we can apply this to unknown system. Here, I take the system of D-V and A-P axis formation in early Drosophila embryo as an example of application of this instrument to a specific developmental system.

§ 2 - 4 - 1 Antero-Posterior(A-P) axis

In early embryogenesis of Drosophila Meranogaster, after egg deposition(AED), almost synchronous nuclear division proceeds without cytoplasmic division. At syncytial blastoderm stage(1:20～2:10 minutes AED 25 C°), developmental fate of each segment is determined from expression pattern of genes on the surface of the embryo. (Fig. 4)[4]

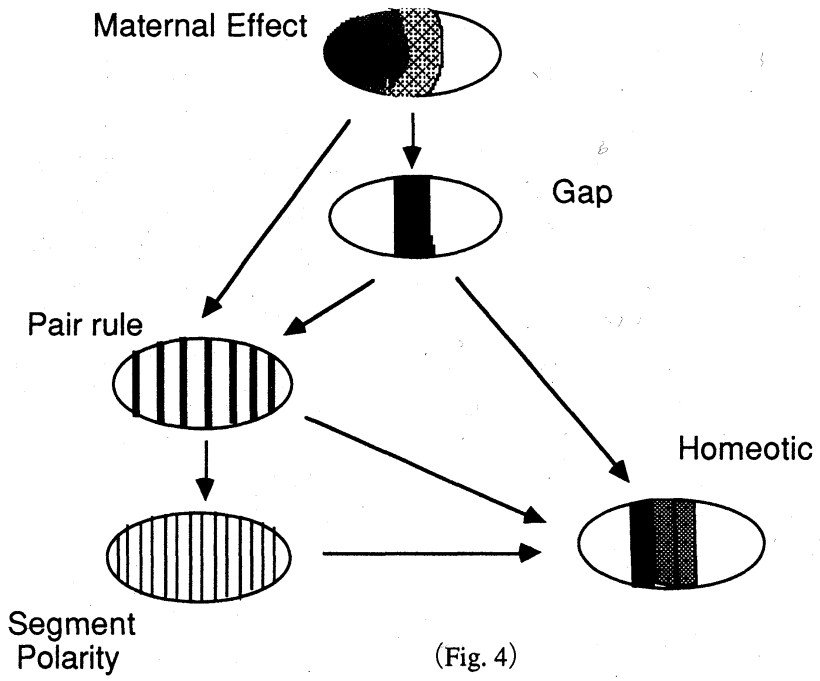
Concretely speaking, the information along A-P axis, which is specified by gradients of maternal effect genes, is transformed into periodic segmental structure by segment polarity genes through gap genes and pair rule genes. In parallel, homeotic genes determine the identities of each segment. These groups of genes shows hierarchical expression pattern along the time course of development. Each gene expresses with characteristic spatio-temporal pattern responding to the spatially inhomogenious input.

§ 2 - 4 - 2 Dorso-Ventral(D-V) axis

The D-V axis of Drosophila embryo is determined by the nuclear concentration gradient of morphogen, which is the gene product of gene named dorsal(dl). The system that makes this gradient consists of six groups of genes. Naturally, dl gene product is uniformly distributed in the cytoplasm along D-V axis associated with cactus gene product. There is a signal from somatic cells at ventral side of the embryo. The signal makes dl gene product transferred into nuclear and produces the gradient of concentration in the nuclear[5]. (Fig. 5)

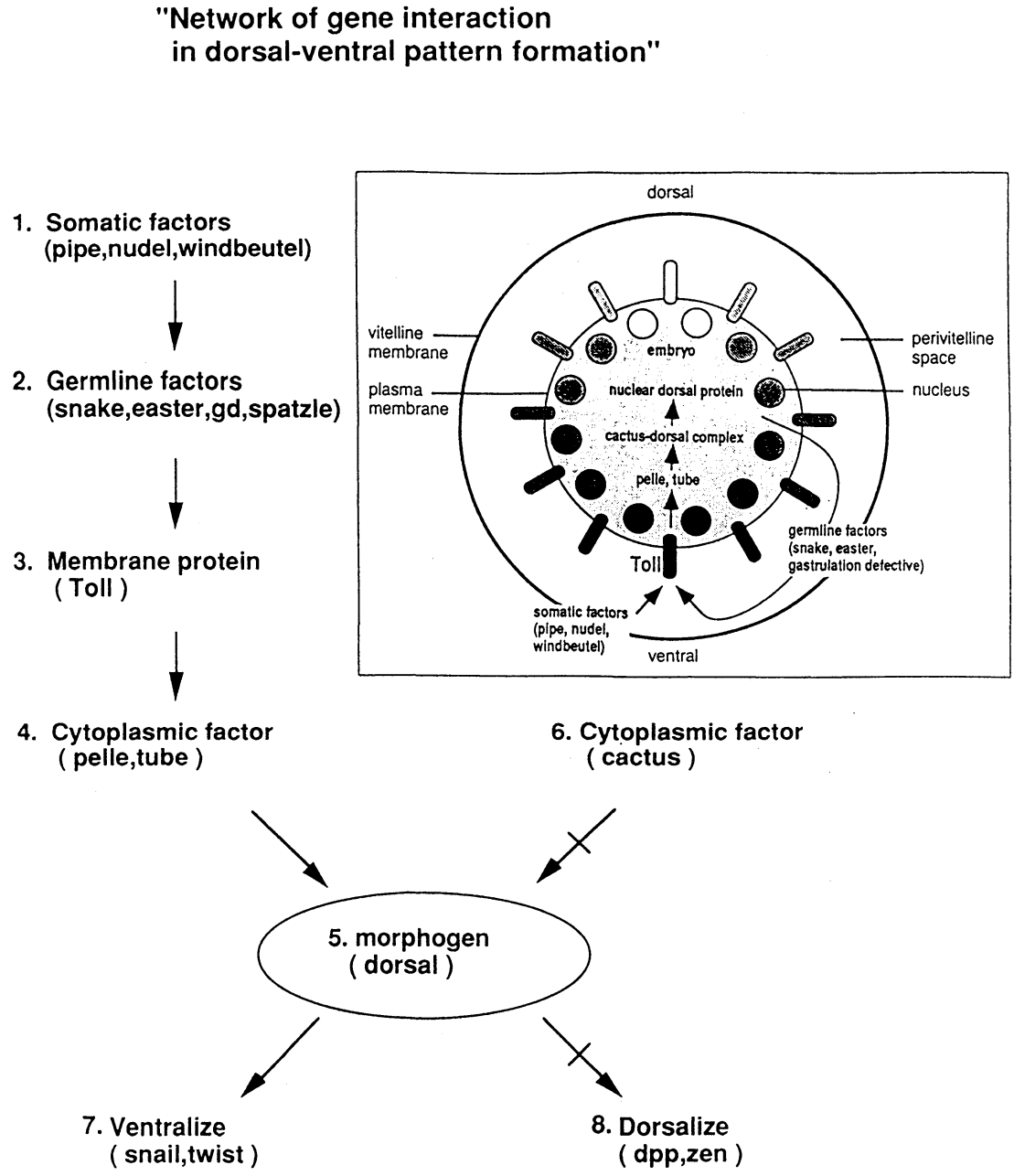
The functions of six identified groups of the genes are inferred from phenotypes of each gene mutant. Each mutation changes the distribution of

dorsal gene product along D-V axis so that the mutant embryo shows ventralized or dorsalized phenotype.



(Fig. 4)

Fig. 4 : The diagram illustrates the expression pattern of the groups of the genes that act to make segmented body plan of Drosophila in early stages of its development. Arrows indicate the existence of regulatory interactions between members of these classes of genes.



(Fig. 5)

Fig. 5 : The interaction among the groups of the genes related to D-V axis formation. Somatic factors(1), and Germline factors(2) that are secreted from egg cell form the ventrally localized signal. Membrane protein, Toll(3) receives the signal and transduces it to cytoplasmic factor(4). On the other hand, in the cytoplasm, there exists dorsal(5), morphogen for ventralizing, associated with cactus(6). If there's the signal (1)→(4), (5) is transferred into nuclear and activates the ventralizing genes(7) and inhibits the dorsalizing genes(8). In the figure, →indicates positive, ⊣indicates negative effect, respectively.

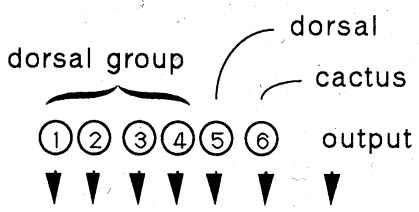
§ 3 Examples-

§ 3 - 1 Dorso-Ventral axis

Each function of the groups of the genes related to D-V axis formation is almost understood. (Fig. 5) Here, I infer the interaction among the groups of the genes from the mutant phenotypes. The input and output for learning is as follows.

Input :
(for each groups of the genes)
mutant ⇒ 0
wild type ⇒ 1

Output : ventralized ⇒ 1
dorsalized ⇒ 0
lateralized ⇒ 0.5



{{1,1,1,1,1,1},{1}}
{{0,1,1,1,1,1},{0}}
{{1,0,1,1,1,1},{0}}
{{1,1,0,1,1,1},{0}}
{{1,1,1,0,1,1},{0}}
{{1,1,1,1,0,1},{0}}
{{1,1,1,1,1,0},{1}}
{{0,0,1,1,1,1},{0}}
{{0,1,0,1,1,1},{0}}
{{0,1,1,0,1,1},{0}}
{{1,0,0,1,1,1},{0}}
{{1,0,1,0,1,1},{0}}
{{1,1,0,0,1,1},{0}}
{{0,1,1,1,1,0},{0.5}}
{{1,0,1,1,1,0},{0.5}}
{{1,1,0,1,1,0},{0.5}}
{{1,1,1,0,1,0},{0.5}}
{{1,1,1,1,0,0},{0}}

(Fig. 6)

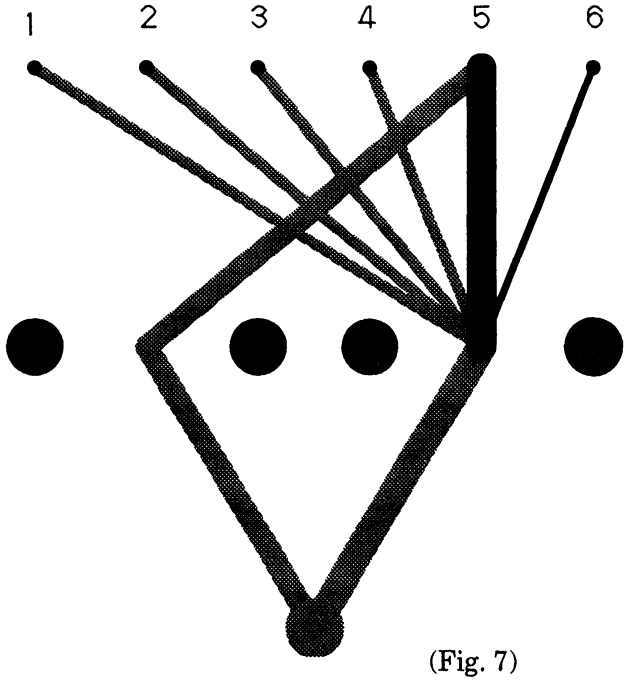
If some of the six groups of the genes lose its function, whole embryo is ventralized or dorsalized. In other words, a change of output follows a change of input. There exists 2^6 (= 64) I/O relation in all. But usually, because of experimental difficulty, we can get double mutants at best. Therefore, the network could learn only 18 input patterns. (Fig. 6) Each learning pattern is a vector whose elements have value 1 or 0, and number of 0 element is not more than two. The vector whose all elements have value 1 represents the wild type ventral side input.

In figure 7, a typical learned network is shown. Even if the learning starts from a different initial network state, same topology of network is realized as long as the learning converges. The thickness of connection indicates the absolute value of strength of connection. The radius of each unit indicates the absolute value of threshold. Gray indicates negative. Black indicates positive.

What can we get from the learned network? At first sight, we can see that dorsal that corresponds to the fifth input unit plays a special role. Also, the 1-st to 4-th input units(= dorsal group) have the same effect to the output. It is understood that they are the factors of promoting ventralization, because they inhibit the hidden unit that inhibits the output. Conversely, cactus, the sixth

input unit, is the inhibitor of ventralization, for it activates the hidden unit that inhibits the output.

The problem is whether dorsal, the 5-th input unit, is the morphogen of this system or not. Here, the term 'morphogen' means the factor whose concentration critically determines the output of the system. To certify this, the I/O of this network is plotted. (Fig. 8)



(Fig. 7)

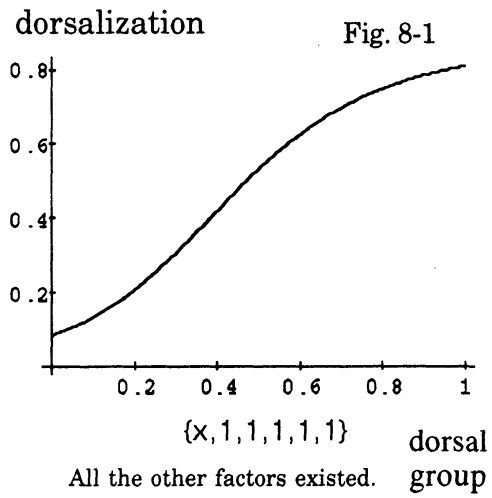


Fig. 8-1

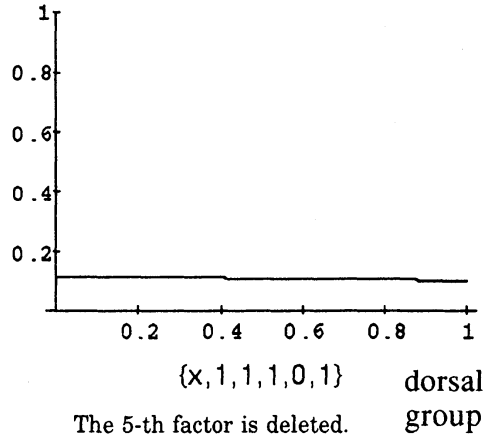


Fig. 8-2

Figure 8 plots the outputs of the network changing the state of one input unit from 0 to 1 with the other states of input units fixed. Figure 8-1, 8-3, 8-5 is the output when all the other factors exist. It is easily seen that dorsal group and dorsal activate ventralization, and cactus inhibits it. To show which the

morphogen of this system is, figure 8-2, 8-4, 8-6 plots the outputs of the network when one of the other factors is fixed to 0, namely, deleted. In dorsal⁻, both dorsal group and cactus make no influence on the output of the network. On the contrary, even if the other inputs are deleted, the network responds to changes of the dorsal input as same as they exist. Therefore, it is concluded that the morphogen of this system is dorsal. It seems significant that useful information can be got from experimental data such as figure 6.

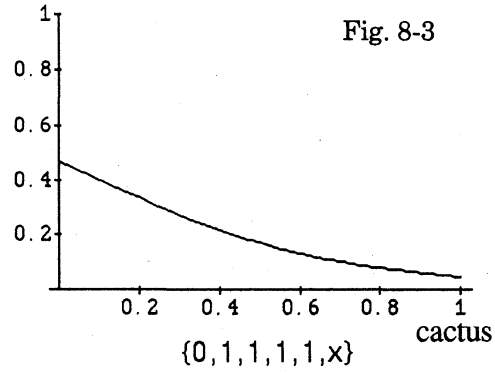


Fig. 8-3

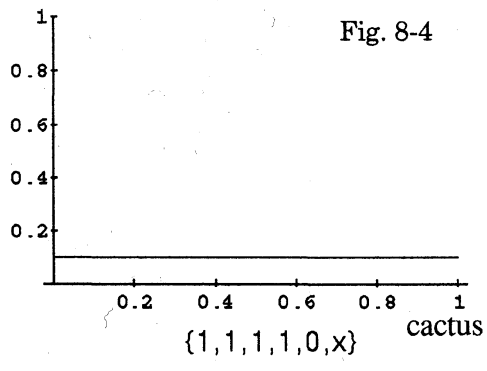


Fig. 8-4

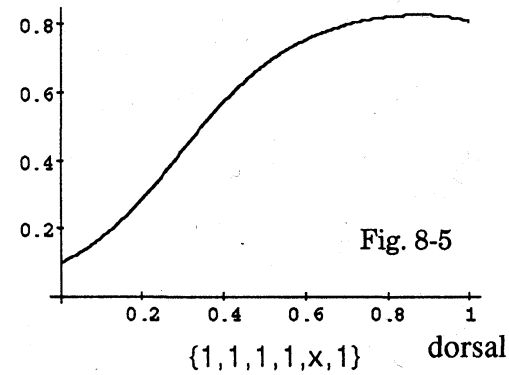


Fig. 8-5

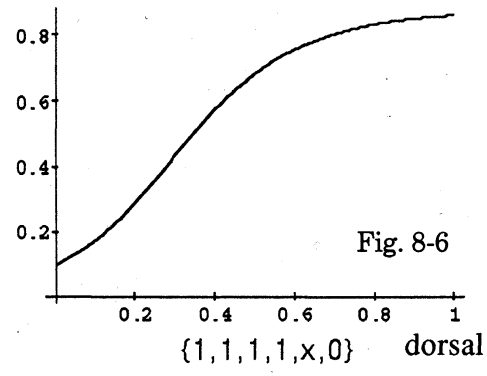
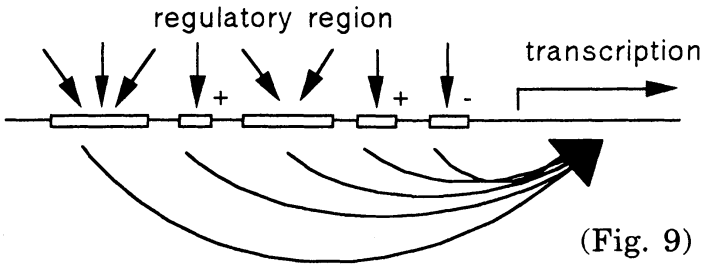


Fig. 8-6

§ 3 - 1 Antero-Posterior axis

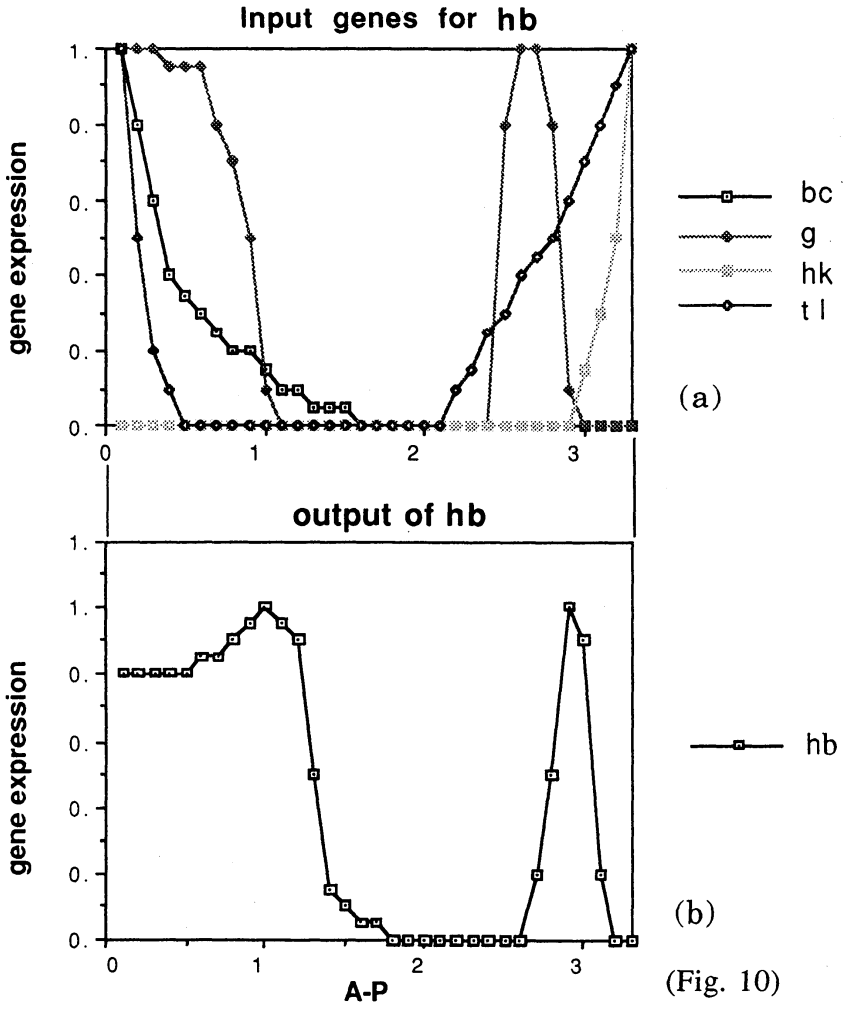
In A-P axis formation, several groups of genes interacts hierarchically along the time course of development. Each gene shows characteristic spatio-temporal expression responding to spatially inhomogenious input. Most of the genes have DNA binding site and regulate the gene expression of the same or lower layers of hierarchy. In the regulatory region of a gene, there are many binding site for the genes in the upper layer. Particularly, it is found that there exists localized domain of regulatory region that regulates the gene expression of specific region of the embryo.[6)~8)] Then, a gene can be assumed to be a two layered Feedforward network. The factors that have influence on the gene expression are assigned to the input layer. Each domain in the regulatory

region that responds to various inputs takes as a hidden unit. Synthesizing process in the regulatory region that determines the gene expression is approximated in the output layer. (Fig. 9)



(Fig. 9)

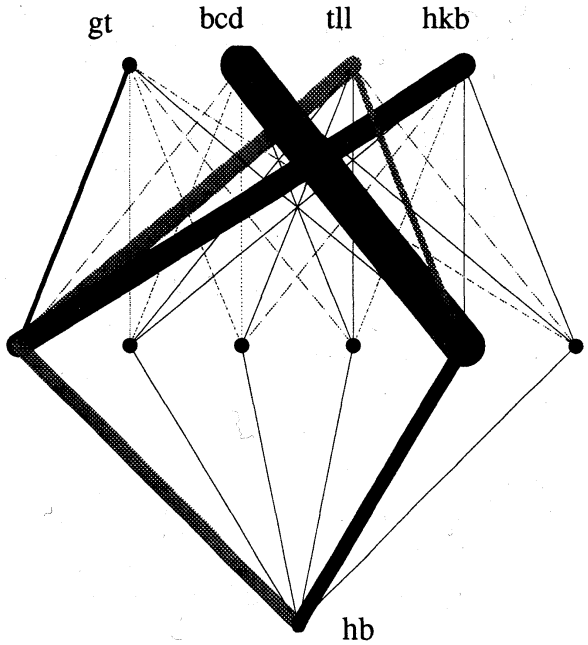
According to the above description, let's infer the response of each gene for inputs from the gene expression pattern on the surface of the embryo.



(Fig. 10)

§ 3 - 2 - 1 hunchback hb

Hunchback(hb) that belongs to the groups of gap genes expresses in the two places of surface of the embryo. (Fig.10) Both (a) and (b) plot the normalized gene expression of wild type embryo along the A-P axis. Left side is anterior, and right side is posterior. (a) is the expression patterns of the genes that seem to affect the hb expression. (b) is the expression of hb that responds to (a). To learn the I/O relation, the A-P axis is discretized into 33 intervals. Normalized expression of genes in each interval is used as learning pattern and teaching signal. The result of learning is following network. (Fig. 11)



(Fig. 11)

The structure of learned network shows that

1) bicoid(bcd) activates hb expression
2) tailless activates hb expression, and hackbein(hkb) represses this effect.
3) giant(gt) contributes almost nothing to hb expression

1) corresponds to the expression of anterior part of the embryo[9], and 2) corresponds to that of posterior part of it[10]. These things are proved to be correct by experiment. There's no experimental data for 3) by now, but it is natural that gt that belongs to the group of gap genes has weak effect on hb expression.

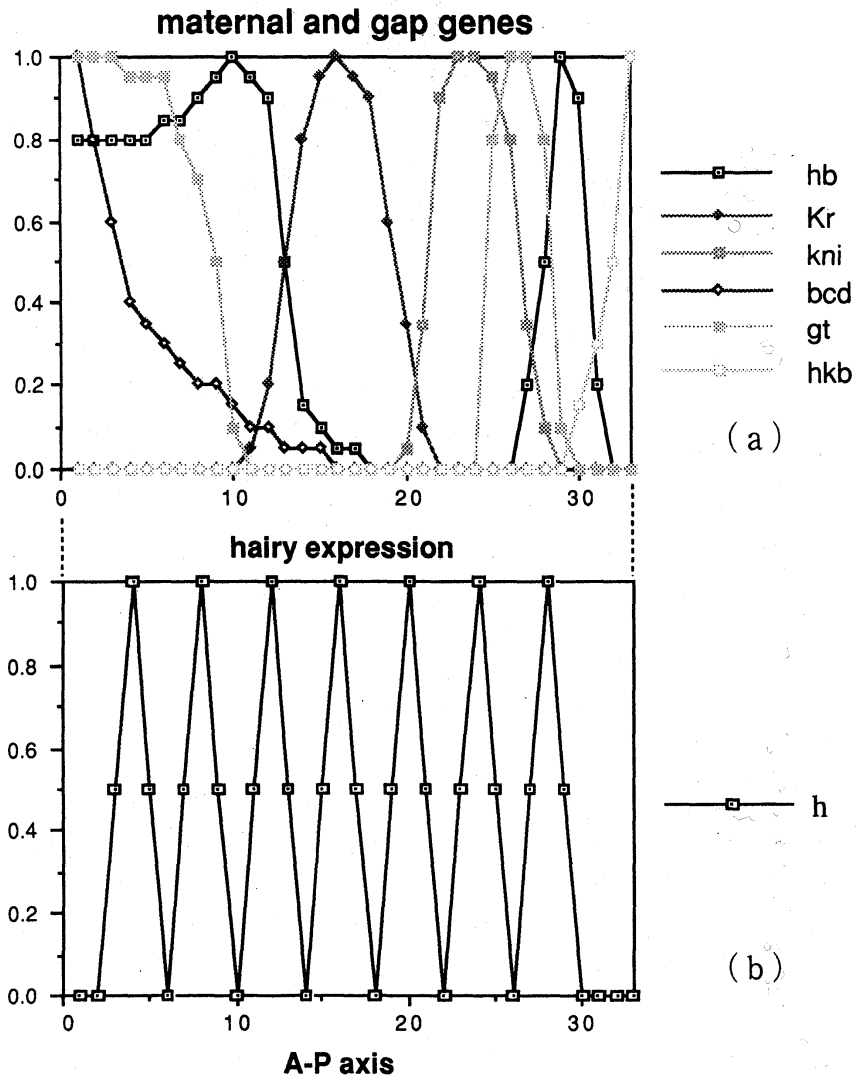
§ 3 - 2 - 2 Pair-rule gene

As explained above, hb expression is determined by threshold function of concentration of maternal effect genes, bcd, tll, and hkb to a certain degree.

Therefore, it's close to Boolean function. Contrast to this, the group of pair-rule genes that express seven striped pattern receive spatially complicated input of maternal effect and gap genes. (Fig. 12) How the primary pair-rule genes, such as hairy (h), even skipped (eve), decode spatial distribution of gap genes?

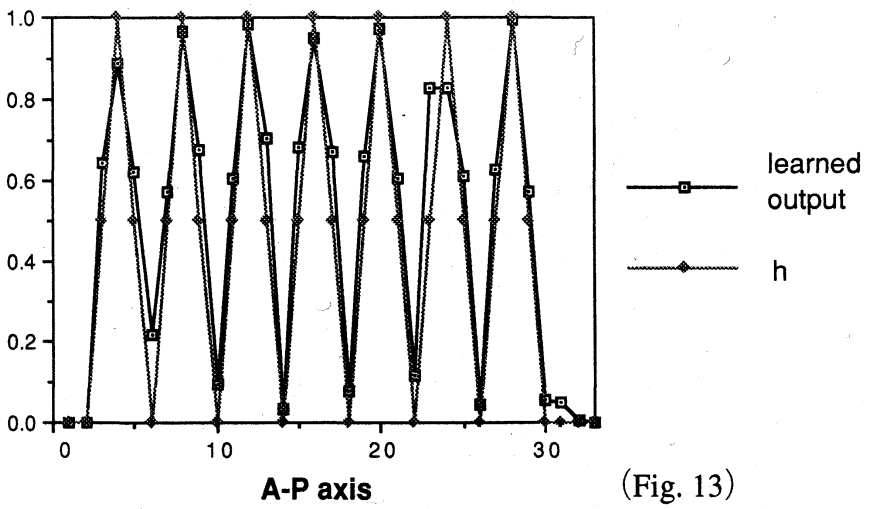
There's a model that explains the seven striped pattern by mutual inhibition and long range activation of pair rule genes[11]. It assumes that the pair rule genes have interaction to make a periodic pattern with characteristic wave length, and the gap genes set its boundary condition. But it is not sure that the real system uses that kind of reaction diffusion system.

Here, I ignore the interaction between the pair rule genes for the present and assume they express striped pattern depending on specific concentration relation of gap genes. Let's infer the function form of them from I/O relation experimental data.



(Fig. 12)

Figure 12 plots the expression of the genes that regulate hairy expression (a) and hairy itself (b) by the same way of figure 10[12]. The I/O relation of hairy is learned using (a) as input and (b) as output. Figure 13 overwrites the output of learned network along the A-P axis. It is found that the information of six genes on the embryo is enough to express seven striped pattern of h. The problem is what to the extent the learned network reproduces the function of hairy gene.

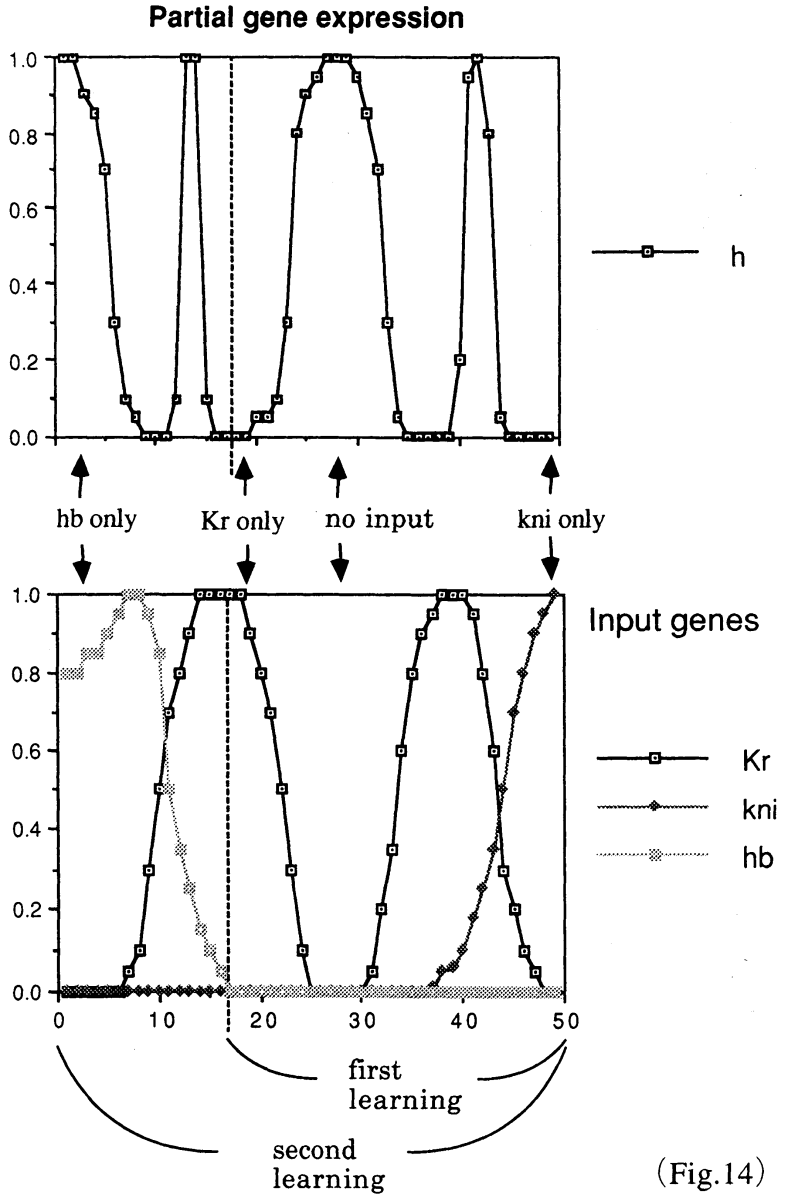


(Fig. 13)

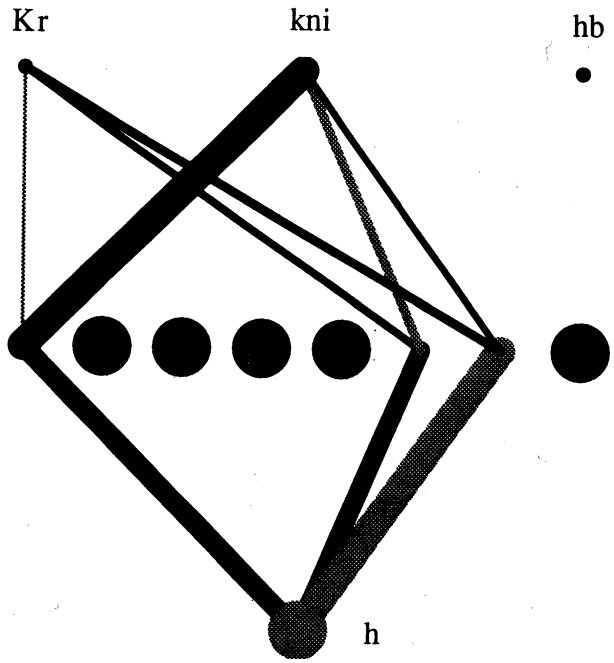
The learned network should be able to reproduce the mutant phenotype. But from experiment, the effect of a gene is often different from it working alone, when there exists other gene products. In case of h, triple mutant, hb$^-$, Kr$^-$, kni$^-$, has h uniform expression in the middle of its embryo, and in double mutant, hb$^-$, kni$^-$, its expression is repressed. Accordingly, Kr is the inhibitor of h expression when it works alone.[13] But in wild type embryo, the 4-th stripe of h is expressed in the middle of the domain of Kr expression. (See Fig. 12) As a conclusion, to make the learned network more confident, it is necessary to give the network all known information including mutant phenotype to learn. Then we can check the structure of the learned network, or its behavior for unknown input pattern. In other words, because the more information exists, the less the number of realization of the network that satisfies required I/O relation becomes, it is necessary to give the network mutant phenotype information to restrict the realized solution form.

The problem is how to learn whole data efficiently and how to make the network easily understood. Let's take a method to construct a simple network from mutant data for first and construct more complex one by addition of new data. To test the performance of the instrument and see how analogue function embedded in the learned network, I gave the network partial data to learn. (Fig. 14)

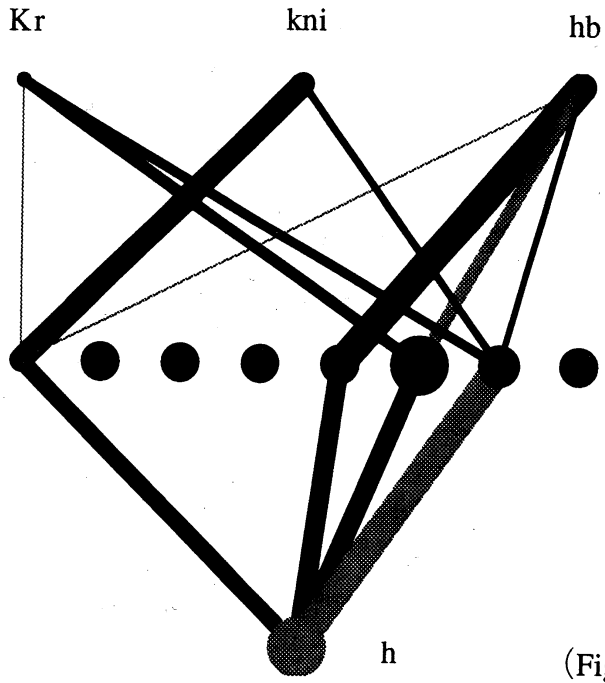
In figure 14, mutant data and the expression around specific stripe of h is summarized. In the first step, the effect of Kr and kni that is in the right side of dashed line in Fig. 14 is learned. The learned network is shown in figure 15. In the second step, started with learned network of the first step, whole data in figure 14 is learned, that means addition of hb information to the network. The learned network is shown in figure 16. Let's consider how to read out the information embedded in the learned network. How the network shows analogue interaction in which the gene expression only occurrs when the two gene products have specific concentration relation?



(Fig.14)



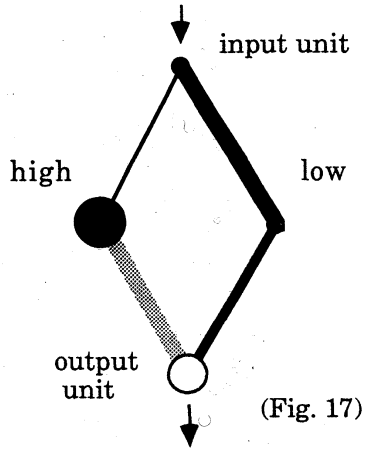
(Fig.15)



(Fig. 16)

The learned networks show different connection patterns according to initial states, but the number of hidden units is nearly constant. Here, to restrict the realized network structure, the parameters of the network are constrained. From experiment, it is found that h turns on this time of development even if there's no input. Therefore, while the learning proceeds, the thresholds of neural units are fixed, to negative for output unit, to positive for hidden ones, respectively.

From fig. 14, it is obvious that the same gene product has different effect on the output depending on its concentration or existence of other factors. This fact is seen on the learned network as an input unit has plural connections to hidden units. A thick connection to a hidden unit that has small threshold (= easily activated) means the low concentration effect of its input unit, and a thin connection to a hidden unit that has large threshold (= hardly activated) means the high concentration effect of its input unit.



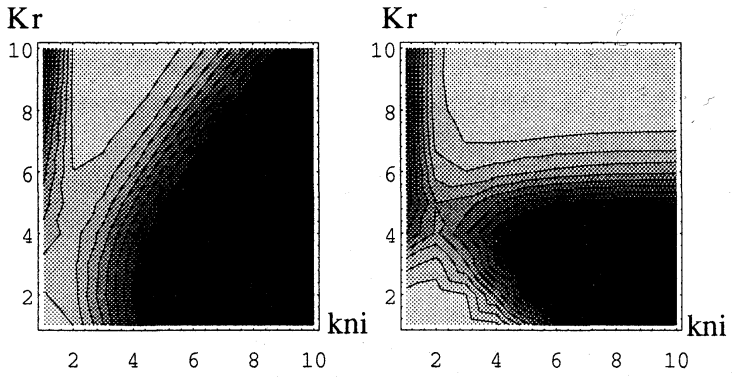
(Fig. 17)

Figure 17 illustrates an input unit that activates output unit at low concentration and inhibits output unit at high concentration. In figure 16, kni has the same kind of effect when it works alone on the expression of h.

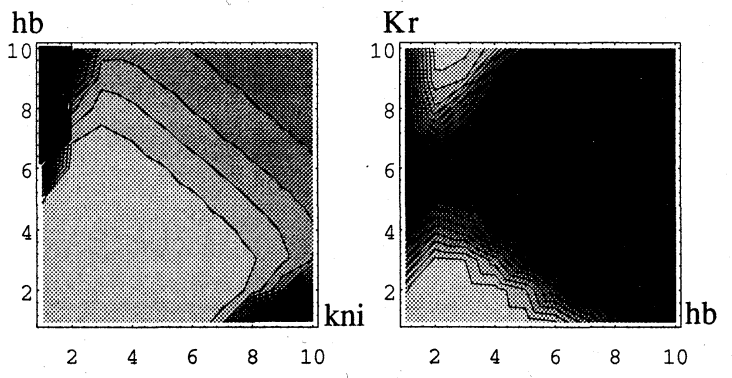
In the actual system, similar facts reported for eve expression. [13)~14)] Eve gene has nearly 10 kb regulatory region upstream of its transcriptional region. There's necessary and sufficient domain for specific stripe expression in it. If the cloned regulatory region is transfected into the embryo with reporter gene, the expression of reporter gene at the site of a specific stripe strictly depends on existence of that region. Eve 2, 3, 7-th stripes have corresponding domain in its regulatory region. Eve's 2-nd and 3-rd stripes are expressed in the different concentration relation of hb and Kr.

There are many binding sequences of hb and Kr gene product in the regulatory domain of eve 2-nd and 3-rd stripe, which suggests the threshold functional regulation by cooperative effect. The affinity of binding sequence in each regulatory domain, has negative correlation to the concentration of its binding factors at that stripe. For example, the affinity of Kr binding sequence in the regulatory domain for 2-nd eve stripe where the concentration of Kr is low, is higher than that for 3-rd eve stripe where that is high. This fact is realized on the learned network as different hidden units like in figure 17.

The way of thinking above, we can read the learned network considerably. In addition, if you want, two or three body interaction can be drawn out from the learned network. (Fig. 18, Fig. 19)



(Fig. 18)



(Fig. 19)

In figure 18, left picture shows the result of the first learning. Right picture of figure 18 and figure 19 show the result of the second learning. All of them illustrate the output of the learned network with varying the normalized input of two factors. Lower left side of each graph is origin, namely, no input region. The higher the output is, the brighter the graphics becomes. One can see easily in what area of input space the gene h expresses. Adding new information to learn, the network can be used as predictor for unknown input pattern.

Reference

(1) Funahashi. K : On the approximate realization of continuous mappings by neural networks, Neural Networks, Vol. 2, 183-192 (1989)

(2) 麻生英樹 : 誤差逆伝播学習の数理的性質, 電気情報通信学会技術研究報告, PRU89-14 (1990)

(3) 石川眞澄 : 忘却を用いたコネクショニストモデルの構造学習アルゴリズム, 人工知能学会誌, 5巻, 5号, pp595-603 (1990)

(4) Wilkinson. D. G and Krumlauf. R : Molecular approaches to the segmentation of the hindbrain, TINS, Vol. 13,No. 8 (1990) modified.

(5) Govine. S & Steward. R : Dorsoventral pattern formation in Drosophila, Trends in Genetics, Vol. 7, No. 4, 119-125 (1991)

(6) Beardsley. T : Smart Genes, Scientific American, Vol. 265, No. 2, 72-81 (1991)
(7) Pankratz. M. J and Jäckle. H : Making stripe in the Drosophila embryo, Trends in Genetics, Vol. 6, No. 9 : 287-292 (1990)
(8) Stanojevic. D, Small. S, Levine. M : Regulation of a segmentation stripe by overlapping activators and repressors in the Drosophila embryo, Science, Vol. 254, 1385-1387 (1991)
(9) Driever. W and Nüsslein-Volhard. C : The bicoid protein is a positive regulator of hunchback transcription in the early Drosophila embryo, Nature, Vol.337, 138-143 (1989)
(10) Casanova. J : Pattern formation under the control of the terminal system in the Drosophila embryo, Development 110 : 621-628 (1990)
(11) Meinhaldt. H : Models for maternally supplied positional information and the activation of segmentation genes in Drosophila embryogenesis, Development 104 suppl : 95-110 (1988)
(12) Carroll. S. B. and Vavra. S. H : The zygotic control of Drosophila pair-rule gene expression II. Spatial repression by gap and pair-rule gene products, Development 107 : 673-683 (1989)
(13) Stanojevic. D, Hoey. T, Levine. M : Sequence-specific DNA-binding activities of the gap proteins encoded by hunchback and Krüppel in Drosophila, Development 107 : 673-683 (1989)
(14) Small. S, Kraut. R, Hoey. T, Warrior. R, Levine. M : Transcriptional regulation of a pair-rule stripe in Drosophila, Genes & Development 5 : 827-839 (1991)