

## RNA 干渉の被食者捕食者モデルによる動態解析

上田 尚学<sup>†1</sup> 荒木 紀子<sup>†1</sup>  
吉川 智啓<sup>†1</sup> 富永 大介<sup>†2</sup>

RNA 干渉 (RNAi) の動態を表現する数理モデルとして、分子機構に基づいた微分方程式系がいくつか提案されている。これらは RNAi に関わる各要素について立式しているためパラメータ数が多く、また実際の観測が不可能である変数を複数含んでおり、モデルを同定するためのデータの取得は容易ではない。

我々は RNAi の作用が被食者/捕食者系として捉えられることに着目し、ロトカ・ヴォルテラ型の競争モデルが十分に RNAi の動態を表現できることを、HeLa 細胞のアポトーシス系の観測のモデリングにより確認した。これにより、直接には観測できない siRNA の効果の強さを表すことができる。

### Prey-Predator model for dynamics of RNA interference

TAKANORI UEDA,<sup>†1</sup> NORIKO ARAKI,<sup>†1</sup>  
TOMOHIRO YOSHIKAWA<sup>†1</sup> and DAISUKE TOMINAGA<sup>†2</sup>

Several Ordinary Differential Equation (ODE) models have been proposed to represent RNA Interference (RNAi) dynamics to today. These consist of equations for all each element involved in RNAi, so they have many parameters. They also have 'hidden variables' which can not be observed by experiments. These facts make identification of model parameters difficult.

RNAi phenomenon can be considered as 'Prey/Predator' system. We applied the modified competitive Lotka-Volterra model to dynamics of apoptosis of HeLa cells by siRNA transfection. We show the model represents the dynamics of the phenomenon and also represents hidden dynamics of siRNA effect.

<sup>†1</sup> 株式会社サイトパスファインダー  
CytoPathFinder, inc

<sup>†2</sup> 独立行政法人産業技術総合研究所

National Institute of Advanced Industrial Science and Technology

### 1. はじめに

真核生物のいくつかのモデル生物において、生きて細胞の中に短い siRNA (small interfering RNA、短鎖干渉 RNA) と呼ばれる長さが 21 から 23 塩基対の二本鎖 RNA が細胞外から入ることにより、細胞内の特定の RNA が分解される現象が知られている。これは RNA 干渉 (RNA interference, RNAi) と呼ばれる現象であり<sup>1)2)</sup>、これを利用すると生細胞内の任意の RNA を人為的に分解でき、したがって任意の特定の遺伝子の機能を解析することができることから、分子生物学の研究において非常に広く知られている。

細胞には本来的に、細胞外の RNA を取り込む機構が備わっており、それによって細胞内に入った siRNA は二本の一本鎖に分かれ、複数のタンパク質とともに RISC (RNA-induced silencing complex) と呼ばれる複合体を作り、標的となる RNA 分子に結合することで、それを切断する。標的 RNA が mRNA であれば、それから作られるタンパク質は機能しないものになるか、あるいは翻訳が行われない。

RNAi の分子機構は簡潔なものではなく、未解明の部分もある。また細胞の中に入った siRNA の量やその効果を、定量的に測定することは非常に困難である。しかし、siRNA に影響を受けた細胞の挙動を数理モデル化して解析することでそこにアプローチできる。

RNAi の動態の数理モデルはこれまでにいくつか提案されている<sup>3)4)5)6)</sup>。これらはいずれも質量作用則に基づく線形の常微分方程式であり、十分な観測データがあれば、あるいはデータにレポータータンパクや siRNA/dsRNA の式に含まれるパラメータ、感度解析などの結果を合わせて見れば、siRNA の働きの強さが分かり、シミュレーションもできる。RNAi の分子機構は現在も詳細が解明されているとは言えないが、これらのモデルは分かっている範囲での分子機構を基に作られている。そしてそのために、方程式系の規模が大きくパラメータも多い (表 1)。さらに観測データを得ることが出来ない変数 (隠れ変数) が含まれている。そのため現実的にはモデルの自由度に比べて十分なデータを用意することは容易ではなく、パラメータ値の同定を困難にしている。

RNAi の機構は分子生物学分野において現在盛んに研究されており、これまで知られていなかった分子機構がしばしば見つかる。これまでに提案されたモデルはそういった系には適用できず、したがって RNAi の完全なモデルを作ることは現在ではできない。この状況におけるよいモデルとは、一つには、少なくとも現在までに観測されている現象をよく再現し、かつ冗長さのない簡素なモデルである。

RNAi は siRNA が特定の RNA を分解する働きである。siRNA および標的となる RNA

の両方はそれぞれ、人為的あるいは自然の機構により供給され、分解されるが、これは食物連鎖における捕食者と被食者の関係に似ている。そこで我々は、RNAi のダイナミクスを被食者/捕食者モデルで表現できると考えた。

被食者/捕食モデルはロトカ・ヴォルテラ・モデルとその発展系として多く提案されているが<sup>7)</sup>、我々はその中でも変数の数とパラメータ数が少ないものに着目した。被食者と捕食者の個体数をそれぞれ  $x$ 、 $y$  としたとき、もっとも素朴なロトカ・ヴォルテラ・モデル

$$\frac{dx}{dt} = x(a - by) \quad (1)$$

$$\frac{dy}{dt} = y(cx - d) \quad (2)$$

は環境許容力 (carrying capacity) を考慮しておらず、捕食者の影響が非常に小さくなったときに被食者数が発散していくという欠点がある。これは、相手がいないときの被食者、捕食者の個体数はロジスティック・モデルに従うようにモデルを修正することで解消される。修正したモデルはロジスティック・ロトカ・ヴォルテラ・モデル<sup>8)</sup> などと呼ばれ、以下の式で表される。

$$\frac{dx}{dt} = x(a - bx - ey) \quad (3)$$

$$\frac{dy}{dt} = y(f - cx - dy) \quad (4)$$

ここで RNAi の標的 RNA として、生細胞内でアポトーシス (機能的細胞死) を抑制する働きをしているものを考えると、siRNA によってこれが分解されるとアポトーシスが生じる、つまり細胞が死ぬことになり、上式の  $x$  は細胞の数、 $y$  は siRNA の効果の強さ (細胞を殺す働き) と解釈できる。細胞数は多くの場合、環境許容力の制約を受けるため、ロジス

表 1 RNAi の微分方程式モデルにおける規模の比較。equations: 各モデルの連立微分方程式系としての次数、parameters: 各モデルに含まれる自由パラメータ数。  
 Table 1 comparison of dimensions and numbers of parameters of ODE models of RNAi. equations: a number of differential equations which consist of each model, parameters: a total number of parameters in a model.

	equations	parameters
Bergstrom <sup>3)</sup>	4	12
Bartlett <sup>4)</sup>	12	27
Groenenboom <sup>5)</sup>	17	14
Marshall <sup>6)</sup>	4	8
<b>our model</b>	<b>2</b>	<b>4</b>

ティック・モデルで表現される。一方で siRNA の効果の強さは、個体数や分子数などのなんらかの具体的な物理量を表現するものではなく、無次元量であり、制約を設けるのが適切かどうかはわからない。そこで、ロジスティックな挙動 (飽和する挙動) を細胞数にだけ取り入れたモデルを RNAi 現象のモデルとして選ぶことにすると、上式からパラメータを二つ減らすことができ、モデルは以下ようになる。

$$\frac{dx}{dt} = x(a - bx - y) \quad (5)$$

$$\frac{dy}{dt} = y(cx - d) \quad (6)$$

以降このモデルを上田モデルと呼ぶ。 $x$  (細胞数) が観測できる量であるのに対し、 $y$  (siRNA の効果の強さ) は直接には観測できない隠れ変数である。この論文では、上田モデルは RNAi によるアポトーシスを観測した系にうまくフィットさせることができ、隠れ変数である siRNA の働きがモデリングされることを示す。

## 2. 方 法

### 2.1 アポトーシスの観測

まず、モデリングの対象となる現象を観測し、時系列データを取得する。ここでは HeLa 細胞のアポトーシスについて行う。市販の 6 種類の siRNA (いずれも HeLa 細胞においてアポトーシスを引き起こすもの) を固相トランスフェクション<sup>9)</sup> (図 1) により細胞に導入し、細胞数の増減を時間を追って測定する。この測定をポジティブ・コントロールと呼ぶ。ここで

表 2 観測に用いた siRNA。ACD は 4 種類の siRNA の混合、VHP、HP、MP は同じ CDC2 遺伝子の mRNA をターゲットとするが siRNA の配列を変更することで RNAi の効率を調整したもの、KIF と PLK はそれぞれ同名の遺伝子の mRNA をターゲットとするものである。

Table 2 siRNAs for the experimental observation. ACD is a cocktail of 4 siRNAs. A target of VHP, HP and MP is same (CDC2 gene). RNAi efficiency of these three siRNAs varies by changing sequences. KIF and PLK target genes which has same name.

Positive control	
ACD	QIAGEN:1027299 (AllStars Hs Cell Death Control siRNA)
KIF	QIAGEN:SI03019793 (Hs_KIF11.8)
PLK	QIAGEN:SI02223837 (Hs_PLK1.6)
VHP	QIAGEN:1027273 (Very High Potency Hs_CDC2 siRNA)
HP	QIAGEN:1027274 (High Potency Hs_CDC2 siRNA)
MP	QIAGEN:1027275 (Moderate Potency Hs_CDC2 siRNA)
Negative control	
Neg	QIAGEN:1027310 (Negative Control)

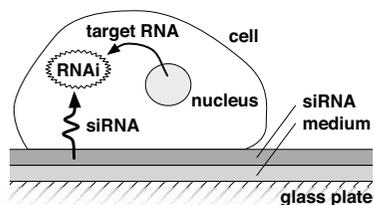


図 1 固相トランスフェクション。ガラスプレートの上に培地、その上に siRNA と導入効率を向上するための試薬の混合物を塗布し、その上に細胞を置くことで、培養しながら siRNA を細胞内に導入する。細胞内に入った siRNA は核から出てきたターゲットの RNA を分解する (RNA 干渉現象)。

Fig. 1 Solid state transfection. Glass plate is coated by culture medium and siRNA. siRNA is mixed with matrix which improves transduction efficiency. Transduced siRNA decomposes RNAs which comes from the nucleus.

は米 Dharmacon 社から販売されている siGENOME SMARTpool の Apoptosis Library に含まれている米国 QIAGEN 社のものを用いた (表 2)。これは一つの標的となる細胞内 RNA に対し、1~4 種類の配列の siRNA を混合したものであり、複数の配列の混合の場合はそれらを同時に細胞内に導入させることで目的の効果を向上するものである。

また対照実験として、細胞内のどの RNA も分解しない siRNA を導入した場合を観測する。これを (どの RNA も分解しない、RNAi 効果を引き起こさないという意味で) ネガティブ・コントロールと呼ぶ。これは、細胞外から siRNA の分子を細胞内に取り込むことによる影響を受けた状態での細胞数の挙動を測定するものである。これもポジティブ・コントロール条件と同じ QIAGEN 社の製品である Negative Control siRNA (表 2) を用いた。

細胞を培地に播種し培養を開始した時刻を 0 として、2.4、5.4、8.0、20.7、29.0、44.9、52.7、68.7、76.8、92.7 時間の各時刻に細胞数を測定する。細胞数は、培地に光を当てその透過量を見ることで、相対的な無次元量として測定される。細胞内には、播種の時点から siRNA が進入していき、それによりアポトーシスが起り始めることで、siRNA の効果が落ちていくと同時に細胞の数も減少していく。

なおネガティブ・コントロール、ポジティブ・コントロールのいずれの場合も、4 回の繰り返し実験を行う。つまり同一条件 (同じ siRNA) 下で上田モデルにおける変数  $x$  が 4 回観測され、各観測ではサンプリング時刻は 10 点である。

## 2.2 モデル

上田モデルは微分方程式系による被食者捕食者系のモデルであり、被食者は生きている細

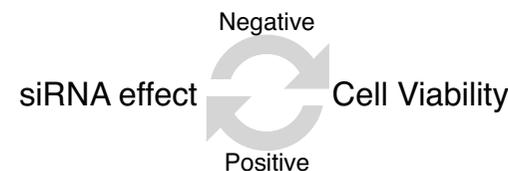


図 2 上田モデルにおける被食者と捕食者の解釈。被食者は細胞 (あるいは細胞の活性)、捕食者は siRNA (あるいは siRNA の細胞内での活性) である。siRNA の活性が強くなれば細胞の数 (細胞の活性) が減り、それにより siRNA の細胞内への取込む活動が弱くなる。

Fig. 2 Physical interpretation of Ueda model. Prey is cell (or viability of cells) and predator is siRNA (or siRNA effect in the cell). High siRNA effectiveness damages cells, then transduction of siRNA into cells decreases.

胞、捕食者は siRNA の効果の強さという解釈に基づいている (図 2)。したがって被食者は捕食者からネガティブな影響を受け、かつ捕食者は被食者からポジティブな影響を受け、それぞれの量を経時的に変化させる。

式 (5)(6) で表されるモデルには 4 個のパラメータ  $a$ 、 $b$ 、 $c$ 、 $d$  が含まれているが、まず最初に対照実験のデータから  $a$  と  $b$  を求め、次に  $a$  と  $b$  は既知であるとして 6 種類の siRNA それぞれについて  $c$  と  $d$  を求める。

対照実験においては siRNA による細胞死が起こらないので、式 (5)(6) において  $y$  を省略したモデルを当てはめればよい。この場合のモデルは以下の形になる。

$$\frac{dx}{dt} = x(a - bx) \quad (7)$$

観測データと微分方程式の振る舞いを比較するためには、微分方程式系を数値積分することが必要であり、そのためには変数の初期値が必要である。ここではパラメータ  $a$ 、 $b$  および変数  $x$  の初期値  $x_0$  の最適値を発見的探索手法である遺伝的アルゴリズム (genetic algorithm, GA) により決定する。GA における探索空間はパラメータ値の対数値とし、遺伝子型には実数値を直接用いる (Real-coded GA)<sup>10)</sup>。交叉操作には UNDX<sup>11)</sup> を、世代交代モデルには MGG<sup>12)</sup> を用いる。観測データとして時系列データが与えられるとき、これは上田モデルにおける変数  $x$  の観測値と解釈される。評価関数は、観測データと数値積分結果の相対誤差の平均値とし、数値積分には四次のルンゲ・クッタ法を用いる。観測データは繰り返し実験の数だけ与えられるので、数値積分の結果は (同一条件下で観測された) 4 つの時系列データ (サンプリング点数 10) と比較され、その 40 点における相対誤差が平均されて評価関数値となる。

次に、上で求めた  $a$ 、 $b$ 、 $x_0$  を既知であるとして、 $c$  と  $d$  および  $y$  の初期値  $y_0$  を同時にチューニングすることで細胞死を起こす siRNA を導入した場合のデータにモデルをフィッティングする。これらの値も  $a$ 、 $b$ 、 $x_0$  と同じ最適化手法により決定する。

### 3. 結 果

まず siRNA として細胞内のどの RNA も分解しないものを導入した場合 (対照実験) を観察し、細胞数の変化を上田モデルにおける変数  $x$  にフィッティングすることでパラメータ  $a$  と  $b$  を求めた。観測データとフィッティングカーブを図 3 (A) に示す。

そして細胞内の RNA を分解することで RNAi 現象をおこし、その結果として細胞死を誘導する siRNA を用いた場合 (ポジティブ・コントロール条件) に対して、上田モデルにおける  $a$  と  $b$  を先に求めた値で固定し、 $x$  を観測値にフィッティングさせることでパラメータ  $c$  と  $d$  の最適値を求めた。siRNA はネガティブ・コントロールを入れて 7 種類あり、それぞれ 4 回のリピート実験を行った。各 siRNA について、4 セットのデータに対する二乗誤差和が最小となるようなパラメータの値を探索した。観測データとフィッティングカーブを図 3 (B-G) に示す。

フィッティングしたカーブは観測データにおける標準偏差の範囲に入っており (図 3 B-G の各プロット中の破線)、フィッティングは成功したと言える。上田モデルにおける変数  $y$ 、つまり siRNA 効果の強さのダイナミクス (同図、各点線) が、フィッティングにより得られたパラメータ (表 3) から計算される。これは直接に観測することの出来ない隠れ変数であるが、隠れ変数は  $y$  ひとつのみであり、可観測の変数  $x$  から直接に計算される。 $y$  が得られたことから、RNAi による細胞死における siRNA の効果の動態をモデリングできたと言える。

### 4. 考 察

対照実験 (ネガティブ・コントロール) における上田モデル (式 7) は、パラメータが 2 個であり、変数の初期値を入れて自由度が 3 のモデルである。これは微分方程式としては解析的に原関数を求めることができ、それはロジスティック関数 (シグモイド関数)、

$$x(t) = \frac{ax_0}{bx_0 + (a - bx_0) \exp(-at)} \quad (8)$$

になる。ここではサンプリング・データが 10 点あることから、パラメータ値はほぼ一意に決まると言ってよい。そして、これに基づいて求められるポジティブ・コントロール条件で

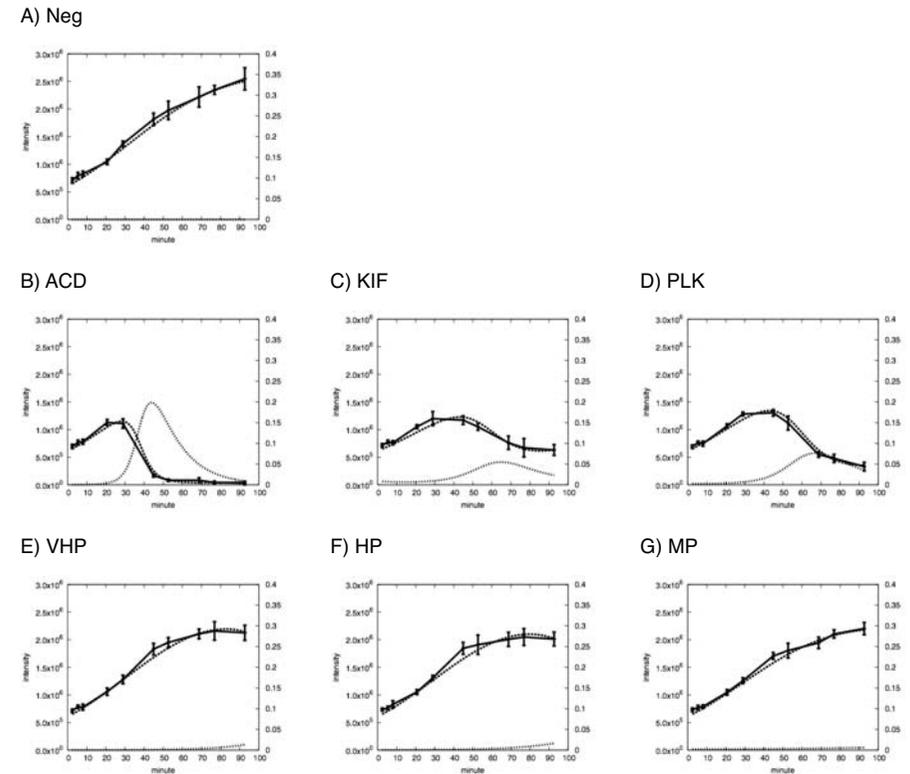


図 3 ネガティブ・コントロール、ポジティブ・コントロールの各 siRNA 導入時の細胞数の変化 (観測データ、実線) と細胞数 (破線) と siRNA の効果 (点線) の上田モデルによる表現。ネガティブ・コントロール条件では siRNA が RNAi を引き起こさないため、点線はプロットされない。

Fig. 3 Changing cell counts in time under the conditions of negative and positive controls (solid) and curves of optimized Ueda model for each condition (dash: cell counts, dot: siRNA effect). Neg siRNA does not cause RNAi in the negative control condition, therefore no dotted line is plotted in 'A) Neg'.

のパラメータも、やはり 10 点のデータに対して自由度は 3 であり、求められたパラメータ値や隠れ変数  $y$  の挙動は、大域解であるか、もしくは他の局所解を探すのは容易ではないと考えている。

上田モデルの 4 つのパラメータ ( $a, b, c, d$ ) のうち、 $b$  はロジスティック・ロトカ・ヴォルテラ・モデルにおいて環境収容力と呼ばれている<sup>8)</sup> が、上田モデルでも同じ意味に解釈できる。パラメータ  $c$  は細胞の生命力が siRNA の強さに及ぼすポジティブな影響にかかっており、ここに示した実験系においては、siRNA の細胞外から細胞内への取込みの効率を表す係数であると解釈できる。

パラメータ  $c$  および  $d$  の値を対数空間にプロットすると、二つのグループに分かれることが示される (図 4)。これはアポトーシスを引き起こす効果の強いものと弱いものにわかれている。多くの siRNA についてパラメータ値を求めることで、siRNA の機能や効率の数値化や分類ができるかもしれない。

パラメータ  $c$  と  $d$  について、その値および対数値の相関係数は低い (0.25795 および 0.10525) ため、これらが線形の関係にあるとは考えにくく、上田モデルよりも更にパラメータの少ない簡潔なモデルを求めるのは容易ではないと考えている。また上田モデルは、すでに簡潔なロジスティック・ロトカ・ヴォルテラ・モデルから、さらにパラメータを一つ減らしたものである。この点からもこれよりもモデルを簡略化することの困難さが予想される。

パラメータ・フィッティングにおける評価関数は、上田モデルにおける変数  $x$  の数値解と実験観測データの相対二乗誤差和である。数値解が得られない場合は評価関数の値は未定義になるが、パラメータ空間のどこがそういった領域になるかは、数値解を計算してみるまで分からない。そのため最急降下法などの解析的な手法を使うために必要な探索出発点を定めるための手がかりがなく、発見的な探索手法を用いざるを得ない。また出発点がうまく

表 3 各 siRNA の場合を表す上田モデルのパラメータ。  $a, b, x_0$  はすべての実験条件で共通である。  $c, d, y_0$  はネガティブ・コントロール条件では定義されない。

Table 3 Parameter values of Ueda model for each siRNA condition.  $a, b$  and  $x_0$  are common for all conditions.  $c, d$  and  $y_0$  are undefined for the negative control condition.

	Neg.	ACD	KIF	PLK	VHP	HP	MP
$a$				3.949e-2			
$b$				1.445e-8			
$x_0$				6.512e+5			
$c$	-	2.694e-7	1.784e-7	1.337e-7	2.498e-8	2.243e-8	7.258e-9
$d$	-	8.087e-2	1.506e-1	9.045e-2	3.091e-7	1.118e-7	6.509e-8
$y_0$	-	2.947e-4	8.382e-3	2.837e-3	3.458e-4	6.760e-4	2.305e-3

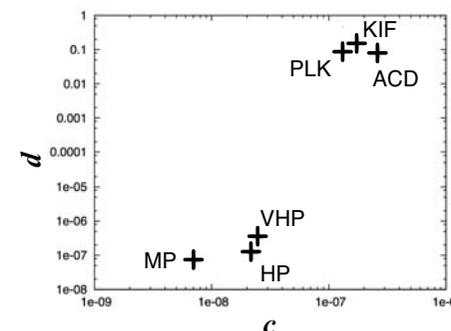


図 4 各実験条件によるパラメータ値の分布。上田モデルにおいては  $a$  および  $b$  はここでのすべての実験条件について共通なので、残りの二つのパラメータ  $c$  および  $d$  のプロットを示す。

Fig. 4 Distribution of parameter values of each condition. Since values of parameters of Ueda model  $a$  and  $b$  is common to all positive and negative conditions, values of other two parameters,  $c$  and  $d$ , are shown.

決められたとしても、本論文における最適化問題の評価関数のプロファイルは、最適解の近傍で非常に急峻 (steep) であり、かつ他の領域では勾配が小さい。解析的な手法はこういった評価関数には適していない。微分方程式系のパラメータ・フィッティングは長く研究されており、遺伝的アルゴリズムの有効性が示されている<sup>13)</sup>。

本研究で提案する上田モデルは、RNAi の分子機構に基づくものではない。同様に詳細な分子機構に基づかずに系のネットワークスキームを反映する正規系の微分方程式モデルとして S-system<sup>14)</sup> があるが、今後、S-system によるモデリングの可能性を検証したいと考えている。

## 参考文献

- 1) Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., Tuschl, T.: Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells, *Nature*, Vol.411, pp. 494–498 (2001).
- 2) Fire, M., Mello, C. C.: Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*, *Nature*, Vol.391, pp. 744–745 (1998).
- 3) Bergstrom, C. T., McKittrick, E., Antia, R.: Mathematical models of RNA silencing: Unidirectional amplification limits accidental self-directed reactions, *Proceedings of the National Academy of Sciences*, Vol.95, pp. 1518–1523 (1998).

- ings of the National Academy of Sciences of the USA*, Vol.100, No.20, pp. 11511–11516 (2003).
- 4) Bartlett, D. W., Davis, M. E.: Insights into the kinetics of siRNA-mediated gene silencing from live-cell and live-animal bioluminescent imaging, *Nucleic Acids Research*, Vol.34, No.1, pp. 322–333 (2006).
  - 5) Groenenboom, M. A. C. Hogeweg, P.: The dynamics and efficacy of antiviral RNA silencing: A model study, *BMC Systems Biology*, Vol.2, 28 (2008).
  - 6) Marshall, W. F.: Modeling Recursive RNA Interference, *PLoS Computational Biology*, Vol.4, No.9, e1000183 (2008).
  - 7) Berryman, A. A.: The Origins and Evolution of Predator-Prey Theory, *Ecology*, Vol.73, No.5, pp. 1530–1535 (1992).
  - 8) Gause, The struggle for existence, 1934 Gause, G.F.: *The struggle for existence*, Hafner Publishing, NY (1964).
  - 9) Yoshikawa, T., Uchimura, E., Kishi, M., Funeriu, D. P., Miyake, M., Miyake, J.: Transfection microarray of human mesenchymal stem cells and on-chip siRNA gene knockdown, *Journal of Controlled Release*, Vol.96, pp. 227–232 (2004).
  - 10) Ueda, T., Koga, N., Ono, I., Okamoto, M.: Application of Numerical Optimization Technique Based on Real-coded Genetic Algorithm to Inverse Problem in Biochemical Systems, *Proceedings of Genetic and Evolutionary Computation Conference (GECCO 2002)*, p. 701 (2002).
  - 11) 小野 功, 佐藤 浩, 小林 重信: 単峰性正規分布交叉 UNDX を用いた実数値 GA による関数最適化, *人工知能学会誌*, Vol.14, No.6, pp. 1146–1155 (1999).
  - 12) 佐藤 浩, 小野 功, 小林 重信: 遺伝的アルゴリズムにおける世代交代モデルの提案と評価, *人工知能学会誌*, Vol.12 No.5 pp. 734–743 (1996).
  - 13) Kikuchi S, Tominaga D, Arita M, Takahashi K and Tomita M: Dynamic modeling of genetic networks using genetic algorithm and S-system, *Bioinformatics*, Vol.19, No.5, pp.643–650 (2003).
  - 14) Savageau M: *Biochemical systems analysis: a study of function and design in molecular biology*, Addison Wesley, Reading, Massachusetts (1976).
-