

DNA 演算シミュレータの構築

下田 明宏[†] 涌井 智寛[†] 星野 哲男[†]
畠山 正行^{††} 荒木 俊郎^{††}

DNA 計算の分野において計算アルゴリズムが幾つか提案されている。提案された DNA 計算アルゴリズムは分子生物学的な DNA 計算は実験によって実現されるものであり、実験を行って検証しなければならない。しかし実験にはいくつかの困難があり容易ではない。そこで DNA 計算シミュレータがこれらの実験の前段階として有効である。しかし現存するシミュレータは種類も少なく、また、時間的に連なる複数の操作手続きを連続して実行できるようには作られていない。そこで我々は連なった操作が行える様な仕組みを実現した DNA 計算シミュレータを開発した。

開発したシミュレータを既存の VNA シミュレータと比較した。その結果、計算精度は共通機能についてはほぼ同じ精度であることが分かった。それに加えて、開発したシミュレータは、連続した操作を実現した以外にも幾つかの新しい特徴を持っている。それは入出力の DNA 分子の種類数を大幅に増やしたために、DNA 分子のうち従来のシミュレータでは無視されていたごく少数の DNA 分子も保持でき、シミュレーションに組み入れられたことである。今後の課題は本シミュレータの妥当性と実現性を分子生物学的な実験と比較して検証すること、及び、他の複雑な計算シミュレーションに応用できるように機能を拡張することである。

A Development of DNA Computing Simulator

AKIHIRO SHIMODA [†], TOMOHIRO WAKUI [†], TETSUO HOSHINO [†]
, MASAYUKI HATAKEYAMA ^{††} and TOSHIROU ARAKI^{††}

In the DNA computing field, some DNA computing algorithms have been proposed. The DNA computings can be realized based on the bio-molecular experiments. Then, these proposed DNA computing algorithms must be verified using these experiments. A DNA computing simulator is valuable as the preceding stage tool of these experiments. The existing simulators are, however, only a few, and the existing ones can not be used for plural sequential operation procedures. Therefore we have developed a DNA computing simulator that can be used for sequentially constituted DNA computing operation procedures.

The developed simulator has been compared with the VNA simulator and the other one. The results show that the arithmetic accuracy is almost the same in the common functions. In addition, the simulator that we have developed has some new features. That is, since the numbers of input/output molecule kinds have drastically be increased, the small numbers of molecule kinds that are neglected in the existing simulators can also be maintained in the program and treated in the simulation. The future work is to verify the reasonableness and realizability of this simulator by comparing with the bio-molecular experiments, and expanding the functions of this simulator to apply other complex computing simulations.

1. はじめに

1994年に発表された DNA 計算に関する Adleman

[†] 茨城大学大学院理工学研究科
Graduate School of Science and Engineering, Ibaraki University

^{††} 茨城大学工学部情報工学科
Department of Computer and Information Sciences,
Ibaraki University

の研究¹⁾以来、DNA 計算に関する研究が盛んになっている。これらの研究では論理的なアルゴリズムの構築だけでなく分子生物学的な実験を伴っていることも多い。なぜならば分子生物学的な実験こそが DNA 計算の実現そのものだからであり、理論的に構築された DNA 計算のアルゴリズムに沿って計算できるか否かは分子生物学的実験によって確認されることが望ましいからである。

しかし、分子生物学的な実験を行うには多くのコス

トが必要となるため、実験を行う前に計算機を用いたシミュレーションで実現可能性を検証することが良策であると考えた。そこで、我々は考案されたアルゴリズムの実現可能性を検証するためのシミュレータを構築した。

2. 既存のシミュレータとその問題点

2.1 VNA シミュレータ

VNA シミュレータ⁶⁾では、具体的に塩基配列を指定するのではなく、数塩基をまとめたものを仮想塩基として扱う。仮想塩基はアルファベットで表され、大文字と小文字のアルファベットが相補関係にあるものとして扱われる。DNA 分子の濃度は微分方程式に基づいて数値的に計算され、計算結果が指定した閾値より低くなった DNA 分子は無視される。これにより、DNA 分子の組み合わせ爆発を防いでいる。VNA シミュレータではハイブリダイゼーション、熱変性、切断、伸張、ライゲーションを行うことができ、ユーザが指定できるパラメータは 14 種類である。

2.2 VNA シミュレータの問題点

VNA シミュレータには以下のような問題点がある。

- (1) 入力では一本鎖 DNA 分子しか扱えないため、出力される DNA 分子が入力として扱えない場合がある。
- (2) 扱える種類数が、入力が最大 14 種類、出力が最大 10 種類と少ない。
- (3) 仮想塩基では DNA 分子の長さが考慮されていない。
- (4) 制限酵素による切断面を指定できない。
- (5) 濃度が低い DNA 分子が無視される。

3. DNA 演算シミュレータの構築

本研究では VNA シミュレータの持つ問題が起きないような、新規の DNA 演算シミュレータを開発した。

我々のシミュレータは以下の二つから構成される。一つ目は DNA を用いて演算をする際に必要不可欠である塩基配列を生成する配列設計支援ツールである。二つ目は配列設計支援ツールによって生成された塩基配列を実際にアルゴリズムに沿ってシミュレーションする DNA 演算ツールである。本論文では以上二つのツールを総称して DNA 演算シミュレータと呼ぶ。配列設計支援ツールの動作画面を図 1 に、DNA 演算ツールの動作画面を図 2 に示す。

配列設計支援ツールにおいては、配列設計を行う際の方針が有田ら²⁾らによって、ハミング距離による結合力の評価が Garzon ら³⁾によって示されており、

これらを参考とした。配列設計支援ツールの仕様は、

- (1) 遺伝的アルゴリズムを用いて、ユーザが指定した評価値を満たすか、もしくは 1000 世代に達するまで準最適解を求め続ける。
- (2) 生成される DNA 配列は完全な二本鎖 DNA 分子であるが、配列中の塩基を一部削除することにより部分二本鎖 DNA 分子や一本鎖 DNA 分子にも対応できる。
- (3) 生成する配列の長さ及び種類数はユーザが任意に決めることができる。
- (4) 遺伝的アルゴリズムのパラメータである個体数や交叉確率及び突然変異確率はユーザが指定する。
- (5) 特定の場所にユーザが指定する塩基配列を入れることができる。それにより認識部位などをユーザが自由に入れることができる。

である。

DNA 演算ツールでは塩基配列を対象とし、それらを確率的に扱う。出力可能な構造を有する DNA 分子のすべては入力可能であり、さらに出力された DNA 分子は手動で再入力することなく入力として扱うことができる。また、制限酵素の認識部位や切断面も扱うことができる。DNA 演算ツールの仕様は、

- (1) 図 3 に示すような 5 種類の構造の DNA 分子が扱える。
- (2) 扱える DNA 分子の種類数は、CPU は Pentium4 の 3.2GHz、メモリは 2.0GB で OS は Windows2000 の計算機において名前が半角英数字 20 文字で 20bp の DNA 分子が二百万強である。
- (3) 投入する DNA 分子の名前はユーザが任意に定めることができ、シミュレーションの途中で名前の変更もできる。

である。

4. 既存シミュレータとの比較・考察

4.1 ハイブリダイゼーションのシミュレーション

比較のために VNA シミュレータ及び本研究で提案する演算ツールの両者に対して同一の実験に対するシミュレーションを行った。実験は図 4 に示すように "seq1seq2", "seq1", "seq2" の 3 種類の論理配列を投入しハイブリダイゼーションを行うものとした。そして両者のシミュレータによる結果を図 5 に示す。

ハイブリダイゼーションのシミュレーション結果は VNA シミュレータと本研究の両者とも目的となる DNA 分子が多く存在し、他の DNA 分子の存在割

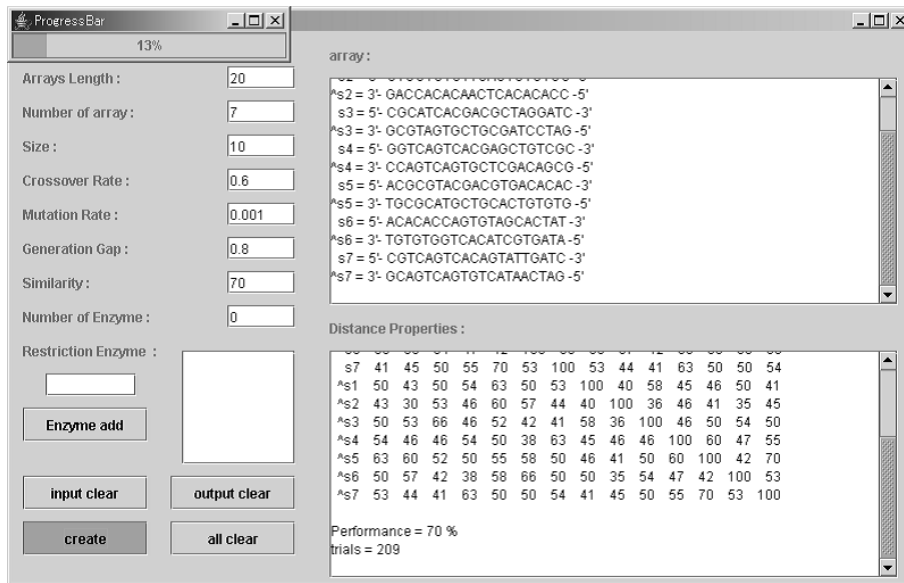


図 1 配列設計支援ツール動作画面
Fig. 1 DNA sequence design support tool in execution

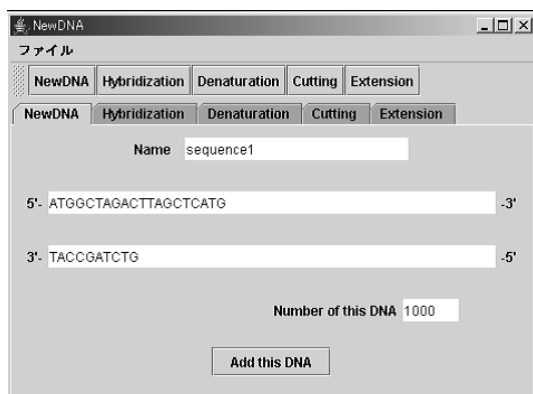


図 2 DNA 演算ツール動作中画面
Fig. 2 DNA computing tool GUI window in execution

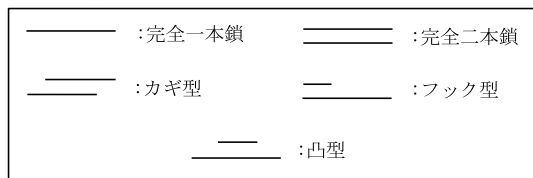


図 3 利用できる DNA 分子の形状
Fig. 3 Straightly formed molecular sequence figuration

合は低くなっているという点で同じだと言える。したがって、この点では両シミュレータはほぼ同様な傾向を持つと言えよう。また、本シミュレータは塩基配列を用いてシミュレーションをしているので仮想塩基を

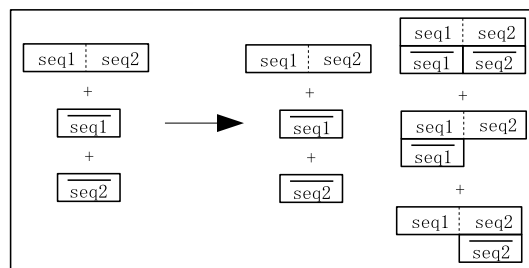


図 4 DNA 演算シミュレータ比較のためのシミュレーション例
Fig. 4 Simulation example for comparing DNA computing simulators

用いている VNA シミュレータよりもミスハイブリダイゼーションの検知の確度が高くできるだろう。

4.2 PCR のシミュレーション

次に DNA の基本操作となっている PCR のシミュレーションを本シミュレータで行った結果を表 1 に示す。なお、表中の "Tem" はテンプレート、" \overline{Tem} " はテンプレートの相補配列、"Pri1" は " \overline{Tem} " に付加するプライマーで "Pri2" は "Tem" に付加するプライマーである。初期投入数は "Pri1" 及び "Pri2" は 10000 個とし、"Tem" 及び " \overline{Tem} " は 100 個とし、ハイブリダイゼーションの試行回数は各サイクル 20000 回とした。

次に VNA シミュレータにおいても同様のシミュレーションを行った結果を表 2 に示す。入力には "Pri1" の仮想塩基配列は "A" で "Pri2" の仮想塩基配列は "b" とし濃度の初期値をそれぞれ 1.0 とし、"Tem" の仮想塩

抽出分子	反応後の存在比率	
	VNA	本研究
$\frac{\text{seq1} \cdot \text{seq2}}{\text{seq1} \cdot \text{seq2}}$	0.830	0.854
$\frac{\text{seq1} \cdot \text{seq2}}{\text{seq2}}$	0.041	0.035
$\frac{\text{seq1} \cdot \text{seq2}}{\text{seq1}}$	0.041	0.036
$\frac{\text{seq1} \cdot \text{seq2}}{\text{seq1} \cdot \text{seq2}}$	0.002	0.001
$\frac{\text{seq1}}{\text{seq1}}$	0.043	0.037
$\frac{\text{seq2}}{\text{seq2}}$	0.043	0.037
	1.000	1.000

図 5 両者のシミュレータの比較
Fig. 5 Comparisons of both simulators

基配列は"AB"で" Tem "の仮想塩基配列は"ba"とし濃度の初期値をそれぞれ 0.01 としてシミュレーションを行ったが、表 1 と比較のために各 DNA 分子ともに 10000 倍してある。さらに、結果の DNA 分子に二本鎖 DNA 分子もあるがその DNA 分子は全て変性されているものとみなしている。表 2 において総ステップ数は 600 ステップとし、50 ステップごとの値を記述してある。

表 1 と表 2 の" Tem "分子の数の変化を図 6 に示す。なお表 2 の 50 ステップを 1 サイクルに対応させている。同図を見て分かるように両シミュレータから得られた結果は似たような増加傾向を持っていた。また、最終的に得られるテンプレートの分子数の理論値が 10100 に対して、DNA 演算ツールによるシミュレーションの結果は約 9500 であり、VNA シミュレータによるシミュレーションの結果は濃度を分子数に換算すると約 9570 となった。このことから DNA 演算ツールと VNA シミュレータではほぼ同じ値に収束すると考えられる。

我々の構築した DNA 演算ツールでは 1 試行で 1DNA 分子に対する 1 反応を扱うが、VNA シミュレータでは 1 ステップで複数反応を同時に扱う。PCR は 3 つの反応からなるため、この点が増加傾向の差異として現れていると考えられる。

4.3 部分二本鎖 DNA 分子を入力として用いるアルゴリズムのシミュレーション

部分二本鎖 DNA 分子を入力として用い、反応後の結果を用いてさらに別の反応を行うアルゴリズム⁴⁾が

表 1 DNA 演算ツールにおける PCR のシミュレーション
Table 1 PCR simulation in our simulator

サイクル数	Pri1	Pri2	Tem	Tem
1 サイクル	9935	9938	165	162
2 サイクル	9842	9824	258	276
3 サイクル	9661	9658	439	442
4 サイクル	9406	5683	725	700
5 サイクル	8989	8934	1141	1136
6 サイクル	8341	8256	1789	1814
7 サイクル	7345	7274	2785	2796
8 サイクル	5959	5908	4171	4162
9 サイクル	4263	4277	5867	5793
10 サイクル	2645	2595	7485	7475
11 サイクル	1376	1350	8754	8720
12 サイクル	627	620	9503	9450

表 2 VNA シミュレータにおける PCR のシミュレーション
Table 2 PCR simulation in VNA simulator

ステップ数	Pri1	Pri2	Tem	Tem
50 ステップ	9903	9903	192	192
100 ステップ	9746	9746	346	346
150 ステップ	9469	9469	617	617
200 ステップ	8998	8998	1078	1078
250 ステップ	8241	8241	1821	1821
300 ステップ	7127	7126	2920	2921
350 ステップ	5686	5683	4349	4352
400 ステップ	4115	4110	5919	5924
450 ステップ	2698	2691	7347	7354
500 ステップ	1628	1620	8433	8440
550 ステップ	925	919	9150	9157
600 ステップ	506	506	9578	9573

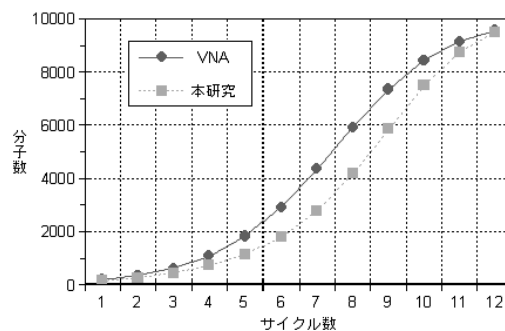


図 6 両シミュレータにおける" Tem "分子の数の変化
Fig. 6 Changes of " Tem " molecules in both simulator

ある。そのアルゴリズムは図 7 に示されるような加算を行うアルゴリズムである。

表 3 に示すパラメータを用いて、配列設計支援ツールで図 7 に示す加算アルゴリズムをシミュレーションするための塩基配列を設計した。表 3 中の"配列の長さ"は例えば図 7 中の"繰上=0"や"+1"などの矩形で

表 3 配列設計指定パラメータ

Table 3 DNA sequence design designated parameter values

パラメータ	入力値
配列の長さ	20
配列の種類数	7
個体の個数	10
交叉確率	0.6
突然変異確率	0.001
世代ギャップ	0.8
類似度	70
認識部位の数	0

囲まれたものの塩基数がそれぞれ 20 塩基であることを示している。

配列設計した結果を表 4 に示す。

表 4 配列設計の結果

Table 4 Result of DNA sequence design

パラメータ	塩基配列
繰上=0	5'-ATCGTCGGCGCAAGACCGCT-3'
繰上=1	5'-CTGGTGTGTTGAGTGTGTGG-3'
0	5'-CGCATCACGACGCTAGGATC-3'
1	5'-GGTCAGTCACGAGCTGTCCG-3'
+0	5'-ACGCGTACGACGTGACACAC-3'
+1	5'-ACACACCAGTGTAGCACTAT-3'
+2	5'-CGTCAGTCACAGTATTGATC-3'

この結果を用いて DNA 演算ツールで図 7 に示すアルゴリズムのシミュレーションを行った。投入する DNA 分子は各 1000 個ずつとし、ハイブリダイゼーションの試行回数は 20000 回とした。なお、抽出の課程では目的の DNA 分子を全て抽出できたものと仮定した。

最終的に抽出される DNA 分子は図 7 中の解分子であり、個数は 721 個であった。理想値と比較すると約 72%抽出できた。一方、VNA シミュレータで同じアルゴリズムのシミュレーションを行った場合には、解分子は理想値の約 44%であった。

VNA シミュレータでの検証において理論値の半分以下しか解分子を得られなかった要因として、VNA シミュレータでは DNA 分子の長さが考慮されていないことが挙げられている。仮想塩基が相補であると判定されると、本来は結合が安定しないアルファベット同士でも結合してしまう。そのため、本来は生成され難いはずの DNA 分子の数が増加し、生成されるべき解分子の数が減少してしまう。

それに対して我々が構築した DNA 演算ツールでは、ユーザが指定するワトソン・クリック相補性の充足率に基づいて結合するか否かを決定する。適切にワトソン・クリック相補性の充足率を指定することで、本来

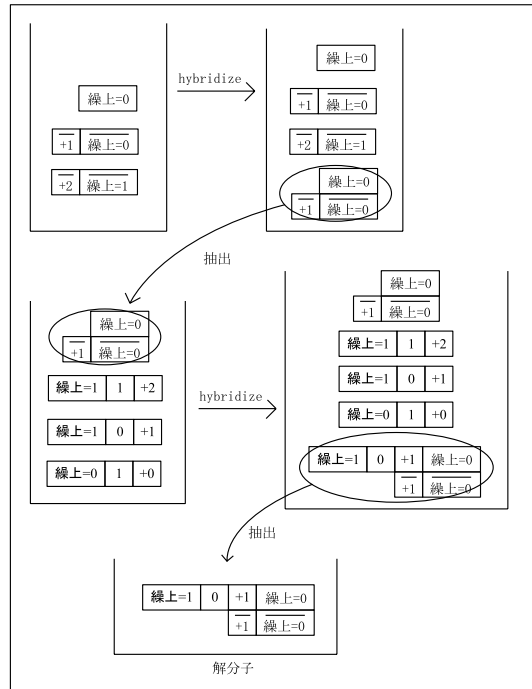


図 7 全加算器を用いたアルゴリズムにおける 2 進数の 1+1 のシミュレーション例

Fig. 7 $1_{(2)} + 1_{(2)}$ computing simulation example using full adder algorithm

は結合が安定しない塩基配列同士では結合することはない。

また、VNA シミュレータでは二本鎖の部位をもつ DNA を入力できないため、二本鎖の部位を除いた DNA を入力する必要があった。それに対して、我々が構築した DNA 演算ツールでは二本鎖の部位をもつ DNA を入力可能であるため、出力された DNA を加工することなくシミュレーションを継続することができた。

以上から、部分二本鎖 DNA を扱える点と、DNA 分子の長さが考慮される点で我々の構築した DNA 演算ツールの方が優れているといえる。

5. 結 論

我々が構築した DNA 演算ツールは、一本鎖と二本鎖の DNA に関しては VNA シミュレータとほぼ同様なシミュレーション結果を出力することが確認できた。さらに以下の点で VNA シミュレータより優れている。

- 具体的な塩基配列を用いるため、類似度とワトソン・クリック相補性の充足率を適切に設定することにより、ハイブリダイゼーションの確度が高く

なる

- 二本鎖 DNA 分子を入力することが可能である
- 組み合わせ爆発を防ぎつつ, 多くの種類数の DNA 分子を入出力可能である
- シミュレーションの途中で, 必要に応じて任意の DNA 分子を追加可能である
- 制限酵素による切断面が詳細に扱える

以上から, 本研究で構築した DNA 演算シミュレータは, 従来のものと比較して信頼性の面で同等であると考えられる。また, 二本鎖の部位もつ DNA を入力する必要があるアルゴリズムなどの内, VNA シミュレータではシミュレーション不可能なアルゴリズムに対しても有効なシミュレータであるといえる。

6. 今後の課題

今後の課題として以下の項目が挙げられる。

- (1) ゲル電気泳動シミュレーション機能の実装
- (2) シミュレーション手順のファイル入出力
- (3) DNA 演算ツールにおけるデータのファイル入出力
- (4) 分子生物学的実験データとの比較
- (5) 2 つのツールの統合

参 考 文 献

- 1) L.M.Adleman: Molecular Computation of Solutions to Combinatorial Problems. *Science*, 226(November 1994),1021-1024
- 2) 萩谷昌己:分子コンピュータの現状と展望, サイエンス社,2004.
- 3) M. Garzon, R. Deaton, L. F. Nino, Ed Stevens : "Encoding Genomes for DNA Computing", Proc. of the Third Annual Genetic Programming Conf., pp. 684-690, 1998.
- 4) 涌井智寛:DNA を用いた全加算器の提案とその設計/検証, 平成 16 年度茨城大学工学部情報工学専攻修士論文,2005
- 5) 星野哲男, 涌井智寛, 下田明宏, 畠山正行:DNA を用いた乗算の方法の提案とその計算手順の設計, 第 54 回数理モデル化と問題解決研究発表会, 名古屋,2005
- 6) Akio Nishikawa, Masayuki Yamamura and Masami Hagiya: DNA Computation Simulator Based on Abstract Bases, *Soft Computing* Vol.5 No.1, 2001, pp.25-44
- 7) 萩谷 昌己, 横森 貴: DNA コンピュータ, 培風館,(2001)
- 8) G. パウン,G. ローゼンバーグ,A. サローマ:DNA コンピューティング 新しい計算パラダイム, シュプリンガー・フェアラーク東京,(1999)