Available online at www.sciencedirect.com
ScienceDirect
Progress in Natural Science 18 (2008) 633–637
Progress in Natural Science
www.elsevier.com/locate/pnsc

Special report

DNA computing model based on lab-on-a-chip and its application on solving the timetabling problem

Fengyue Zhang [a,*,1], Bo Liu [b,1], Wenbin Liu [c], Qiang Zhang [d]

[a] *Department of Biomedical Engineering, School of Life Science and Technology, Beijing Institute of Technology, Beijing 100081, China*
[b] *Chinese Academy of Inspection and Quarantine, Beijing 100029, China*
[c] *College of Computer Science and Engineering, Wenzhou University, Wenzhou, Zhejiang 325027, China*
[d] *Liaoning Key Laboratory of Intelligent Information Processing, Dalian University, Dalian, Liangning 116622, China*

Received 21 December 2007; received in revised form 16 January 2008; accepted 17 January 2008

Abstract

The essential characteristic of DNA computation is its massive parallelism in obtaining and managing information. With the development of molecular biology technique, the field of DNA computation has made a great progress. By using an advanced biochip technique, laboratory-on-a-chip, a new DNA computing model is presented in the paper to solve a simple timetabling problem, which is a special version of the optimization problems. It also plays an important role in education and other industries. With a simulated biological experiment, the result suggested that DNA computation with lab-on-a-chip has the potential to solve a real complex timetabling problem.

Keywords: DNA computing; Timetabling problem; Lab-on-a-chip

1. Introduction

Since Adleman demonstrated the possibility of solving NP-complete problems by encoding a computer style problem in DNA sequences [1], the DNA computation has become a new vista of computation that bridged computer science and molecular biology. DNA computing provided a massive computational parallelism that allowed us to tackle intractable combinatorial problems exhaustively, in contrast to the exponentially increasing time required by a Turing machine. After Adleman's work, the field of DNA computation has made a great progress, which included mainly some theoretical studies [2–4], solution-based DNA computation [1,5–7] and surface-based DNA computation [8–14]. Generally, the surface-based DNA computation, which manipulates DNA strands immobilized on a surface, usually uses more advanced biological technology and has a potential to be automatized in contrast with the solution-based DNA computing.

As a very promising area of research, the focus of DNA computation is whether it could solve a hard arithmetic (like NP-hard) problem in reality. However, the most complex problem ever solved by DNA computation was a 20-Variable 3-SAT problem by Braich [15]. It was recognized that powering the theoretical computing capacity in a large reaction vessel was meaningless unless a reaction achieved the expected level of performance. Therefore, based on lab-on-a-chip which simply defined as a chip integrated with heaters, valves, pumps, microfluidic controllers, electrochemical and electroluminescent detectors [16–21], this paper proposed a DNA computing model in theory to solve a simple timetabling problem, which is a special version of the optimization problems found in real life situations but is not studied by DNA computation.

* Corresponding author. Tel./fax: +86 010 68914941.
E-mail address: zfywuhan@mail.hust.edu.cn (F.Y. Zhang).
[1] These authors contributed equally to this work.

1002-0071/$ - see front matter © 2008 National Natural Science Foundation of China and Chinese Academy of Sciences. Published by Elsevier Limited and Science in China Press. All rights reserved.
doi:10.1016/j.pnsc.2008.01.007

2. DNA computing model based on lab-on-a-chip to solve a simple timetable

As a special version of the optimization problems found in real life situations, the timetabling problem, which plays an important role typically in education, has been periodically faced by every schools, colleges and universities all over the world. Like other optimization problems, the problem is computationally NP-hard. General timetabling problem could be defined as the scheduling of a set of lectures to which several groups of students must attend over a preset period of time, using some resources and satisfying a certain set of constraints [22].

A lot of studies have been carried out on timetabling problem [22–26]. Nowadays, this problem is still studied due to its variety and its complexity.

A simple timetable problem could be described as follows: a school with m teachers $X_1, X_2, \ldots, X_m$ and n classes $Y_1, Y_2, \ldots, Y_n$, known definitely that teacher X_i needs to have p_{ij} lectures to class Y_j, scheduling a set of lectures with as less hours as possible.

The DNA computing model can be described as follows:

Step 1. Encoding single-strand DNA chain to represent the class.
Step 2. Designing a lab-on-a-chip to be a logical calculator.
Step 3. Computing by adding single-strand DNA chain to the chip, if the value of logical calculation is true (i.e. satisfying the hard and soft constraints), keep the value, or delete the value.
Step 4. Repeating the DNA computing process mentioned above until all known lectures p_{ij} are obtained. The needed number of cycles would be the less class hours.

A simple example is given to describe the model in details: three teachers X_1, X_2 and X_3, four classes Y_1, Y_2, Y_3 and Y_4, the required teaching array $P = [p_{ij}]$ is

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

The hard constraints are

C_0 – A teacher could only give one lecture at a time.
C_1 – A room could only host one lecture at a time.
C_2 – A class could only attend one lecture at a time.
C_3 – Room capacities that Y_1 and Y_4 could attend lecture only in room a are respected.
C_4 – Teacher X_3 having a lecture to class Y_2 do so must on the second class hour.

The soft constraints are

S_0 – Enough rooms could be used for lectures.
S_1 – Class attending lecture should obey certain order that Y_j preceded Y_{j+1}.

A DNA computing model was proposed according to this special timetabling problem:

Step 1. Single-stranded DNA chains of Y_1, Y_2, Y_3 and Y_4 representing four classes were encoded and constructed as well as their complementary chains of $\overline{Y}_1$, $\overline{Y}_2$, $\overline{Y}_3$ and $\overline{Y}_4$. The designed DNA chains, not the complementary chains, were fixed on surface as probes, in which each designed single-stranded DNA chain included two sub-regions (Y_j or $\overline{Y}_j$ and E or $\overline{E}$), representing the certain class (Y_j) and its complementary strand ($\overline{Y}_j$) as shown in Fig. 1(a) and (b), respectively. A restriction enzyme site of $E/\overline{E}$ was designed at the end of double chains of $Y_j/\overline{Y}_j$, and the 5′ end of the $\overline{Y}_j$ chain was labeled with fluorescence which would be a positive reaction when activated by laser as shown in Fig. 1(c). According to the Watson–Crick principle, the complementary chains would form duplex that could be cut away from the surface in the restriction enzyme site if restriction enzyme reaction was operated as shown in Fig. 1(d). Therefore, the biological characteristic of DNA strands designed in this step would satisfy the hard constraints C_0, C_1 and C_2.

Step 2. The lab-on-a-chip was designed and it included three pools of hybridization representing three teachers X_1, X_2 and X_3. According to the required teaching array, single-stranded DNA chains were sited on the surface of the pools in which Y_1, Y_3 and Y_4 were sited on pool of X_1; Y_1 and Y_4 were on pool of X_2; and Y_2, Y_3 and Y_4 were on pool of X_3 as shown in Fig. 2(a).

Adding $\overline{Y}_1$ to the chip while closing all channels except those to X_1. $\overline{Y}_1/Y_1$ would form duplex region. After the pool X_1 was washed by buffer in a certain experimental condition, the fluorescence labeled on $\overline{Y}_1$ was a positive reaction when activated by laser as shown in Fig. 2(b). So the value of X_1Y_1 was true. Then close all channels of X_1 and go on to the next operation by adding $\overline{Y}_2$. $\overline{Y}_2$ would go to X_2 directly since the channels to X_1 had been closed. The result is illustrated in Fig. 2(c), in which X_2 is a negative reaction and X_3 is a positive reaction, i.e. X_3 Y_2 has true value. Repeating the operation by adding $\overline{Y}_3$ to the chip, there is a negative reaction in pool of X_2, which is a positive reaction by adding $\overline{Y}_4$, as shown in Fig. 2(d). In this step, the DNA computing model would satisfy the hard constraints C_0, C_1, C_2, C_3 and C_4 as well as the soft constraints S_0 and S_1 by controlling the adding order of

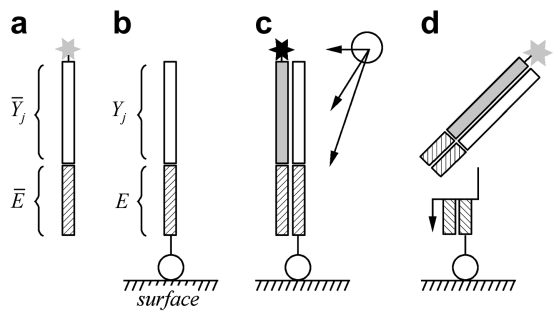

Fig. 1. The designed DNA strands on surface.

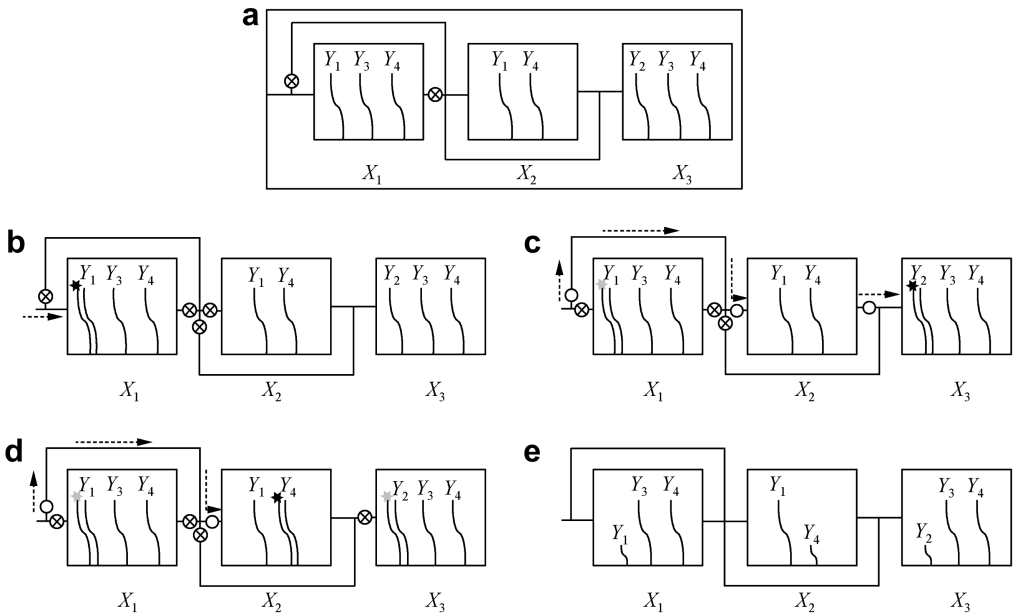

Fig. 2. The designed lab-on-a-chip and its computing processes.

DNA strands and closing or opening the channels of hybridization pools.

Step 3. After the first class hour or one cycle, all $\overline{Y}_1$, $\overline{Y}_2$, $\overline{Y}_3$ and $\overline{Y}_4$ were operated, the arrangement of lecture was ascertained, i.e. X_1Y_1, X_3Y_2 and X_2Y_4. The restriction enzyme was added to the chip. All these duplex DNA chains would be cut in restriction enzyme site and be washed away in order for next operation.

Step 4. Repeating steps 2 and 3 till all lectures were achieved. The arrangement of the second and third class hour were X_2Y_1, X_1Y_3, X_3Y_4 and X_3Y_3, X_1Y_4, respectively, after two cycles of operation.

Step 5. The needed least class hour was three to the given timetabling problem.

The main operation can be described as follows:

```
m = 3;
n = 4;
For (i = 1; i <=n; i++){
   input Y_i;
   for (j = 1; j <=m; j++){
      if there exist ( Y_i in X_j which
makes (Y_i/Y)_i(= 1)){
      close X_j;
         break;
            }
         }
      }
```

For this given example, there was not only one arrangement order of course. In fact, the arranged order of X_i in the chip or the adding order of Y_j would result in different assembled lectures of each class hour, but it would not affect the less class hours.

3. A simulation of biological experiment

A regular biological experiment was performed to simulate the computing procedures of lab-on-a-chip to verify the relative biological operation of the proposed model. In this experiment, the positively charged nylon membranes were used to substitute the surface of lab-on-a-chip, and DIG labeled and NBT/BCIP colored technique substitute for fluorescence labeled and laser activated technique. The reason for using regular biological technique to substitute for the advanced biological technique was that the essential principle of these biological operations was the same.

3.1. Preparation of biological experiment

Synthesizing DNA strands eight kinds of single-stranded DNA encoded to this model were synthesized as follows:

Y_1: BamHI (ggatcc)-gctattcgagcttaaagcta;
$\overline{Y}_1$: BamHI (cctagg)-cgataagctcgaatttcgat-DIG-11-ddUTP;
Y_2: BamHI (ggatcc)-tgaggcagtatgagcacgag;
$\overline{Y}_2$: BamHI (cctagg)-actccgtcatactcgtgctc-DIG-11-ddUTP;
Y_3: BamHI (ggatcc)-ggtatggtcgcacgaagcaa;
$\overline{Y}_3$: BamHI (cctagg)-ccataccagcgtgcttcgtt-DIG-11-ddUTP;
Y_4: BamHI (ggatcc)-cataggagcatatcgtaccg;
$\overline{Y}_4$: BamHI (cctagg)-gtatcctcgtatagcatggc-DIG-11-ddUTP.

Labeling DNA strands DNA strands of $\overline{Y}_1, \overline{Y}_2, \overline{Y}_3$ and $\overline{Y}_4$ were labeled DIG-11-ddUTP on 3′ terminal with termi-

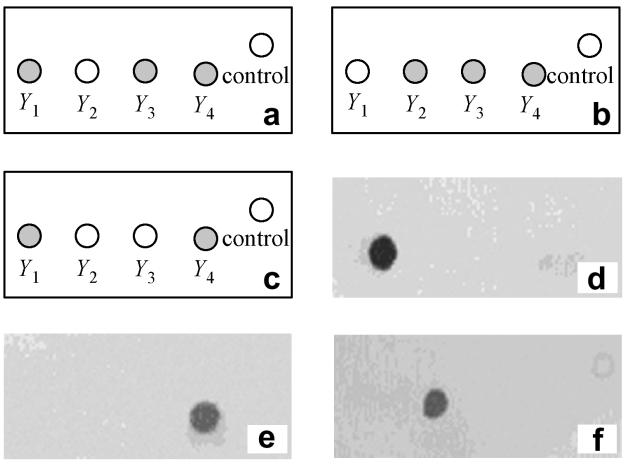

Fig. 3. The hybridize matrix and hybridize results on positively charged nylon membranes.

nal deoxynucleotidyl transferase (TdT) by using the DIG Oligonucleotide 3′-End Labeling Kit, according to the manufacturer's instruction, Roche Co.

Arranging hybridization matrix according to the model, three positively charged nylon membranes were prepared to perform molecule biological Southern Blot operation, in which the matrix of hybridization is shown in Fig. 3. For example, the matrix of Fig. 3(a) as addressed with DNA strands which indicates that the teacher X_1 has three classes Y_1, Y_3 and Y_4.

3.2. DNA Southern Blot operation

After taking the Southern Blot operation by adding $\overline{Y}_1$ to X_1, the membranes were colored with NBT/BCIP and washed with buffer (these detection operations were performed by using DIG High Prime DNA Labeling and Detection Starter Kit according to the manufacturer's instruction, Roche Co.).

3.3. Results

The experimental result of $\overline{Y}_1$ to X_1 hybridization has positive reaction on address of Y_1 as shown in Fig. 3(d), and the value of X_1Y_1 was true. After $\overline{Y}_1, \overline{Y}_2, \overline{Y}_3$ and $\overline{Y}_4$ were performed, the experimental results of X_3 and X_2, as shown in Fig. 3(e) and Fig. 3(f), respectively, are that Y_2 and Y_4 have a positive reaction. Therefore, the arranged lectures of the first class hour X_1Y_1, X_3Y_2 and X_2Y_4 are ascertained, and this simulation is consistent with the proposed DNA computing model.

4. Discussion and conclusion

In this study, only a lab-on-a-chip model was given in theory without a real finished product. Although a simulated biological experiment was carried out, many techniques are needed to manufacture a chip which we had not described in detail here. Even with a lab-on-a-chip, it was problematic to hold the large amounts of DNA molecules required to support massively parallel computing and provide vast memory capacity. However, we believe that developing a compact hybrid DNA computer that uses microchips by taking advantage of both lab-on-a-chip and DNA computing technologies is worthy of a further research.

For a real timetabling problem, such as the general university course timetabling problem (UCTP) known to be NP-hard, it usually must satisfy a large and diverse array of supplementary constraints which are difficult to describe in a simple example. The presented model for solving timetabling problem in this study mainly helps us learn more about the nature of computation, and more research is required to develop a better DNA computing model, which is capable of solving a wide range of complex problem.

Acknowledgments

This work was supported by National Natural Science Foundation of China (Grant Nos. 30370356, 30740036), the open funds of Liaoning Key Lab of Intelligent Information Processing, Dalian University (Grant No. 2005-1), Zhejiang Provincial Natural Science Foundation of China (Grant No. Y106654) and the Basic Research Foundation of Beijing Institute of Technology (Grant No. 20070642002). The authors sincerely thank Prof. Guangzhao Cui for his insightful comments.

References

[1] Adleman LM. Molecular computation of solutions to combinatorial problems. Science 1994;266:1021–3.
[2] Calude CS, Paun G. Computing with cells and atoms. London: Taylor and Francis Publisher; 2000.
[3] Paun G, Rozenberg G, Salomaa A. DNA computing: new computing paradigma. Berlin: Springer Verlag; 1998.
[4] Rozenberg G, Salomaa A. Handbook of formal languages. Berlin: Springer Verlag; 1997, p. 1–3.
[5] Lipton RJ. DNA solution of hard computation problem. Science 1995;268:542–5.
[6] Ouyang Q, Kaplan PD, Liu SM, et al. DNA solution of the maximal clique problem. Science 1997;278:446–9.
[7] Ying ZX, Zhang FY, Xu J. A Chinese Postman problem based on DNA computing. J Chem Inf Comput Sci 2002;42:222–4.
[8] Cai W, Anne E, Robert M, et al. The power of surface-based DNA computation. In: Proceedings of the first annual international conference on computational molecular biology; 1997. p. 67–74.
[9] Frutos G, Smith M, Corn M, et al. Enzymatic ligation reactions of DNA "word" on surface for DNA computing. J Am Chem Soc 1998;120:10277–82.
[10] Smith M, Corn M, Anne E, et al. A surface-based approach to DNA computation. Comput Boil 1998;5:255–67.
[11] Liu QH, Wang LM, Fruto AG, et al. DNA computing on surfaces. Nature 2000;403:175–9.
[12] David HW, Catherine L, Taylor C, et al. Universal biochip readout of directed hamiltonian path problems. In: DNA 8, LNCS 2568. Berlin Heidelberg: Springer-Verlag; 2003. p. 168–81.
[13] Yoichi T, Akihiro H. Shortening the computational time of the fluorescent DNA computing. In: DNA 8, LNCS 2568. Berlin Heidelberg: Springer-Verlag; 2003. p. 85–94.

[14] Zhang FY, Yin ZX, Xu J. Application of DNA chip on 0–1 planning problem. Prog Biochem Biophys 2003;30(3):412–5.
[15] Braich R, Chelyapov N, Johnson C, et al. Solution of a 20-variable 3-SAT problem on a DNA computer. Science 2002:499–502.
[16] Cheng J, Edward LS, Wu L, et al. Preparation and hybridization analysis of DNA/RNA from *E. coli* on microfabricated bioelectronic chips. Nat Biotechnol 1998;16:541–6.
[17] Quake SR, Scherer A. From micro- to nanofabrication with soft materials. Science 2000;290:1536–40.
[18] Han J, Craighead HG. Separation of long DNA molecules in a microfabricated entropic trap array. Science 2000;288:1026–9.
[19] Thorsen T, Maerkl SJ, Quake SR. Microfluidic large-scale integration. Science 2002;298:580–4.
[20] Stavis SM. Detection and identification of nucleic acid engineered fluorescent labels in submicrometre fluidic channels. Nanotechnology 2005;16:314–23.
[21] Craighead H. Future lab-on-a-chip technologies for interrogating individual molecules. Nature 2006;442:387–93.
[22] Werra D. An introduction to timetabling. Eur J Oper Res 1985;19:151–62.
[23] Abramson D. Constructing school timetables using simulated annealing: Sequential and parallel algorithms. Manag Sci 1991;37:98–113.
[24] Colorni A, Dorigo M, Maniezzo V. Genetic algorithms and highly constrained problems: the time-table case. In: Parallel problem solving from nature – proceedings of 1st workshop, PPSN 1, LNCS 496. Berlin: Springer-Verlag; 1991. p. 55–9.
[25] Corne PR, Fang H-L. Evolutionary timetabling: practice, prospects and work in progress. In: UK Planning and Scheduling SIG Workshop, 1994.
[26] Gaspero L, Schaerf A. Tabu search techniques for examination timetabling. In: Proceedings of the 3rd international conference on practice and theory of automated timetabling (PATAT 2000), LNCS 2079. Berlin: Springer-Verlag; 2001. p. 104–17.