DNA computing using a difference of melting temperature among DNA fragments

이지연, 신수용*, 장병탁*, 박태현
서울대학교 응용화학부, 컴퓨터공학부*
전화 (02) 880-8020, FAX (02) 888-7295

ABSTRACT

We propose new encoding method for numerical data in DNA using temperature gradient. To represent numerical values in DNA sequences, we introduce melting temperature. Since DNA strands representing smaller values have a lower Tm, they tend to denature with ease and also easily amplified by denaturation temperature gradient PCR. We also implement a local search molecular algorithm using temperature gradient, which is contrasted to conventional exhaustive search molecular algorithms. The proposed methods are verified by solving an instance of the travelling salesman problem. We could effectively amplify the correct solutions and the use of temperature gradient made the detection of solutions easier.

INTRODUCTION

DNA computing is an unfamiliar term to biologist. DNA computing is one of the newly proposed computing method with biomolecules to overcome the disadvantages of conventional computation based on serial and irreversible properties. DNA, composed of four bases has a potentially gigantic memory capacity. We can generate duplex DNA strands by simple hybridization and manipulate them using specific enzymes. Since biochemical reactions in cells are massively parallel, DNA computing has an inborn computational power itself.

We proposed an encoding method for numerical data in DNA using temperature gradient and we also designed the modified biochemical tools to implement our scheme. After the Adleman's pioneering work[1], DNA computing has been applied to various fields of research. Most of these applications do not consider the problem of representing numerical data in DNA molecules. However, many real world applications involve graph problems which have weighted edges. To evaluate our method and implementation tools, we selected a traveling salesman problem (TSP) as a benchmark problem. The objective of the TSP is to find the cheapest way of visiting all the cities and returning to the starting point this when a number of cities to visit and the traveling cost between each pair of cities are given. TSPs are interesting in that their solution requires to represent the path weights. We are solving the TSP with 7 nodes and 23 edges as shown in Fig. 1. Each oligomer sequence was designed using a sequence generator, NACST, which was developed to generate orthogonal oligomer sequences fitted to solve TSP[2]. Representation of weight is a key factor to solve the TSP with DNA molecules. We represented the weights using melting temperature gradient, and implemented a local search molecular algorithm with the temperature gradient method. To exactly determine the Tm of oligomers, we used both conventional GC content method and nearest-neighbor method[3]. DNA sequence with lower weight has a lower Tm and is easily denatured. Therefore, the Tm of the DNA strands that contain the smaller sum

of weight is lower among the sequences having the same length and consequently those are more amplified by denaturation temperature gradient PCR. This strategy effectively amplified the correct solution and made the detection easier.

MATERIALS AND METHODS

Oligomer synthesis All 7 vertexes and 5 weights are designed in 20 mer ssDNA. And 23 edges in 40 mer ssDNAs are designed according to their vertexes and weight sequences. All oligomers are 5'-phosphorylated. Other 5'-biotinylated vertexes were prepared for affinity-separation. All oligomers were synthesized at Bioneer Co.

Hybridization and ligation Oligomer mixture was heated to 95℃ and cooled to 20℃ by 1℃ per 1 minute. Mix 5 μl of hybridized mixtures and 350 units of T4 DNA ligase (TaKaRa), ligase buffer (66 mM Tris-HCl, pH 7.6, 6.6 mM $MgCl_2$, 10 mM DTT, 0.1 mM ATP), and add H_2O to 10 μl. We incubated the reaction mixture for 16 hours at 16℃.

PCR amplification and gel electrophoresis All PCR amplifications were performed on a PTC-200 DNA Engine (MJ Research). For normal PCR, 1 μM of primer and AccuPower PCR PreMix (Bioneer) were dissolved to distilled water to a total volume of 20 μl. And PCR was processed for 34 cycles at 95℃, (Tm-5)℃, and 72℃ for each 30 seconds. Initial denaturation and prolonged polymerization was executed for 5 minutes. All gel electrophoresis were performed with 2% Agarose-1000 (GibcoBRL) in 0.5X TBE buffer and the gel was ethidium bromide stained. 50 bp DNA Ladder (GibcoBRL) was used as a marker.

Affinity separation We roughly followed the affinity separation protocol in the Adleman's experiments. We obtained the ssDNA by PCR priming with the 5'-biotinylated analog. The amplified product was annealed to streptavidin paramagnetic particles (Promega) and the particles were washed three times. Then the particles were heated to 80℃ in 100 ml of ddH_2O for 5 minutes to denature the bound dsDNA. The aqueous phase was retained. For affinity purification, 1 nmol of 5'-biotinylated vertex strands was annealed to particles as above and washed three times. Particles were washed four times to remove unbound ssDNA and then heated to release ssDNA bound to the complementary vertex 1. The aqueous phase with ssDNA was retained. This process was then repeated with each vertexes. The washing solution was 0.5X SSC.

Denaturation temperature gradient PCR For denaturation temperature gradient PCR (DTG-PCR), the denaturation temperature was kept initially 70℃ and gradually decreased 1 ℃ per one cycle and kept 95℃ for the remaining 10 cycles. Other conditions are identical with normal PCR described in the PCR amplification section.

RESULTS AND DISCUSSION

Random Path Formation and Size Sieving The hybridization and ligation results are shown in Fig. 3(A). Compared with the oligomer mixture, the ligated DNA strands became elongated. But, there are few copies around the 300 bp that indicates the length of paths including 8 vertexes. So we executed the ligation reaction again and obtained an upper shifted ligation product as shown in Fig. 3(B). However, still shorter DNA strands mainly retained. Ligation is an essential step to generate random paths, so efficient ligation is needed to produce DNA strands long enough to solve large problems.

PCR with In and Out Vertex and Affinity Separation After the hybridization and ligation reaction, dsDNAs with sticky ends are generated. We cannot use the primer pair because the start and the end point are identical in TSP and the primer pair is exactly complementary to each other. So we executed PCR with only one primer that is complementary with the end vertex, 0. By this PCR, we could make blunt-ended double-stranded DNAs that end with the vertex 0. We sieved the PCR product (Fig. 3(C)) and picked out the DNA strands around 300 bp. Subsequently, we executed the second PCR with vertex 0, and we could amplify the DNAs that start with vertex 0. As shown in Fig. 3(D), the PCR product appeared in several bands in the gel. We excised and eluted the band around 300 bp and amplified with 5'-biotinylated vertex 0 as a primer. The amplified product was used in the following affinity separation.

Denaturation Temperature Gradient PCR and Latter Implementation Steps We made the Tm of the correct solution to be the lowest among the candidate solutions using the temperature gradient method. Therefore, if we decrease the denaturation temperature to a certain level, mainly DNA strands of correct solutions can be denatured at that temperature, and be amplified. As the denaturation temperature increases, other DNA strands also will be amplified. But the amount of correct solutions will be more increased cycle by cycle, and occupy the major part of the solution. By simple modification of the typical PCR, we can amplify the correct solution and detect it easily. We can show the effectiveness of DTG-PCR by the relative amplification in Fig. 3(E). Shorter DNA strands in the lane 2 are more amplified by DTG-PCR when compared with the normal PCR product in the lane 1. We purified the lower band by excision and elution form the gel, and executed the repetitive DTG-PCR, and obtained the band around 300 bp. This band contains four different DNA strands of the possible hamiltonian paths. Because these DNA strands have the identical length, 300 bp, they cannot be separated by normal gel electrophoresis. However, we can separate these strands by the temperature gradient gel electrophoresis (TGGE), because they have distinct melting behaviors. Their GC content varies from 40.67% to 44.00%, and the Tm of each DNA strands ranges from 84.80℃ to 86.62℃ by NN model, and from 96.51℃, to 97.87℃ by GC content method with 1 M salt concentration and 10 nM oligomer concentration. The separated DNA strands with the lowest Tm can be sequenced to confirm that the whole procedure is correct. If we design the starting and end vertex to contain certain restriction enzyme sites, the sequencing step will become easier. This latter implementation steps will be a future work.

CONCLUSIONS

We introduced a temperature gradient method to solve the problems with numerical data. This method can overcome the restrictions of other encoding methods and be easily implemented by a simple modification of experiments. First, we can more strongly amplify the correct solution relatively in the reaction procedure and the solution can be easily detected. This method drives the DNA pool to contain more correct solutions, rather than to search randomly for the correct solution in the DNA pool. Second, by fixing the length of weight sequences, we can easily scale up the problem size. We also showed the feasibility of our computing method by solving the travelling salesman problem with adequate biochemical

tools. The combination of the orthogonal design of DNA sequences and denaturation temperature gradient PCR provides a novel method to solve general graph problems with weighted edges. This is applicable to any Tm-involved implementation steps to amplify low melting temperature DNA strands.

REFERENCES

1. Adleman, L. M., "Molecular computation of solutions to combinatorial problems" (1994), *Science*, **266**, 1021-1024

2. Kim D, Shin S. Y., Lee I. H., and Zhang B. T., "NACST/Seq: A sequence design system with multiobjective optimization" (2002), *8th International Meeting on DNA Based Computers*, (Submitted)

3. SantaLucia, J, "A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics" (1998), *Proc. Natl. Acad. Sci.*, **95**, 1460-1465

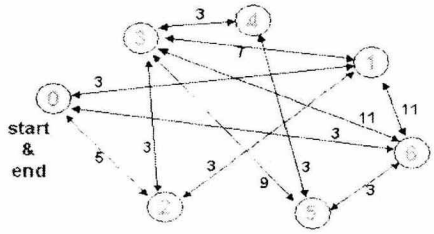

Fig. 1 Target TSP problem (each integer beside the arrow is weight)

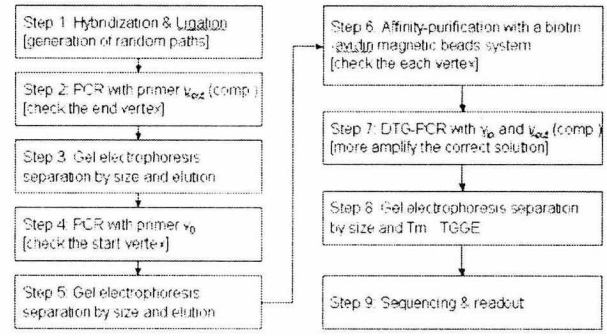

Fig. 2 Experimental strategy

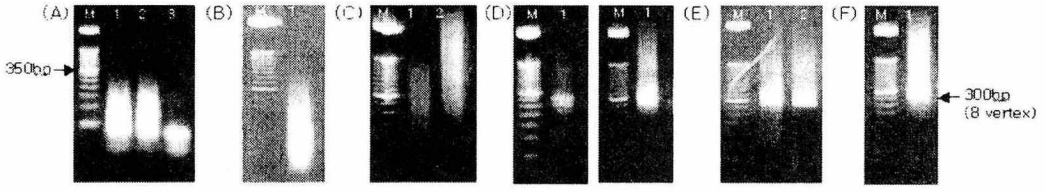

Fig. 3. Gel electrophoresis on a 2% agarose gel. The lanes contain: lane M denotes DNA size marker (50 bp ladder), (A) lanes 1,2: first ligation result, lane 3: oligomer mixture, (B) lane 1: second ligation result, (C) lanes 1,2: first PCR with V_{out} result, (D) lane 1: second PCR with V_{in} result, (E) lane 1: normal PCR result, lane 2: DTG-PCR result, (F) lane 1: the final DTG-PCR result.