J. Biochem. **142,** 605–611 (2007)
doi:10.1093/jb/mvm166

Multiple Hybridization-extension Sequencing (MHES) on Microarray

Zhiqiang Pan, Yanqiang Li, Pengfeng Xiao and Zuhong Lu*
State Key Laboratory of Bioelectronics, Southeast University, Nanjing, 210096, China

Received August 6, 2007; accepted August 17, 2007; published online September 7, 2007

Sequencing-by-synthesis (SBS) by fluorescein-labelled nucleotide incorporating into a target DNA template has been greatly concerned on microarray. The extended fluorophore-base must be required to be quenched prior to sequencing the next one. However, the low quenching efficiency has been an obstacle in length-read. Here, we present a new sequencing strategy, multiple hybridization-extension sequencing (MHES), to resolve the above problem. First, the sequencing primers hybridize to the ssDNA template immobilized on microarray. The first 3–5 bases next to the primer's end are sequenced by SBS of Cy5-dNTP. The extended primers are rapidly removed by λ DNA exonuclease. Then, the same primers hybridize to the same ssDNA templates again. The sequenced bases are polished by natural dNTP. The other 3–5 bases next to the polished primer's end are sequenced. According to this principle, the unknown sequences of a target DNA could be sequenced after primers' hybridization-extension multiple times. Although the fluorescein-labelled nucleotides are also needed, it is unnecessary to quench the fluorophore-bases in the process of sequencing. It has been successfully demonstrated that 10 bp fragment from synthetic template and 10 bp fragment from DTBNP1 gene were accurately sequenced. The new method has a great potential in read-length and high-throughput sequencing on microarray.

Key words: DTBNP1 gene, microarray, multiple hybridization-extension, sequencing.

INTRODUCTION

DNA sequencing has emerged as a very important tool in finding genetic information and has benefited many number of different scientific researches such as molecular biology, genetics, clinical diagnostics and so on (*1*). The success of the Human Genome Project (HGP) largely resulted from the early development of the Sanger sequencing method, and the considerable amount of information has become available from this sequencing method (*2*). Actually, it is clear that the sequence information in future demand will keep increasing. Although the cost of the Sanger method has become low, it is hard to meet the demand for the increasing individual sequences. A barrier to increase the sequencing needs still remains cost and time. To search for alternative faster and cheaper sequencing technology (*3*), a variety of new sequencing technologies have been investigated and developed by many research groups around the world (*4*). The present methods may be classed into three different principles of sequencing. One is sequencing-by-hybridization, in which the sequences are sequenced by hybridizing DNA to a high-density oligonucleotide microarray (*5*). Another is sequencing-by-decomposition, the sequences of target are sequenced by chemically degrading DNA (*6*) or analysing DNA pieces broken by MS (*7–9*). The third method is sequencing-by-synthesis such as pyrosequencing (*10*, *11*), polymerase colonies (*12*) and sequencing of single molecules dye (*13*). These methods have the potential to speed the rate of DNA sequencing. Especially for SBS, it has been significantly concerned due to being great flexibility and feasibility.

SBS involves to immediately read the sequence of a DNA template after each nucleotide incorporating into a growing primer-ssDNA, if each nucleotide incorporated can be identified, then the sequence can be read (*14*). In the case of SBS of fluorescein-labelled nucleotide incorporating into the target DNA, it has the high throughput potential to be applied on microarray (*15–17*). In most cases, fluorescein-labelled nucleotides are all required to be applied, and fluorophore-bases must be quenched after detection in order to extend the next bases well. However, the current quenching reagent or means could not completely remove the fluorophore-bases, and could only partly remove or destroy DNA template or/and primer. Up to the present, there has not been an ideal quenching reagent or means suiting for this task.

In this study, we describe a new sequencing strategy to resolve this problem, it is the multiple hybridization-extension sequencing based on SBS. The sequences of the immobilized DNA target on microarray are read through multiple primer-hybridization and multiple primer-extension. First, the sequencing primers hybridize to the ssDNA template immobilized on microarray. The first 3–5 bases next to the primer's end are sequenced by SBS. The extended primers are rapidly removed by λ DNA exonuclease. Secondly, the same primers hybridize

*To whom correspondence should be addressed. State Key Laboratory of Bioelectronics, Southeast University, Nanjing, 210096, China. Tel: +86-25-83793310, Fax: +86-25-83793779, E-mail: zhlu@seu.edu.cn

to the same ssDNA templates again on microarray. The sequenced bases are polished by natural unlabelled-nucleotides. The other 3–5 bases next to the polished primer's end are sequenced by SBS. According to the principle, the unknown sequences of a target DNA could be read well after primer being hybridized-polished-extended-removed multiple times. Although the fluorescein-labelled nucleotides are also needed, it is unnecessary to quench the fluorophore-bases in the process of sequencing.

As a principle experiment, we have successfully read 10 bp in length from synthetic DNA template. It has been successfully demonstrated that 10 bp in DTBNP1 gene with unknown sequences are accurately sequenced by the new method. The new method has a great potential in read-length and high-throughput sequencing on microarray.

MATERIALS AND METHODS

Preparation of Aldehyde-coated Glass Slides—The glass slides were cleaned for 1 h in a solution consisting of 1/3 hydrogen peroxide (30%) and 2/3 sulphuric acid (18 M) (*18*), rinsed three times in deionized distilled water, left for 10 min in boiling deionized distilled water and dried under an argon flow. The cleaned glass slides were silanized with 2% of 3-aminopropyltriethoxysilane (Sigma) dissolved in 95% acetone for 2 min, washed twice with acetone and baked for 45 min at 75°C. The silanized slides were activated with 5% glutaraldehyde in 0.11 M PBS (pH 7.0) for 2 h, washed thoroughly with distilled water, and dried. The aldehyde-coated slides were stored at 4°C.

Synthesis of Oligonucleotide—Oligonucleotides were synthesized and purified by Invitrogene and summarized in Table 1. PT1 was specially designed to fabricate a DNA microarray for investigating the extension conditions. PT2 was used to prepare the oligonucleotide microarray for demonstrating the principle of sequencing. PT31, PT32 and PT33 had -g, -gg and -ggg difference, respectively in their sequences. PF and PR were designed to amplify the 189 bp DTNBP1 gene fragment in length (Accession No. NC_006473 from 144353 to 144541). PS, PS2 and PS3 were the sequencing primers.

PCR Amplification—DTNBP1 DNA was PCR amplified with primer PF and PR. The profile was consisted of an initial melting step of 2 min at 94°C, followed by 35 cycles of 30 s at 94°C, 30 s at 60°C and 45 s at 72°C; and a final elongation step of 7 min at 72°C. PCR products were purified using a QIA-quick column (Qiagen). The eluted PCR products were dried in vacuum.

Preparation of ssDNA Microarray—Oligonucleotides (PT1 and PT2) and the purified PCR products were respectively diluted a final concentration of 20 μM in sodium carbonate buffer (0.1 M, pH 9.0), and transferred to 384-well-plates for arraying. They spotted on the silanizationed glass slides (*19*) using a PixSys5500 microarrayer (Cartesian Technology). Then, the slides were incubated in a humid chamber at room temperature for 4 h, and at 37°C for 2 h. For PCR microarray, the dsDNA on microarray was denatured by 0.2 M NaOH at room temperature for 5 min. The slides were rinsed once by solution I (0.2% Trion X-100, 2% SSC-0.1% SDS), and rinsed once by deionized water for hybridization.

Table 1. **The synthesized oligonucleotides.**

Oligonucleotide	sequences
PT1	5′-NH_2-$(T)_{12}$-GATCCTTGTCACTGCTACATGTACAGATAG AGGCAGCTCTGACCATGCTTGCATGGTCAGAGCT GCC-3′
PT2	5′-$\overline{NH_2}$-$(T)_9$GTAGTCGGGGGTTTTAAAGGCCTACGGTCGT GACTGGGAAAACCCTGGCGTGCGTGATG-3′
PT31	5′-NH_2-$(T)_{12}$-GATCCTTgTCACTGCTACATGTACAGATAGCT
PT32	5′-NH_2-$(T)_{12}$-GATCCTTggTCACTGCTACATGTACAGATAGCT
PT33	5′-NH_2-$(T)_{12}$-GATCCTTgggTCACTGCTACATGTACAGATAGCT
PF	5′-CATGCTCCACTCCGGTAA-3′
PR	5′-NH_2-CGCCCCTAAATGCCTGTC-3′
PS	5′-PO_4^{3-}-GGCGTTTGCAGGGAGGTGCG-3′
PS2	5′-PO_4^{3-}-CATCACGCACGCCAGGGTTTTCCC-3′
PS3	5′-AGCTATCTGTACATGTAGCAGTGA-3′

PT1, PT2, PT31, PT32 and PT33, directly used to fabricate the ssDNA microarray (Two underline regions can form self-hairpin). PF and PR, a pair of primers used to amplify DTNBP1 gene (Accession No.NC_006473 from 144353 to 144541). PS, the sequencing primer of the target DTNBP1. PS2, the sequencing primer of PT2. PS3, the sequencing primer of PT31, PT32 and PT33.

Sequencing—Primer hybridizing Sequencing Primers (PS, PS2 and PS3) were diluted in hybridization solution [3 : 1 dilution (v/v); Telechem]. Hybridization was conducted in a humid glass chamber sealed with plastic film at 80°C for 5 min, after cooled to room temperature and remained for 5 min. Finally, the slides were rinsed with solution I for 2 min, rinsed once with Taq DNA polymerase buffer [solution II, 20 mM Tris–HCl (pH 8.4), 1.5 mM $MgCl_2$, 20 mM KCl, 10 mM $(NH4)_2SO_4$] for 1 min.

Polishing—Before the next sequencing cycle, the determined bases were polished by natural dNTP using Taq DNA polymerase at 65°C for 5 min. The reaction system contained 1 × solution and 10 μM each dNTP, 0.25 U/μl Taq polymerase. After reaction, the slide was rinsed once with solution I at room temperature for 2 min, rinsed once with deionized water for 1 min, followed by rinsing with Klenow polymerase buffer [solution III, 10 mM Tris–HCl (pH 7.5), 5 mM $MgCl_2$, 7.5 mM DTT] for 1 min.

Extending Primer—Extension was performed in 1 × solution III, the corresponding ratio of Cy5-dNTP to nature dNTP, 0.25 U/ul Klenow polymerase (exo^-, New England Biolabs) at 37°C for 3 min. The concentration ratio of Cy5-dCTP to dCTP was 1 : 100 (0.1 μM : 10 μM), Cy5-dUTP (dTTP analog) to dTTP 1 : 100 (0.1 uM: 10 μM), Cy5-dATP to dATP 1 : 300 (0.1 μM : 30 μM), Cy5-dGTP to dGTP 1 : 500 (0.1 μM : 50 μM) (Cy5-dATP, Cy5-dUTP, Cy5-dGTP, Amersham Pharmacia Biosciences and Cy5-dCTP, Perkin–Elmer). After incubation, microarray was rinsed twice with solution I at room temperature for 1 min, followed by rinsing once with deionized water for 1 min. And the slide was blowed with a dry N_2 stream.

The nucleotides A, G, C and T, was added orderly in the reaction.

Removing Primer—In order to obtain an efficient method of removing the primer, three different methods were compared. The extended complex of target PT2 and primer PS2 was used to investigate it. (i) NaOH denaturation: microarray was rinsed with 0.2 M NaOH at room temperature for 5 min, rinsed once with deionized water for 1 min and rinsed once with solution I for 1 min, and rinsed once with solution III for 1 min, (ii) λ DNA exonuclease digestion: digestion was performed in 1 × λ DNA exonuclease buffer, 2 U/ul λ DNA exonuclease at 37°C for 1–60 min, rinsed once with solution I for 1 min and rinsed once with solution III for 1 min, (iii) T7 DNA exonuclease digestion: digestion was performed in reaction containing 1 × T7 DNA exonuclease buffer, 2 U/μl T7 DNA exonuclease at 25°C for 1 h, after digestion, the slide was rinsed once with solution I for 1 min, rinsed once with solution III for 1 min.

Imaging and Analysis—The microarrays images were captured by a scanner with an instrument setting of 70% laser power and 950 PMT except for supplying an explanatory in experiment (ScanArray Lite, Packard BioScience Company, USA), and were analysed with Genepix Pro 3.0 software.

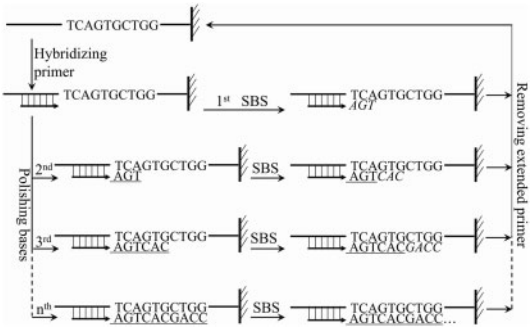

Fig. 1. **Principle of multiple hybridization-extension sequencing**. In the first cycle, the 'AGT' were sequenced by SBS using Cy5-dNTP. The extended primer was removed. In the second cycle, the sequenced 'AGT' were polished by natural dNTP. The 'CAC' were sequenced by SBS using Cy5-dNTP. The extended primer was removed again. In the third cycle, the sequenced 'AGTCAC' were polished by natural dNTP. The 'GACC' were sequenced by SBS using Cy5-dNTP. According to the same principle, the longer sequences of a target template were read after multiple hybridization and extension. 1st, the first sequencing cycle. 2nd, the second sequencing cycle. 3rd, the third sequencing cycle. *n*th, the *n*-th sequencing cycle. The underlined bases, natural dNTP. The italic bases, Cy5-dNTP. SBS, sequencing-by-synthesis.

RESULTS

Sequencing Principle—The principle of sequencing by MHES is shown in Fig. 1. As illustrated, the strategy could determine the unknown sequences in a given gene. First, the immobilized ssDNA templates were hybridized with the sequencing primers on microarray. The first 3–5 bases next to the primer's end are sequenced by SBS. The extended primers are removed through the digestion of λ DNA exonuclease. Secondly, the ssDNA templates were hybridized with the same primers again. The sequenced bases are polished by natural dNTP. The other 3–5 bases next to the polished primer's end are sequenced by SBS. According to the same principle, the unknown sequences of a target DNA could be read after primer being hybridized-polished-extended-removed multiple times.

Proof-of-Principle Experiment—To demonstrate the feasibility of the sequencing of multiple hybridization-extension on microarray, the synthetic oligonucleotide PT2 modified with 5′-amino were used to fabricate the oligonucleotide microarray on the slide. Figure 2 shows the sequencing results for PT2 template. The average relative fluorescent intensity (ARFI) of spots showed the linearly increase with the increasing of the extended bases in each cycle. Three bases of AGT (5′–3′) were sequenced in the first cycle. The ARFI of the added nucleotides were 1597 (A), 3128 (A + G), 3089 (A + G + C) and 4300 (A + G + C + T) respectively after subtracting from background signal (35) (Fig. 2A). Given 1597 equaled to X, 3128 to Y, 3089 to Z and 4300 to L, then $Y-X=1531$, $Z-Y=-39$, $L-Z=1211$. $(Y-X)/X=0.96$, $0.5<0.96<1$, so the result indicate that the signal was positive, resulting in one nucleotide incorporating into the primer. And $Z-Y=-39$, which indicated negative signal, and no extension. Similarly, $(L-Z)/X=0.76$, which was located between $0.5X$ and $1X$, therefore, only one nucleotide incorporated into the template PT2 (the analysis results of each spot were still complied with this regular, data not shown).

In the second cycle, the ARFI of spots was orderly 40 (A), 35 (A + G), 1419 (A + G + C), 1402 (A + G + C + T), 2720 (A + G + C + T + A), 2701 (A + G + C + T + A + G), 3985 (A + G + C + T + A + G + C) and 3955 (A + G + C + T + A + G + C + T), respectively (Fig. 2A). According to the same principle mentioned above, three bases of CAC (5′–3′) could be easily determined.

In the third cycle, the ARFI of spots were 38 (A), 1078 (A + G), 1081(A + G + C), 1085 (A + G + C + T), 1879 (A + G + C + T + A), 1883 (A + G + C + T + A + G), 3712 (A + G + C + T + A + G + C) and 3718 (A + G + C + T + A + G + C + T), respectively. 1879−1078 = 801, 801/1078 = 0.74 indicating one nucleotide incorporation. 3712−1879 = 1833, 1833/1078 = 1.70, the value was located between 1.5 and 2, indicating two repeat bases were extended at the same time on target template. Therefore, the extended bases were determined to be 5′-GACC-3′ in this cycle.

In conclusion, the sequences of nucleotides read along the primer were 5′-AGTCACGACC-3′. Their complementary sequences were 5′-GGTCGTGACT-3′ on the target template. The sequencing results were completely in accordance with the synthetic template PT2.

Effect of the Factors on the Accuracy of Sequencing—The fidelity of polymerases is one key factor for correct incorporation of complementary versus non-complementary bases. Therefore, it is necessary to investigate the effect of Klenow polymerase on misincorporation of different Cy5-dNTP in MHES method. The synthetic template PT1 with self-hairpin structure was applied in

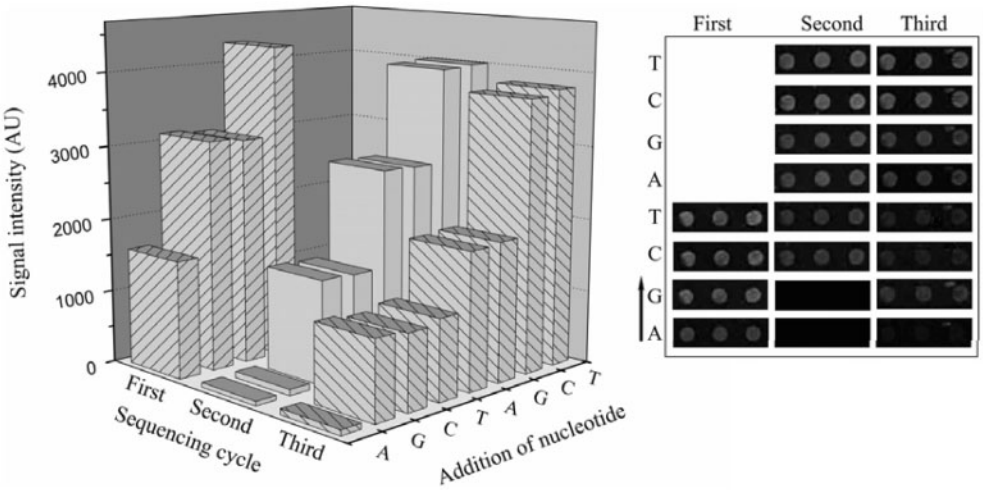

Fig. 2. **Proof-of-principle experiment by sequencing PT2 template on microarray.** (A) The ARFI analysis of scanning image. (B) Scanning image of spots. The addition of nucleotides was 'AGCT' in each cycle.

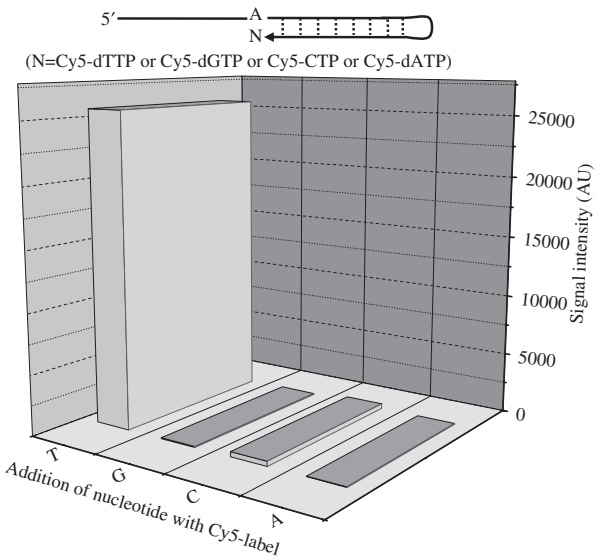

Fig. 3. **Misincorporation of four different Cy5-dNTP by using Klenow polymerase (exo$^-$).** The self-primering PT1 was used to investigate it. Only 'T' was the complementary to the 'A' in hairpin PT1. The other three nucleotides were non-complementary.

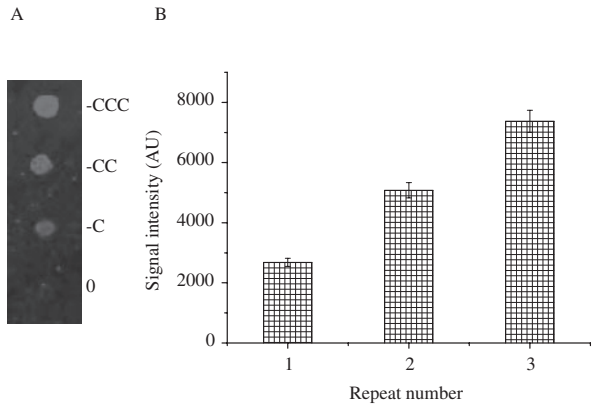

Fig. 4. **Quantification of the ARFI ratio of 1-, 2- and 3-base repeats using the 1:100 Cy5-dCTP/dCTP.** (A) Scanning image of spots. -C, 1C base extended on PT31. -CC, 2C extended on PT32. -CCC, 3C extended on PT33. 0, spot without template. (B) The ARFI analysis of these spots. 1, -C. 2, -CC. 3, -CCC. PT31, PT32 and PT33 were designed to be -g, -gg and -ggg difference in their sequences. Their immobilization concentration was 35 μM.

the experiment. Extension was performed in 1 × solution III, 0.1 μM one of Cy5-dNTP, 0.25 U/μl Klenow polymerase (exo$^-$). The Cy5-dNTP was incorporated along the hairpin-primer using Klenow polymerase. We can see from Fig. 3 that the incorporated nucleotide should be the Cy5-dTTP. The ARFI of spots was 25,700 after the Cy5-dTTP incorporation. When the other three nucleotides non-complementary base were incorporated, the ARFI of spots from Cy5-dGTP was 53 (misincorporation ratio 0.2%), from Cy5-dCTP 411 (1.6%) and from Cy5-dATP 51 (0.2%). Here, the ratio was ratio of the ARFI value from misincorporation to the ARFI value from incorporation value. Obviously, misincorporation rate of Cy5-dGTP and of Cy5-dATP were relative lower. The misincorporation rate of Cy5-dCTP was comparatively high.

For repeat bases in target DNA, we quantified the ratio of their ARFI. PT31, PT32 and PT33 were designed to investigate it. They were only -g, -gg and -ggg difference in their sequences (Table 1). They were respectively diluted to 35 μM. After immobilization, they were hybridized with PS3. The primer extension was performed by using 1: 100 concentration ratio of Cy5-dCTP to dCTP. Figure 4A shows fluorescence image of spots. The ARFI of 1-, 2- and 3-base repeats were 2680, 5078 and 7370, respectively (Fig. 4B). So, their ratios were 1: 1.89: 2.75, which was located in

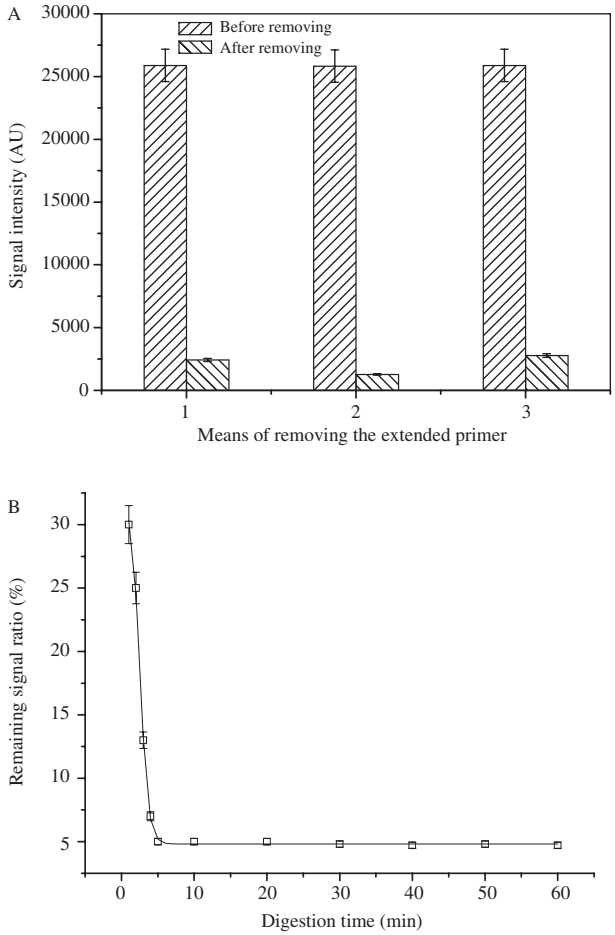

Fig. 5. **Three different means of removing the extended primer.** (A) The ARFI contrast before and after removing the primer. 1, NaOH denaturation. 2, λ DNA exonuclease digestion. 3, T7 DNA exonuclease digestion. (B) The remaining ratio of fluorescent signal changes with the digestion time using λ DNA exonuclease.

1: (1.5–2): (2.5–3). The ratio was also consistent with the reports (*10*, *17*). Here, the concentration ratio of Cy5-dCTP/dCTP yielded ~1.1–2.5% self-quenching in the homopolymeric runs.

The Removal of the Extended Primer—The results of three different means of removing primers are shown in Fig. 5A. In the first means, NaOH was used to denature the complex of the primer and template. The ARFI of spots was 25,880 before removing the extended primer, and decreased to 2419 after treatment. So, the remaining signal ratio after and before removing the primer was 9.4%. In the second means, λ DNA exonuclease was applied to digest the extended primer. The ARFI of spots dropped to 1387 from 25,325 after digestion. Therefore, the remaining signal ratio was only 5.4%. T7 DNA exonuclease was performed in removing the extended primer in the third means. The ARFI of spots declined from 25,773 to 5183 after digesting the extended primer, as a result, there was a relative high remaining ratio of 20.1%. Obviously, these results indicate that λ DNA exonuclease was a relative excellent means of removing the extended primers due to being a relative low remaining ratio. So, we chose the digestion of λ DNA exonuclease as a standard means of removing the primers on microarray in this study.

To obtain the effective time of removing the primer by digestion of λ DNA exonuclease, the digestion time was investigated (Fig. 5B). These results indicate that the remaining ratio decreased with the increase of digestion time, the ratio quickly decreased to the bottom of 5% in 5 min, and maintained stable after 5 min digestion.

Sequencing a Gene Fragment—In order to further demonstrate the accuracy of MHES, we sequenced the fragment of DTNBP1 gene. 189 bp fragment was amplified and purified, and the purified products were spotted on the slide. The sequencing result is shown in Fig. 6. We found that 5′-GAC-3′ could be determined in the first cycle. Similarly, the sequences of 5′-TCTCC-3′ were determined in the second cycle, and 5′-GC-3′ in the third cycle. Therefore, the sequences along the primered template were 5′-GACTCTCCGC-3′. The sequencing results indicated that MHES method is reliable and accurate to be used to sequence the unknown fragment of gene. The result is consistent with the result from a Sanger sequencing (Fig. 7).

DISCUSSION

Quantitatively Determining the Sequencing Signal—To accurate quantification for the fluorescent signal is a key to obtain the sequence information. Four main sources would disturb the determination for the fluorescent signal in sequencing.

The first source is from misincorporation of nucleotide. Since the polymerase typically has multiple opportunities to insert the nucleotide into a non-complementary strand while the correct incorporations are completed. Therefore, if misincorporation happens, subsequent incorporation will be less efficient, moreover, eventually read-length will be limited. In this study, we selected Klenow polymerase as sequencing polymerase. We found that the misincorporation ratio of the heterogenous-labelled nucleotides (Cy5-dATP and Cy5-dGTP) was greatly reduced to 0.2%/extension, and the misincorporation ratio of the homogenous one (Cy5-dCTP) was only 1.6% (Fig. 3). Since the read-length ranged only from 2 to 5 bases in each cycle, although the homogenous misincorporation ratio was a relative high, it could not influence the accuracy of sequencing result. Compared to the current sequencing polymerases such as Thermal polymerase and T7 sequences, Klenow polymerase has multiple advantages. For example, the 25–40°C reaction temperature benefits the firmly binding between the primer and template. Additionally, a relative low price is another merit. The MHES method is superior to the previous SBS methods in terms of the price and the reaction condition.

Additionally, each Cy5-dNTP was diluted to different ratio by the nature dNTP due to their different discrimination of incorporation by the polymerase. Their ratios ranged from 1:100 to 1:500 (see in 'Materials and methods' section). However, the concentration of each Cy5-dNTP was still 0.1 μM, which is only

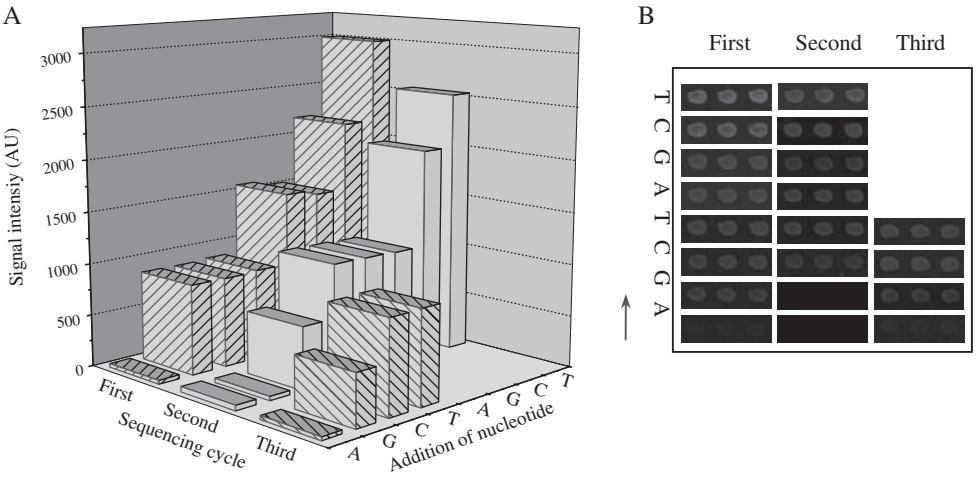

Fig. 6. **Sequencing a fragment in DTNBP1 gene.** (A) The ARFI analysis of scanning image. (B) Scanning image of spots. The addition of nucleotides was 'AGCT' in each cycle.

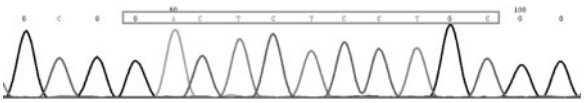

Fig. 7. **Sanger sequencing result for DTNBP1 gene.** The framed region was the sequences read by using MHES method.

2–4% of that reported (*17*, *20*). Our results have shown that the signal intensity linearly increase with the addition of incorporated bases by using these ratios of Cy5-dNTP and dNTP (Figs. 2 and 6). Moreover, using of the low concentration Cy5-dNTP could greatly decrease the sequencing cost as well.

The second is from the removal of primer. Because the remaining primer could cause a false-positive signal in the next cycle, it is most important for the next cycle sequencing to remove the extended primers in MHES method. The results have shown that the removal of the primer by using λ DNA exonuclease was a preferable one due to its lower remaining ratio of primers. Moreover, λ DNA exonuclease digestion was a relative gentle reaction condition in comparison to NaOH denaturation, and could cause little dissociation of templates from microarray.

The third is from self-quenching of the adjacent bases labelled with fluorophore. When two identical fluorescent molecules are in close proximity, their fluorescence emission is quenched due to the intermolecular interaction (*21*). Decreasing the concentration ratio of fluorescein-dNTP to dNTP will decrease the self-quenching, thus increases the linear relation between their fluorescent intensity and the number of bases incorporated. Aksyonov *et al*. (*17*) reported that when the concentration ratio of fluorescein-dNTP to dNTP was decreased to 1 : 1 from 10 : 1, self-quenching in the homopolymeric runs reduced. In this study, we further decreased the concentration ratio of fluorescein-dNTP to dNTP further ranging from 1 : 100 to 1 : 500 (see in 'Materials and methods' section). The results have shown that self-quenching was dramatically fallen. Therefore, the number of base extended could be also accurately determined through accumulating the fluorescence intensity (Figs. 2 and 6).

The fourth is from the fluorescent background on microarray. For fluorescent imaging of microarrays, it is important that surface does not bind any fluorescent species (*e.g*. fluorescein-dNTP, the fluorescing impurities). If it happens, then correct quantification of spot signals is impossible and imaging sensitivity is impaired. To alleviate the amount of the non-covalently bound fluorescent species on the surface of the slides as possibly, two measures were taken. One is to use the fresh reagents including siloxanes and crosslinkers to prepare the surface treatment of the slides. Especially, it is noted that the siloxanes should be stored under dry nitrogen at room temperature, since they are used for glass surface functionalization and are moisture sensitive. Another is washing buffer for microarray. In current work, we found that the washing buffer (0.2% TrionX-100, 2% SSC-0.1% SDS) gave lower background than previous washing buffer (SSC-0.1% SDS-0.2%) (data not shown).

Attachment from Primer-template and from Template-slide—Both the attachments of the primer-template and the template-slide must withstand multiple reactions and rinse cycles. In the case of attachment of the primer-template, they could be prevented from detaching by using the longer sequencing primer such as 24 mers (Table 1) and low extension temperature. If the sequencing primer is shorter, some fraction of the primers could detach from the substrate in reaction (*17*). Additionally, PT1 was designed to easily form a hairpin with self-primer, as a result, the signal loss was significantly decreased due to having no detachment of the primer and template in reaction (Fig. 3). Thus, the more accurate ARFI could be obtained (*16*). The extension temperature (37°C) is far below Tm of sequencing primer (PT and PT2, about 65°C), so, the primer, in dynamics principle, could not detach from the template below Tm. In the case of the template-slide attachment, the controls of slide quality and conditions of DNA immobilization are responsible for the detachment. Large amount of experiments had shown that the stock and store of reagent

were important. The treatment conditions for the slides (see in 'Materials and methods' section) should be strictly abided by. At the same time, the hydration between the silanized slides and templates must be performed in a humid chamber at 37°C for over 2 h, which could improve the efficiency of immobilization DNA.

Throughput and Read-length—In the experiment, we only chose one gene fragment to be sequenced. It has been successfully demonstrated that MHES method was accurate and efficient in sequencing on microarray. Although the 10 bp read-length is relatively short in this study, MHES is an improvement over traditional sequencing-by-synthesis on microarrays. Usually, the diameter of each sample spot is 60 μm on microarray, the distance of two adjacent spots is 300 μm. Then, more than 10,000 of spots could be spotted on a 75 mm × 25 mm slide, indicating more than 15,000 gene fragment could be immobilized onto a slide and be sequenced at the same time. If 50 bases/spot could be sequenced, 750 kb bases would be sequenced by using a glass slide. It is enough for sequencing a genomic DNA of bacteriophage lambda (λ) (~50 kb in length) to need one slide. It would be a great potential to high-throughput sequencing if real-time computation could be applied in MHES method.

CONCLUDING REMARK

Multiple hybridization-extension sequencing is a high promising sequencing technology. The method could accurately read the sequences of the unknown gene on microarray. Compared with the previous SBS (*12*, *16*, *17*), MHES has two main advantages. One is that it is unnecessary to quench fluorophore-base during sequencing. So, it has a great potential to increase read-length in sequencing on microarray. Another advantage is the low cost of sequencing. The use of Klenow polymerase and the lower concentration of Cy5-dNTP would greatly contribute to reducing the cost of sequencing. As a principle experiment, we have successfully sequenced not only 10 bp in length from synthetic template, but also 10 bp fragment in DTBNP1 gene. Although the read-length is relatively short in this study, MHES is an improvement over traditional sequencing-by-synthesis on microarrays. Much work is still needed to increase the read-length and high-throughput, this will make fluorescent SBS on microarray a valuable practical method. MHES method has potential application in resequencing, mutation and SNP research on microarray, which should greatly facilitate the diagnostic exploitation of human genetic information.

This work was funded the Hi-Tech Research and Development Program of China Project (No. 2006AA020702) and the National Natural Science Foundation of China Project (No. 60671017).

REFERENCES

1. Franca, L.T., Carrilho, E., and Kist, T.B. (2002) A review of DNA sequencing techniques. *Q. Rev. Biophys.* **35**, 169–200
2. Shendure, J., Mitra, R.D., Varma, C., and Church, G.M. (2004) Advanced sequencing technologies: methods and goals. *Nat. Rev. Genet.* **5**, 335–344
3. Collins, F.S., Morgan, M., and Patrinos, A. (2003) The Human Genome Project: lessons from large-scale biology. *Science* **300**, 286–290
4. Ji, M., Hou, P., Li, S., He, N., and Lu, Z. (2004) Microarray-based method for genotyping of functional single nucleotide polymorphisms using dual-color fluorescence hybridization. *Mutat. Res.* **548**, 97–105
5. Drmanac, S., Kita, D., Labat, I., Hauser, B., Schmidt, C., Burczak, J.D., and Drmanac, R. (1998) Accurate sequencing by hybridization for DNA diagnostics and individual genomics. *Nat. Biotechnol.* **16**, 54–58
6. Maxam, A.M. and Gilbert, W. (1977) A new method for sequencing DNA. *Proc. Natl Acad. Sci. USA* **74**, 560–564
7. Fu, D.J., Tang, K., Braun, A., Reuter, D., Darnhofer-Demar, B., Little, D.P., O'Donnell, M.J., Cantor, C.R., and Koster, H. (1998) Sequencing exons 5 to 8 of the p53 gene by MALDI-TOF mass spectrometry. *Nat. Biotechnol.* **16**, 381–384
8. Roskey, M.T., Juhasz, P., Smirnov, I.P., Takach, E.J., Martin, S.A., and Haff, L.A. (1996) DNA sequencing by delayed extraction-matrix-assisted laser desorption/ionization time of flight mass spectrometry. *Proc. Natl Acad. Sci. USA* **93**, 4724–4729
9. Edwards, J.R., Itagaki, Y., and Ju, J. (2001) DNA sequencing using biotinylated dideoxynucleotides and mass spectrometry. *Nucleic Acids Res.* **29**, E104–E104
10. Ronaghi, M., Uhlen, M., and Nyren, P. (1998) A sequencing method based on real-time pyrophosphate. *Science* **281**, 363–365
11. Zhou, G.H., Gotou, M., Kajiyama, T., and Kambara, H. (2005) Multiplex SNP typing by bioluminometric assay coupled with terminator incorporation (BATI). *Nucleic Acids Res.* **33**, e133
12. Mitra, R.D., Shendure, J., Olejnik, J., Edyta Krzymanska, O., and Church, G.M. (2003) Fluorescent in situ sequencing on polymerase colonies. *Anal. Biochem.* **320**, 55–65
13. Braslavsky, I., Hebert, B., Kartalov, E., and Quake, S.R. (2003) Sequence information can be obtained from single DNA molecules. *Proc. Natl Acad. Sci. USA* **100**, 3960–3964
14. Hyman, E.D. (1988) A new method of sequencing DNA. *Anal. Biochem.* **174**, 423–436
15. Pastinen, T., Kurg, A., Metspalu, A., Peltonen, L., and Syvanen, A.C. (1997) Minisequencing: a specific tool for DNA analysis and diagnostics on oligonucleotide arrays. *Genome Res.* **7**, 606–614
16. Seo, T.S., Bai, X., Kim, D.H., Meng, Q., Shi, S., Ruparel, H., Li, Z., Turro, N.J., and Ju, J. (2005) Four-color DNA sequencing by synthesis on a chip using photocleavable fluorescent nucleotides. *Proc. Natl Acad. Sci. USA* **102**, 5926–5931
17. Aksyonov, S.A., Bittner, M., Bloom, L.B., Reha-Krantz, L.J., Gould, I.R., Hayes, M.A., Kiernan, U.A., Niederkofler, E.E., Pizziconi, V., Rivera, R.S., Williams, D.J., and Williams, P. (2006) Multiplexed DNA sequencing-by-synthesis. *Anal. Biochem.* **348**, 127–138
18. Zhao, S., Ferrara, K.W., and Dayton, P.A. (2005) Asymmetric oscillation of adherent targeted ultrasound contrast agents. *Appl. Phys. Lett.* **87**, nihms8459
19. Hou, P., Ji, M., He, N., and Lu, Z. (2004) Microarray-based method to evaluate the accuracy of restriction endonucleases HpaII and MspI. *Biochem. Biophys. Res. Commun.* **314**, 110–117
20. Tonisson, N., Kurg, A., Kaasik, K., Lohmussaar, E., and Metspalu, A. (2000) Unravelling genetic data by arrayed primer extension. *Clin. Chem. Lab. Med.* **38**, 165–170
21. Zhuang, X., Ha, T., Kim, H.D., Centner, T., Labeit, S., and Chu, S. (2000) Fluorescence quenching: a tool for single--molecule protein-folding study. *Proc. Natl Acad. Sci. USA* **97**, 14241–14244