Chapter 1

Preparation of Single-Stranded DNA for Pyrosequencing by LATE-PCR

Huiyong Yang, Chao Liang, Zhiyao Chen, Bingjie Zou, Qinxin Song, and Guohua Zhou

Abstract

To establish a simple method for preparing single-stranded DNA templates for pyrosequencing, the Linear-after-the-exponential (LATE)-PCR technology on the basis of Taq DNA polymerase without hot-start capacity was applied to amplify a 78-bp sequence (containing the SNP6 site), and the PCR-enhancing reagents (glycerol and BSA) were used to increase the efficiency and specialization, much more before the reagents A and B were designed to eliminate the impurity (limited primers, PPi, dNTPs, and so on), and 1–2 μL LATE-PCR products with simply treatment can be used in pyrosequencing directly. Then, five SNPs related with human breast-cancers in the BRCA1 gene were investigated, and the programs had no nonspecific signals that were made of theoretic sequences. Moreover, the genotyping of the SNPs could also be distinguished easily. The results indicated that this method can be used to prepare high quality single-stranded DNA templates for pyrosequencing and allows pyrosequencing be lower in cost, simpler in operation, and easier in automation, and the cross-contamination from sample preparation was also reduced.

Key words Pyrosequencing, Linear-after-the-exponential amplification, Preparation of single-stranded DNA template, Single nucleotide polymorphism

1 Introduction

Pyrosequencing is a novel sequencing technology that is based on single-stranded DNA (ssDNA) templates and different from Sanger technology, and this technique uses three enzymes (DNA polymerase, luciferase and ATP sulfurlyase) to catalyze the reactions from their corresponding substrates, such as adenosine 5′-phosphosulfate (APS), luciferin, and the added nucleotides, which was complementary to the template, to produce a flash of light signal with an intensity that is proportional to the number of incorporated bases. dNTPs that fail to incorporate with the template are degraded by apyrase and do not produce light signal. Pyrosequencing can provide the information of DNA sequences in

Guohua Zhou and Qinxin Song (eds.), *Advances and Clinical Practice in Pyrosequencing*, Springer Protocols Handbooks,
DOI 10.1007/978-1-4939-3308-2_1,

a short time [1, 2], whereas it does not require labeled primers or electrophoresis. The pyrosequencing technique has become an ideal platform for sequencing short DNA sequences and has a wide range of applications in the fields of single-nucleotide polymorphism (SNP) detection [3], microbial genotyping [4, 5], CpG methylation analysis [6], and so on. The pyrosequencing technique has also been used in the large-scale DNA sequencing and brings a revolutionary advancement in the field of DNA sequencing [7].

Now, the preparation of single-stranded DNA material generally is dependent on solid-phase beads (streptavidin-coated) method, which has characters of multistep process, low efficiency, and high cost for the using biotinylated primers and streptavidin beads [8, 9], and moreover, the multistep process of samples preparation also increases the risk of contamination in PCR products, so that the application of pyrosequencing in clinical is limited. The asymmetric PCR method can produce ssDNA [10–12], but the method has not been widely used for the low efficiency in amplifying ssDNA. Linear-after-the-exponential-polymerase chain reaction (LATE-PCR) developed asymmetric amplifications by introducing the concentration to compute primers melting temperature (Tm) [13–15]. Under the condition of the Tm of limited primer (PL) being higher than that of excess primers (PX) (Tm (PL) – Tm(PX) ≥ 0 °C) and optimizing the concentration, and Tm of the primers, LATE-PCR can generate large amount of ssDNA, which is sufficient for reaction of probe hybridization, real-time fluorescence PCR detection, and so on [16–19].

Wangh et al. [16] used the LATE-PCR products to pyrosequencing directly with a satisfying result. However, the application of the AmpliTaq Gold Taq polymerase and commercial PyroGold kit increased the cost as well as ignored the cases of the residual PL, and not fully extended dsDNA products that could confuse the results of pyrosequencing. In this experiment, a system of LATE-PCR that is based on Taq DNA polymerase, which was introduced without hot-start capacity and one tube method was carried out for preparing the single-stranded template for pyrosequencing. The two reagents (A and B) were designed to eliminate the interfering components in PCR products. This method was applied to pyrosequence the 78-bp sequence (including SNP6 site) LATE-PCR product, and the five SNPs site in the brca1 gene (one of breast cancer susceptibility gene [20]) were detected.

2 Materials

1. PTC-225 PCR engine (MJ Research, Inc., USA).
2. PowerPac 1000 HV power supply (BIO-RAD, Inc, USA).
3. GenGenius gel imaging system (Syngene, Inc, UK).

4. The apparatus for pyrosequencing were set up according to reference [21]: R6335 photomultiplier (Hamamatsu Photonics K.K, Inc., Japan) and BPCL Ultra-Weak Chemiluminescence Analyzer System (Institute of Biophysics of Chinese Academy of Sciences).
5. Taq DNA polymerase, Deoxynucleotide (dNTPs, 10 mM), 1 × buffer (Mg_2+ free), and $MgCl_2$ (25 mM) were purchased from TaKaRa BioTech (Dalian, China).
6. AmpliTaq Gold DNA polymerase was purchased from Applied Biosystems (Foster City, USA).
7. Klenow Fragment (Exonuclease Minus), polyvinylpyrrolidone (PVP), and QuantiLum recombinant luciferase were purchased from Promega (Madison, WI).
8. Apyrase, D-luciferin, bovine serum albumin (BSA), and adenosine 5′-phosphosulfate (APS) were obtained from Sigma (St. Louis, MO).
9. Sodium 2′-deoxyadenosine-5′-*O*-(1-triphosphate) (dATPαS), 2′-deoxyguanosine-5′-triphosphate (dGTP), 2′-deoxythymidine-5′-triphosphate (dTTP), and 2′-deoxycytidine-5′- triphosphate (dCTP) were purchased from Amersham Pharmacia Biotech.
10. ATP sulfurylase was produced in our lab [22].
11. The other chemicals were of a commercially extra-pure grade.
12. All solutions were prepared with deionized and sterilized water.
13. The blood samples (named g-1, g-2, and g-3) for pyrosequencing were obtained from healthy volunteers in our lab, and the genomes were extracted according to phenol–chloroform method.
14. The sequences containing the five SNPs were obtained from the database of NCBI (NT_010755.15), and the primers for LATE-PCR and sequencing were designed by primer 5.0, whereas the Tm of primers were calculated by OligoAnalyzer 3.0 (a Web-based program) (*see* **Notes 1–5**).

3 Methods

3.1 Conditions of Amplification

1. A total of 25 μL of PCR mixture contains 1×PCR buffer (50 mM KCl, 15 mM Tris–HCl, pH 8.0), 3 mM $MgCl_2$, 0.1 mM dNTP, 1.5 U of Taq DNA polymerase, 1 μL of DNA template (36 μg/l), 0.1 μM limited primer, 13 % glycerol, and 4 % BSA.
2. The reaction procedure consisted of incubation at 95 °C for 5 min, 60 cycles of denaturation at 87 °C for 10 s, the primer annealing at 66 °C for 10 s, extension at 72 °C for 20 s, exten-

sion at 72 °C for 10 min, and was lastly preserved at 4 °C (*see* **Notes 6** and **7**).

3. The annealing temperature of SNP2 and 4 were 64 °C.
4. The results were detected by agarose gel electrophoresis (the dye was Goldview and Agarose Gel was 2 %).

3.2 Preparation of Reaction Reagents

1. Reagent A contains 0.1 M tris-acetate (pH 7.7), 2 mM EDTA, 10 mM magnesium acetate, 0.1 % BSA, 1.0 mM dithiothreitol (DTT), 2 μM adenosine 5′-phospho sulfate (APS), 0.4 g/L PVP, 0.8 mM D-luciferin, 200 U/L ATP sulfurylase, 6 mg/L luciferase, and 18 U/mL Klenow fragment.
2. Reagent B contains reagent A and 1.6 U/mL apyrase.

3.3 Preparation of Single-Stranded DNA Templates

1. For pyrosequencing, ssDNA templates were prepared as the follows: 1–2 μL products of LATE-PCR and 1.2 μL APS (1 mM) were added in 30 μL reagent A, vortexed, and left for 10–15 min.
2. 40 μL reagent B was added, vortexed, and left for 3–5 min.
3. At last the sequencing primer was added (the final concentration was 0.3 PM) to anneal for 10–15 min.

3.4 Pyrosequencing

1. The pyrosequencing reaction used a 4-enzyme cascade system to produce a visible light that was detected by a photo-sensitive device PMT.
2. All the four dNTPs were dispensed individually, and the dispensing order of the dNTPs could be designed beforehand. The light signal was recorded by data BPCL Ultra-Weak Chemiluminescence Analyzer System (Institute of Biophysics of Chinese Academy of Sciences) and pyrograms were produced by Origin 7.5.

4 Method Validation

4.1 Detection of a 78-bp Sequence (Including SNP6 Site) LATE-PCR Product

To investigate the effect of the inexpensive Taq DNA polymerase without hot-start capacity for LATE-PCR, the 78-bp sequence (containing SNP6 site) with TaKaRa-Taq polymerase (without hot-start capacity) and AmpliTaq Gold Taq DNA polymerase were amplified. The results were shown in Fig. 1 (lanes 1 and 3), and the effect was more excellent than that of TaKaRa-Taq polymerase. To improve the amplified efficiency and specialization of the TaKaRa-Taq polymerase, PCR-enhancing reagents (glycerol and BSA) were applied, and the extended time were shortened in the thermal cycle (according to the length of Amplicon and the extended speed 1000 bp min^{-1}), and the denaturalization temperatures were fallen (according to Tm of amplicon by computing

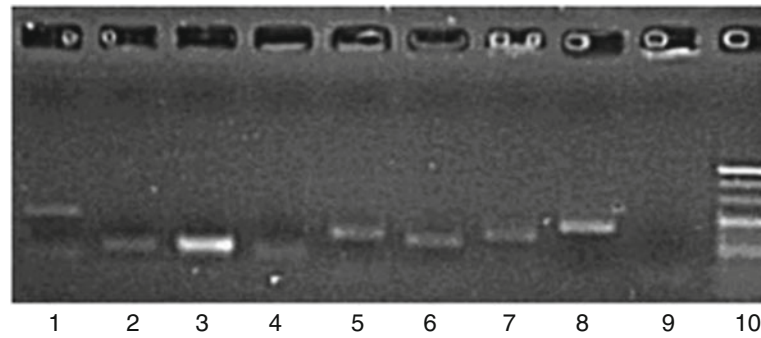

Fig. 1 Agarose electropherogram of amplicons from LATE-PCR at different conditions. Lane 1, the amplicon of the 78 bp sequence from TaKaRa-Taq DNA polymerase; lane 2, the amplicon of the 78-bp sequence from TaKaRa-Taq DNA polymerase (13 % of glycerol and 4 % of BSA); lane 3, the amplicon of the 78 bp sequence from AmpliTaq Gold DNA polymerase; lanes 4–9, the amplicons of 5 SNPs in BRCA1 gene from TaKaRa-Taq DNA polymerase (13 % of glycerol and 4 % of BSA); lane 4, SNP1 (64 bp); lane 5, SNP2 (121 bp); lane 6, SNP3 (102 bp); lane 7, SNP 4 (123 bp); lane 8, SNP 5 (171 bp); lane 9, blank control; lane 10, the DNA Maker (20–200 bp)

based on reference [25]). The lane 2 in Fig. 1 indicted that the amplified efficiency and specialization of the TaKaRa-Taq polymerase were improved obviously. The agarose electropherogram of the five SNP sites (brca1 gene, lane 4–8) also indicated that the Taq polymerase without hot-start capacity was available to be used in LATE-PCR by optimizing reaction condition.

Raw PCR products contain PPi, and dNTPs will consume the APS luciferin metabolized by the pyrosequencing enzymes, resulting in indiscriminant light emission, which not only interfered in the pyrosequencing reactions but also confused sequence determination. To obtain the ideal ssDNA templates, the LATE-PCR products were cleaned up by the followed steps: the PPi was converted to ATP using ATP sulfurylase and APS, and then the ATP and dNTPs were digested by apyrase to AMP and dAMP. However, the extended signals were detected from the PCR products without annealing with sequencing primers (after they were annealed by the two processes), see Fig. 2a. By analysis, the signals were brought on by the residual limited primers or not fully extended dsDNA products, and therefore, one process that used the Klenow DNA polymerase to extend the interfering matters was designed. The results of real-time monitoring the process of removing the interfering matters indicated that there were obvious extending reactions (as shown in Fig. 2b (I)). When the background signal descended and became stabilizing, the four dNTPs were dispensed individually and no obvious extended signals were observed, see Fig. 2b (II). By pyrosequencing the LATE-PCR product after annealing with sequencing primers, the pyrogram obtained was

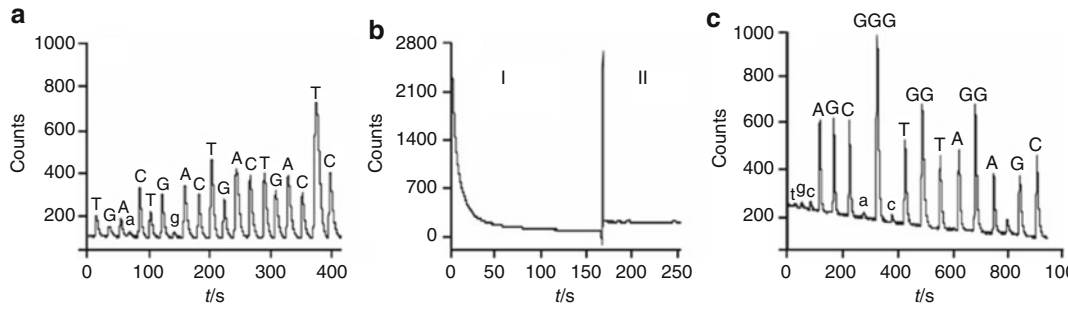

Fig. 2 Pyrograms of LATE-PCR products of 78 bp sequence treated by different ways. (**a**) directly pyrosequenced without treatment by Klenow; (**b**) real-time background signals of the PCR products with treatment by Klenow (I) and the background signals after dispensed the four dNTPs individually (II); (**c**) pyrosequenced after treated according to the protocol in Sect. 3, item 3.3, the referenced sequence: AGCGGGTGGTAGGAGC; the dispensing order of dNTPs: TGCAGCAGTGTAG ATGC. The *capital letters* on the top of each peak indicated that the bases completed the extension reaction, the *small letters* indicated the background signals after adding the dNTPs that did not complete the extension reaction

shown in Fig. 2c. The results were consistent with the theoretical sequence, and it also indicated that the results of reference [16] could be obtained by LATE-PCR with TaKaRa Taq DNA polymerase (without hot-start capacity).

4.2 Detection of Five SNPs (brca1 Gene)

Under the optimized conditions, the five SNPs (brca1 gene), including base transition, insertion, and deletion, were detected. The pyrograms of the sample g-1 obtained were shown in Fig. 3. The heights of the signal peaks were effectively proportional to the number of bases in homopolymeric regions of the templates. However, no signals were observed after adding the dNTP that was not complementary to the template. It is reasonable that to say that introducing the methods that are based on the LATE-PCR products using into pyrosequencing is feasible.

The genotyping of SNPs by pyrosequencing were based on the signal peaks of the pyrograms. In Fig. 3, the genotyping carried out for SNP1, SNP2, SNP3, and SNP4 were wild-type, and the genotypes were AA, CT+/+, C−/−, and CC, respectively. However, when the SNPs locate homopolymeric regions, the results will have confusion to determine. For example, the genotyping of SNP5 from Fig. 3e was the confusion that the TT should be the output of the programs; however, the heterozygosis of CT was also logical according to the height of the signal peaks, and the sequence should be CT(C)1.5(T)0.5GTTTTC in all probability. The result of Sanger method also testified the CT genotyping. Because of this, we designed a new sequencing primer (Bs5b) that 3′-terminal was adjacent to SNP5 site, and the genotyping could be determined easily (Fig. 4).

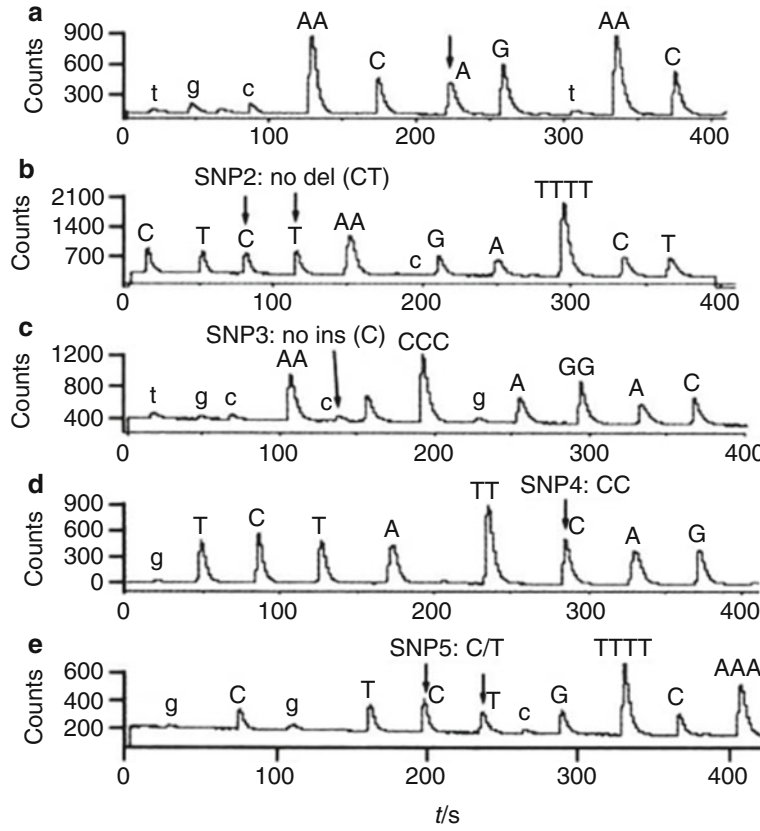

Fig. 3 Pyrograms of five SNPs in BRCA1 gene obtained. (**a**) SNP1: referenced sequence AAC[A > G]GAACT, the dispending order of dNTPs was TGGCACAGTAC and the measured genotyping was AA; (**b**), SNP2: referenced sequence CT[CT > del]AAGATTTTCT, the dispending order of dNTPs was CTCTACGATCT and the measured genotyping was CT+/+; (**c**) SNP3: referenced sequence was AA[C >ins]TCCCAGGAC, the dispending order of dNTPs was TGCACTCGAGAC and the measured genotyping was C–/–; (**d**) SNP4: referenced sequence was TCTATT[C > T]AG, the dispending order of dNTPs was GTCTAGTCAG and the measured genotyping was CC; (**e**) SNP5: referenced sequence was CTC[C > T]GTT TCAAA, the dispending order of dNTPs was GCGTCTCGTCA and the measured genotyping was CT; and others were the same as in Fig. 2)

5 Technical Notes

1. OligoAnalyzer 3.0 is available at http://scitools.idtdna.com/analyzer/.The Tm of primers was estimated by the nearest neighbor formula, and the obtained Tm was closer to the real Tm than that obtained by Primer 5.0; and the Tm was different in a same primer for PCR amplification and for sequencing because the concentration of primers, and the reaction system were different [26].

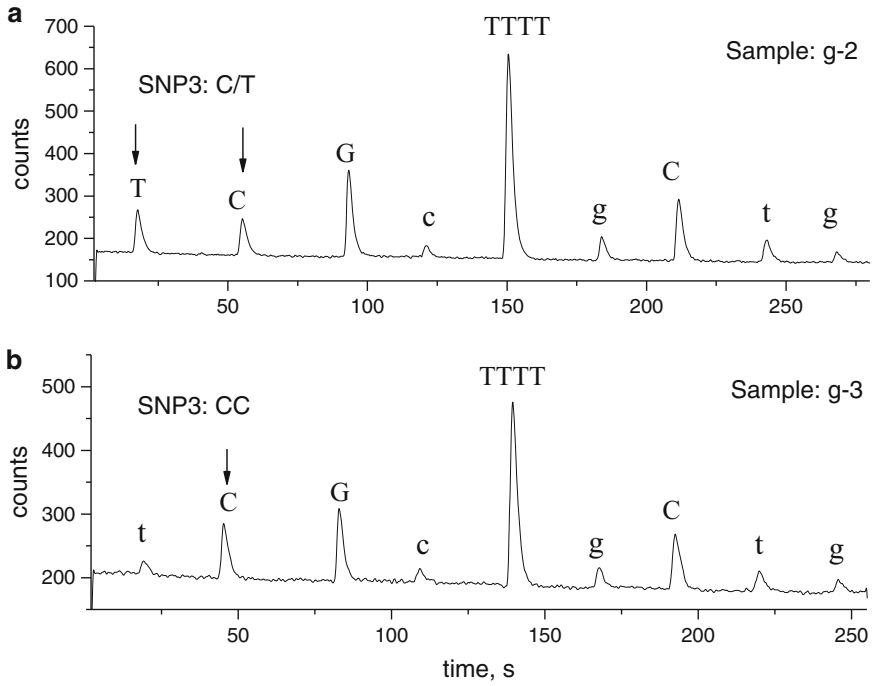

Fig. 4 Pyrograms of SNP5 with annealed by Bs5b sequencing primer. (**a**) the measured genotyping of sample g-2 was TC; (**b**) the measured genotyping of sample g-3 was CC; referenced sequence of SNP5 was [T > C] GTTTTC, the dispending order of dNTPs was TCGCTGCTG and the others were the same as in Fig. 2

2. The efficiency of the ssDNA produced by LATE-PCR is related to the Tm of the primers, so that the design of the primers is a key factor.
3. During the phase of exponential amplification, the limited primer has to be exhausted for pyrosequencing under the condition of Tm(PL) – Tm(PX) ≥ 5 °C. During the phase of linear amplification, the one stranded DNA of PCR products and the excess primer will compete for hybridization to the limiting primer strand [23].
4. Sustained production of single stranded amplification is achieved under the condition of the Tm(Amplicon) – Tm(PX) ≤ 13–18 °C [13, 23].
5. When the adequate Tm of primers cannot be designed to the actual DNA sequence of the target gene because the nucleotides A and T are abundant, the Tm can be increased by substituting one or two guanine bases for adenine near the 5′-end of the primer or adding of cytosine or guanidine to the 5′-end of the primer, irrespective of the target sequence [24]. According to this principle, one or three mismatched bases were introduced in the primers.

6. For LATE-PCR reactions, a higher anneal temperature of the thermal cycle is used during the phase of exponential amplification, and then, a lower anneal temperature is used during the phase of linear amplification. According to the results of the experiment, the ssDNA products will begin to amplify from 30 thermal cycle when the rate of PL:PX is 0.1:1 (the concentration of PX is 1 μM [16]), and the balance for pyrosequencing is required between maximizing the proportion of single-stranded material in the final product and minimizing the number of thermal cycles for minimizing the interfering PCR components.
7. According to the reference [16], the annealing temperatures were adjusted to below 1–2 °C to the Tm of the limited Primers in all thermal cycles.

References

1. Zhang XD, Wu HP, Zhou GH (2006) Pyrosequencing and its application in genetical analysis. Chinese J Anal Chem 34:582–586
2. Ronaghi M (2001) Pyrosequencing sheds light on DNA sequencing. Genome Res 11:3–11
3. Huang XQ, Roder MS, Roder J (2005) Development of SNP assays for genotyping the puroindoline b gene for grain hardness in wheat using pyrosequencing. J Agric Food Chem 53:2070–2075
4. Pai R, Limor J, Beall B (2005) Use of pyrosequencing to differentiate Streptococcus pneumoniae serotypes 6A and 6B. J Clin Microbiol 43:4820–4822
5. Pourmand N, Diamond L, Garten R, Erickson JP, Kumm J, Donis RO, Davis RW (2006) Rapid and highly informative diagnostic assay for H5N1 influenza viruses. PLoS ONE 1, e95
6. Tost J, Gut IG (2007) DNA methylation analysis by pyrosequencing. J Nat Protoc 2:2265–2275
7. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380
8. Diggle MA, Clarke SC (2003) Pyrosequencing™. Mol Biotechnol 24:221–224
9. Royo JL, Hidalgo M, Ruiz A (2007) Pyrosequencing protocol using a universal biotinylated primer for mutation detection and SNP genotyping. Nat Protoc 2:1734–1739
10. Fujii T, Ohta M, Kono M, Hoshina S, Fukuhara K, Tsuruoka M (1999) Rapid detection of the gene of Legionella pneumophila using the fluorescence polarization with the asymmetric PCR. Nucleic Acids Symp Ser 42:59–60
11. Barratt K, Mackay JF (2002) Improving real-time PCR genotyping assays by asymmetric amplification. J Clin Microbiol 40:1571–1572
12. Wei Q, Liu S, Huang J, Mao X, Chu X, Wang Y, Qiu M, Mao Y, Xie Y, Li Y (2004) Comparison of hybridization behavior between double and single strand of targets and the application of asymmetric PCR targets in cDNA microarray. J Biochem Mol Biol 37:439–444
13. Pierce KE, Sanchez JA, Rice JE, Wangh LJ (2005) Linear-after-the-exponential (LATE)-PCR: primer design criteria for high yields of specific single-stranded DNA and improved real-time detection. J Proc Natl Acad Sci USA 102:8609–8614
14. Allawi HT, SantaLucia J Jr (1997) Thermodynamics and NMR of internal G.T mismatches in DNA. Biochemistry 36:10581–10594
15. Peyret N, Seneviratne PA, Allawi HT, SantaLucia J (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. J Biochem 38:3468–3477
16. Salk JJ, Sanchez JA, Pierce KE, Rice JE, Soares KC, Wangh LJ (2006) Direct amplification of single-stranded DNA for pyrosequencing using linear-after-the-exponential (LATE)–PCR. J Anal Biochem 353:124–132
17. Abravaya K, Huff J, Marshall R, Merchant B, Mullen C, Schneider G, Robinson J (2003) Molecular beacons as diagnostic tools: technol-

ogy and applications. Clin Chem Lab Med 41:468–474

18. Sanchez JA, Abramowitz JD, Salk JJ, Reis AH Jr, Rice JE, Pierce KE, Wangh LJ (2006) Two-temperature LATE-PCR endpoint genotyping. J BMC Biotechnol 6:44
19. Hartshorn C, Eckert JJ, Hartung O, Wangh LJ (2007) Single-cell duplex RT-LATE-PCR reveals Oct4 and Xist RNA gradients in 8-cell embryos. J BMC Biotechnol 7:87
20. Starita LM, Parvin JD (2003) The multiple nuclear functions of BRCA1: transcription, ubiquitination and DNA repair. J Curr Opin Cell Biol 15:345–350
21. Zhou G, Kamahori M, Okano K, Harada K, Kambara H (2001) Miniaturized pyrosequencer for DNA analysis with capillaries to deliver deoxynucleotides. Electrophoresis 22:3497–3504
22. Luo J, Wu WJ, Zou BJ, Zhou GH (2007) Expression and purification of ATP sulfurylase from saccharomyces cerevisias in escherichia coli and its application in pyrosequencing. Chinese J Biotech 23:623–627
23. Rice JE, Sanchez JA, Pierce KE, Reis AH Jr, Osborne A, Wangh LJ (2007) Monoplex/multiplex linear-after-the-exponential-PCR assays combined with PrimeSafe and dilute-'N'-Go sequencing. Nat Protoc 2:2429–2438
24. Thornhill A (2007) Single cell diagnostics, 7th edn. Humana Press Inc, Totowa, NJ, pp 65–85
25. Wetmur JG (1991) DNA probes: applications of the principles of nucleic acid hybridization. Crit Rev Biochem Mol Biol 26:227–259
26. Agah A, Aghajan M, Mashayekhi F, Amini S, Davis RW, Plummer JD, Ronaghi M, Griffin PB (2004) A multi-enzyme model for pyrosequencing. Nucl Acids Res 32, e166