

Direct amplification of single-stranded DNA for pyrosequencing using linear-after-the-exponential (LATE)–PCR

Jesse J. Salk^{a,b,c,*}, J. Aquiles Sanchez^a, Kenneth E. Pierce^a, John E. Rice^a,
Kevin C. Soares^a, Lawrence J. Wangh^{a,*}

^a Department of Biology, Brandeis University, Waltham, MA 02454, USA

^b Human Biology Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98019, USA

^c Microscale Life Science Center, Department of Electrical Engineering, University of Washington, Seattle, WA 98195, USA

Received 7 January 2006

Available online 28 February 2006

Abstract

Pyrosequencing is a highly effective method for quantitatively genotyping short genetic sequences, but it currently is hampered by a labor-intensive sample preparation process designed to isolate single-stranded DNA from double-stranded products generated by conventional PCR. Here linear-after-the-exponential (LATE)–PCR is introduced as an efficient and potentially automatable method of directly amplifying single-stranded DNA for pyrosequencing, thereby eliminating the need for solid-phase sample preparation and reducing the risk of laboratory contamination. These improvements are illustrated for single-nucleotide polymorphism genotyping applications, including an integrated single-cell-through-sequencing assay to detect a mutation at the globin IVS 110 site that frequently is responsible for β -thalassemia.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Pyrosequencing; LATE–PCR; Asymmetric PCR; Single-stranded DNA; Single-cell; Single-nucleotide polymorphism; IVS 110; β -Thalassemia

Pyrosequencing [1,2] is a sequencing-by-synthesis technology that employs four enzymatic reactions to quantitatively monitor nucleotide incorporation in real time. Nucleotides are sequentially added to a homogeneous solution containing a template–primer hybrid plus several enzymes (Klenow DNA polymerase, potato apyrase, ATP sulfurylase, and firefly luciferase) and their corresponding substrates (adenosine 5'-phosphosulfate [APS]¹ and luciferin). The addition of a nucleotide that is complementary to the base on the template strand immediately

following the 3' end of the sequencing primer results in incorporation of this nucleotide by the polymerase and release of an equimolar amount of inorganic pyrophosphate (PPi). The pyrophosphate produced in this process is then rapidly used by ATP sulfurylase to convert APS to ATP, which in turn is metabolized by luciferase/luciferin to produce a flash of light with an intensity that is proportional to the number of incorporated bases. Nucleotides that fail to incorporate are degraded by apyrase and a new cycle of addition begins. Sequence is determined from the order and relative magnitude of the light peaks resulting from each nucleotide addition.

Pyrosequencing has advantages over traditional sequencing methods for many applications in that it is faster, provides more quantitative allele ratio information, and does not require labeled primers or electrophoresis. The pyrosequencing technique has become more widely used during the past several years as reaction chemistry [3,4] and understanding of reaction kinetics [5,6] have improved

* Corresponding authors. Fax: +1 206 667 6132 (J.J. Salk), +1 781 736 3107 (L.J. Wangh).

E-mail addresses: jsalk@fhcrc.org (J.J. Salk), wangh@brandeis.edu (L.J. Wangh).

¹ Abbreviations used: APS, adenosine 5'-phosphosulfate; PPi, inorganic pyrophosphate; SNP, single-nucleotide polymorphism; ssDNA, single-stranded DNA; LATE–PCR, linear-after-the-exponential–PCR; dsDNA, double-stranded DNA; Pi, inorganic phosphate.

and as higher throughput hardware/software (www.pyrosequencing.com) and chemistries [7,8] have been developed. Published applications have included microbial genotyping [9–12], mutation detection [13], sequencing of heteroplasmic DNA [14], single-nucleotide polymorphism (SNP) genotyping [15,16], SNP allele frequency determination [17,18], SNP-based allelic loss mapping [19], gene copy number change [20], and CpG methylation analysis [21–23].

The pyrosequencing reaction typically is carried out isothermally near 28 °C because three of the enzymes involved are heat labile. At these low temperatures, accessibility of the sequencing primer to the DNA target becomes a concern [3]. Most users minimize this problem by sequencing single-stranded DNA (ssDNA) templates where primers do not need to compete with a tightly annealed complementary strand. Single-stranded material generally is obtained from double-stranded PCR products by end-labeling one of the PCR primers with biotin and then using streptavidin-coated beads with either magnetic or vacuum equipment [8] to separate the labeled amplicon strands from unlabeled strands as well as PCR by-products. Although the development of robotically controlled pyrosequencing machines has helped to increase sample throughput, a processing bottleneck remains at this strand separation step. The multistep nature of this sample preparation process also increases the risk of contaminating a laboratory with PCR products. This is of particular concern when later amplifications are initiated with a small number of starting templates and even minimal contamination represents a large fraction of the input sample.

Several alternative approaches for template preparation have been developed previously in an effort to further automate and accelerate sample preparation. One group has reported successful sequencing of double-stranded PCR products after removing interfering PCR components with enzymes and/or blocking oligonucleotides [24–26]. Although attractive in principle, this method has not come into widespread use, presumably due to the difficulty in optimizing each reaction to yield adequate and consistent signals [27]. Another approach generates single-stranded material by using one 5'-phosphate-labeled primer for PCR, followed by digestion with lambda exonuclease [28]. Although this generates single-stranded template, workup is lengthy and PCR by-products must be removed by solid-phase filtration.

Linear-after-the-exponential (LATE)-PCR is a novel, highly robust, asymmetric PCR method that uses unequal concentrations of primers to generate large amounts of ssDNA together with a small fixed amount of double-stranded DNA (dsDNA) [29,30]. Although the concept of asymmetric PCR has been in the literature for more than 15 years [31], optimizing reactions to be efficient and reproducible historically has been problematic. LATE-PCR introduces an improved primer design algorithm that results in reliable asymmetric amplifications by taking into

account the fact that primer T_m is concentration dependent. We recently described additional aspects of LATE-PCR primer design protocols to maximize product specificity and single-stranded product yield [32].

Here we introduce a rapid, single-tube method for preparing single-stranded template for pyrosequencing and eliminating other interfering PCR components using LATE-PCR and a straightforward, 1-min enzymatic cleanup. Direct amplification of single-stranded material in this approach obviates the need for strand isolation and allows pyrosequencing to be accomplished in the same reaction chamber, thereby significantly reducing the risk of contamination. We demonstrate this technique for biologically relevant SNP genotyping down to the single-cell, single-molecule level.

Materials and methods

Cells and DNA

Human lymphoblast cell lines heterozygous (GM07425) and homozygous (GM07405) for the β -globin IVS 110 mutation were obtained from the Coriell Cell Repositories (Camden, NJ, USA) and prepared as described previously [33]. Single cells were transferred directly into PCR tubes containing 10 μ l of a PCR-compatible, proteinase K-based lysis solution [34] and were incubated at 50 °C for 30 min and then at 95 °C for 10 min. Previously extracted DNA from frozen gastric biopsies [35] was obtained from a repository maintained by the Seattle Barrett's Esophagus Program at the Fred Hutchinson Cancer Research Center (Seattle, WA, USA).

PCR primers

LATE-PCR primer sets were designed to amplify a 191-bp region containing the β -globin IVS 110 site (human chromosome 11p), a 95-bp region containing the rs858521 SNP site (human chromosome 17p), and a 78-bp region containing the rs2270517 SNP site (human chromosome 17p) using previously described methods [29,30,32]. Briefly, primer pairs were selected such that $\Delta T_m = (T_m^{\text{limiting}} - T_m^{\text{excess}}) \geq 0$ based on concentration-adjusted values. In addition, amplicons were selected with the limitation that $(T_m^{\text{amplicon}} - T_m^{\text{excess}}) \leq 13$ °C. Primer and amplicon T_m values were estimated by the nearest neighbor formula [36] and a salt adjustment [37] using input concentrations of 1000 nM for the excess primer and amplicon, 100 nM for the limiting primer, and 50 mM for monovalent cations (OligoAnalyzer 3.0, a Web-based program to compute these values, is available at <http://scitools.idtdna.com/analyzer/>). This algorithm underestimates T_m in PCR buffers containing Mg^{2+} but is helpful in determining ΔT_m of primer pairs. The following sequences were purchased from Operon Biotechnologies (Huntsville, AL, USA). Concentration-adjusted T_m estimations are as noted:

IVS 110 excess:

5'-TGGGTTTCTGATACGCACTGACTCTCTC-3'
(62.6 °C)

IVS 110 limiting:

5'-GGCCATCACTAAAGGCACCGAGCACT-3'
(63.7 °C)

rs858521 excess:

5'-CAATCCCTTGACCTGTTGTGGAGAGAA-3'
(61.8 °C)

rs858521 limiting:

5'-TCCCCAGAGCCCAGCCGGTGTCATTTTC-3'
(66.5 °C)

rs2270517 excess:

5'-GATGGGTGGAGCTTGTCTTGAGG-3' (61.4 °C)

rs2270517 limiting:

5'-GGTCAGCGCCGGGCTGCAAGTGTAGA-3'
(67.3 °C).

PCR conditions

LATE-PCRs were carried out in a final volume of 25 μ l with final concentrations of 3 mM MgCl₂, 100 μ M each of dATP, dGTP, dCTP, and dTTP, 100 nM limiting primer, 1000 nM excess primer, and 1.5 U AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA, USA) in supplied 1 \times PCR buffer (50 mM KCl, 15 mM Tris-HCl, pH 8.0). In addition to the above, the rs858521 and rs2270517 reactions contained a PCR-enhancing reagent that will be described elsewhere. Reactions were initiated with 20 genome equivalents worth of DNA or with directly prepared single cells as noted. Thermocycling was carried out on an MJ Research PTC-200 thermocycler using the following profiles. IVS 110 amplicon: 95 °C for 10 min, followed by 60 cycles (65 for single cells) of 95 °C for 10 s, 66 °C for 15 s, and 72 °C for 20 s; rs858521 and rs2270517 amplicons: 95 °C for 10 min, followed by 60 cycles of 95 °C for 10 s, 66 °C for 10 s, and 72 °C for 20 s. Contamination control measures for reaction preparation were as described previously [33].

Template preparation

PCR products were prepared for pyrosequencing either enzymatically or using DNA binding columns. In the first method, residual dNTP and PPI left over from LATE-PCR are removed using apyrase, ATP sulfurylase, and luciferase present in the standard reaction volume of PyroGold enzyme mix (Biotage, Uppsala, Sweden). For this approach, 5.0–7.5 μ l of LATE-PCRs was transferred to wells on a PSQ HS 96-well optical plate (Biotage) containing 6.64–9.96 μ l of 20 mM Tris-acetate (pH 7.6) with 2 mM MgAc₂ and was placed in a PSQ HS 96A Pyrosequencer (Biotage). A standard pyrosequencing reaction volume of PyroGold enzyme mix and approximately twice (2.09 \pm 0.07 [3 SD]) as determined by analytical balance) the standard volume of substrate mix (Biotage) was injected into each well using the following instrument set-

tings: enzyme mix pulse time, 23.5 ms; substrate mix pulse time, 44.0 ms; reagent dispensation pressure, 400 mbar. Samples were incubated in the pyrosequencing machine until light output dropped below background signal (\leq 1 min), indicating complete digestion of dNTP and PPI. Afterward, 0.36–0.54 μ l of 10 μ M sequencing primer was added to a final concentration of 300 nM for a total reaction volume of 12–18 μ l. Samples were then heated to 80 °C for 2 min to heat denature all enzymes and were allowed to cool to room temperature for 10 min. Sequencing primers (Operon Biotechnologies) were as follows:

IVS 110: 5'-GACCACCAGCAGCCTAAG-3'

rs858521: 5'-CCCCTTCAGCTCAAA-3'

rs2270517:

5'-GGTCAGCGCCGGGCTGCAAGTGTAGA-3'.

In addition, IVS 110 reactions contained 900 nM of a 3' phosphorylated version of the limiting LATE-PCR primer to prevent amplicon self-priming at the 3' end. Alternatively, for separation of leftover dNTP and PPI from PCR using DNA binding columns, PCR products were isolated from 25- μ l reactions on MinElute PCR purification columns (Qiagen, Valencia, CA, USA) and eluted in an equivalent volume of 10 mM Tris-HCl (pH 8.5). Then five μ l of this elution was used in 12- μ l sequencing reactions as prepared above.

Pyrosequencing

Pyrosequencing was carried out at 28 °C on an automated PSQ HS 96A Pyrosequencer using PyroGold reagents according to the manufacturer's protocol. After an initial dispensation of enzyme and substrate mixes, the sequencing procedure was accomplished by a stepwise elongation of the primer strand through iterative additions of deoxynucleoside triphosphates and simultaneous detection of resulting light emission.

Results

Logic of LATE-PCR assay design for pyrosequencing

Although several previous reports [24–26] have described the use of dsDNA templates for pyrosequencing, most pyrosequencing is carried out using ssDNA templates to boost reaction efficiency by preventing complementary strands from competing with the sequencing primer for hybridization to the template strand. LATE-PCR generates more single-stranded product than double-stranded product, thereby making it an attractive method for generating pyrosequencing substrate without an additional strand separation step. The actual amount of single-stranded amplicon produced by the end of the linear phase of a LATE-PCR depends on the number of double-stranded amplicons generated during the initial exponential phase

of the reaction, the efficiency of linear amplification, and the number of linear cycles used. The number of double-stranded amplicons, in turn, is fixed by the initial concentration of limiting primer. Thus, in practice the total number of single-stranded molecules produced and the final ratio of double-stranded material to single-stranded material can be adjusted by altering either the number of linear cycles or the ratio of limiting primer to excess primer. In designing LATE-PCRs for pyrosequencing, we sought to strike a balance between maximizing the proportion of single-stranded material in the final product and minimizing the number of thermal cycles required to achieve that level. We found that ample single-stranded DNA is generated when 30 linear cycles are employed with a 100:1000-nM ratio of limiting primer to excess primer.

Direct comparison of LATE-PCR and symmetric PCR products for pyrosequencing

Side-by-side symmetric and LATE-PCR amplifications were carried out using three sets of LATE-PCR primers to directly compare pyrosequencing of dsDNA to pyrosequencing of predominantly ssDNA. LATE-PCR and symmetric PCRs were set up as described in Materials and methods, with the latter reactions using equal primer concentrations of 1000 nM. All reactions were initiated with 20 genomes worth of DNA. After 60 cycles, half of the LATE-PCR amplifications were spiked with an additional 1 μ l of limiting primer such that the total amount of incorporated and unincorporated limiting primer equaled the amount of incorporated and unincorporated excess primer (i.e., 1000 nM). An equivalent volume of distilled water was added to the other half of the LATE-PCRs and to all of the symmetric amplifications. All reactions were then subjected to one more cycle of 95 °C for 10 s, 65 °C for 20 s, and then 72 °C for 10 min to convert the single strands in the spiked LATE-PCRs with added limiting primer to a double-stranded form. All samples were then column purified, heat denatured, annealed with a sequencing primer, and subjected to pyrosequencing.

Representative pyrograms from three sequences amplified under LATE-PCR, double-strand-converted LATE-PCR, and symmetric PCR conditions are shown in Fig. 1. Plots within each row are adjusted to the same relative light unit (*y* axis) scale. Average single-base incorporation peak heights from quadruplicate PCRs are tabulated in Table 1. In symmetric and double-strand-converted LATE-PCRs, the average's fraction of the corresponding LATE-PCR signal is noted in parentheses.

Because both the symmetric and converted LATE-PCR samples contain only dsDNA, differences in signal between the two reaction types using the same amplicon are due solely to differences in the amount of final product. In the cases of the rs858521 and IVS 110 symmetric amplifications, average signal strength was slightly better in the symmetric PCR samples than in the corresponding converted samples; hence, slightly more product was made than in

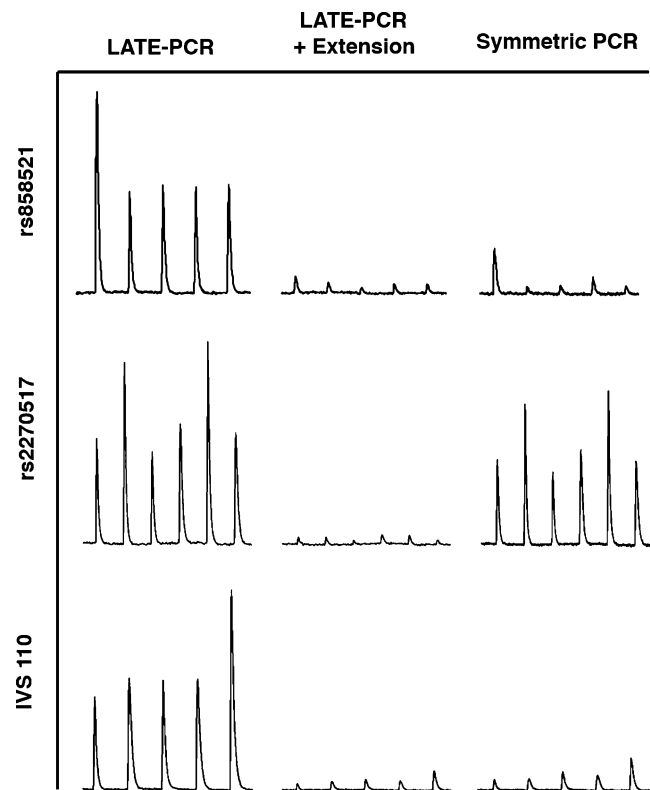


Fig. 1. Pyrosequencing results from different methods of PCR amplification. Plots from the same amplicon within each row are normalized to the same relative light unit (*y* axis) scale.

Table 1
Average signal base incorporation peak heights from different methods of PCR amplification

Amplicon	LATE-PCR (primarily ssDNA)	Double-strand-converted LATE-PCR (dsDNA)	Symmetric PCR (dsDNA)
rs858521	110.5	9.7 (0.09)	11.1 (0.10)
rs2270517	120.0	8.7 (0.07)	92.9 (0.77)
IVS 110	183.6	14.8 (0.08)	24.5 (0.13)

Note. Values are given in relative light units as defined by the pyrosequencing software.

the LATE-PCRs. The symmetric amplification of the rs2270517 amplicon yielded 10-fold better signal than the corresponding converted LATE-PCR, indicating that approximately 10 times more symmetric product than LATE-PCR product was made. In all three cases, however, the unmodified LATE-PCR products exhibited higher overall signals than the symmetric PCR products. Only one of three symmetric amplifications (rs2270517) produced strong enough pyrosequencing signals to be used for accurate genotyping, whereas this was the case for all three LATE-PCR amplifications. We conclude that although more template typically is generated by symmetric PCR than by LATE-PCR, less target strand is available in symmetric PCR to hybridize to the sequencing primer due to reannealing of the two complementary strands.

The impact of complementary strand competition can be quantified by comparison of the LATE-PCR and

converted LATE-PCRs that contain the same number of pyrosequencing template strands but differ in the extent to which these are single-stranded or double-stranded. Assuming perfectly efficient primer extension during strand conversion, the converted samples contain entirely dsDNA, whereas the unconverted LATE-PCR samples contain primarily ssDNA with a small amount of dsDNA in the background. As seen in Table 1, the average pyrosequencing signal from converted dsDNA was 7 to 9% of that from the same amount of LATE-PCR product. Thus, on a per-molecule level, LATE-PCR products give a signal that is up to 14-fold stronger than an equivalent number of double-stranded amplicons. This difference may be even greater given that the efficiency of converting ssDNA to dsDNA may, in fact, have been less than 100% under our experimental conditions.

LATE-PCR product cleanup

Raw PCR products are not generally used for pyrosequencing because by-products of the amplification reaction, most significantly unused primers, deoxynucleoside triphosphates, pyrophosphates, and buffers, interfere with the pyrosequencing reaction. Although traditional biotin/streptavidin separation or other cleanup methods, such as DNA binding columns and ethanol precipitation, can effectively remove these leftover reaction components, they are not conducive to high-throughput applications. Therefore, we sought alternative sample preparation methods that could be accomplished efficiently using single-tube, liquid-phase chemistry.

Residual primers in symmetric PCRs need to be either removed or blocked before pyrosequencing because they can be extended by the Klenow DNA polymerase and can confuse sequence determination. In a LATE-PCR amplification, this is unnecessary because none of the limiting primer and only a fraction of the excess primer remain unincorporated at the end of the reaction. If properly designed, the excess primer can prime only at the 3' end of the limiting primer strand. The extent to which this occurs is minimal because the limiting primer strand is significantly less abundant (generally \leq one-tenth) than the excess primer strand and also is bound up in a double-stranded form where priming is inefficient. In fact, in accord with the results described above, no signal above background is observed due to excess primer extension along the limiting primer strand in control reactions.

Pyrophosphate and dNTPs left over from PCR also prevent accurate sequence determination because they are metabolized by the pyrosequencing enzymes, resulting in indiscriminant light emission. In addition, dNTPs can be incorporated at the end of the sequencing primer in an unregulated manner and can desynchronize templates for subsequent base additions. Previous investigators have reported the combined use of potato apyrase and yeast pyrophosphatase followed by heat denaturation to effectively break down dNTPs and PPi in PCR products prior

to pyrosequencing [24–26]. In our hands, apyrase and pyrophosphatase were found to efficiently metabolize dNTPs and PPi, respectively, but only apyrase could be adequately inactivated by heating. Pyrophosphatase-treated PCRs retained substantial PPi degrading activity even after 10 min at 100 °C. Consistent with other reports [27], this residual activity was found to negatively affect the accuracy, reproducibility, and signal strength in pyrosequencing reactions (data not shown). As an alternative approach for pyrophosphate digestion, a newly available alkaline phosphatase derived from the Antarctic strain TAB5 [38] (New England Biolabs, Ipswich, MA, USA) was tested. Although fully heat labile, the TAB5 pyrophosphatase activity reported by the manufacturer was found to be insufficient for rapid PCR product cleanup (data not shown).

The most effective enzymatic method for pyrophosphate removal was found to be conversion of PPi to ATP using ATP sulfurylase and APS followed by digestion with apyrase. A convenient source of these reagents is the pyrosequencing enzyme and substrate mixes themselves. The luciferase and luciferin contained in these solutions provide a useful system for monitoring the breakdown of PPi as well as dNTPs. Because both ATP and dATP serve as substrates for luciferase, termination of sample light output serves as a good indicator for cleanup completion. Although the optimal pH for pyrosequencing chemistry is 7.6 and the pH of most PCR products is greater than 8.0, it was found that no adjustment was needed when mixing equal volumes of standard pyrosequencing reaction buffer and PCR products.

Breakdown and monitoring of PPi and dNTPs through the ATP sulfurylase and luciferase pathways requires both APS and luciferin. Different amplicons have different requirements for these reagents during cleanup due to varying amounts of leftover PPi and dNTP after PCR caused by differences in PCR product length, reaction efficiency, and other factors in individual reactions. For the three LATE-PCR amplicons described here, an amount of pyrosequencing substrate mix equivalent to twice that used in a standard pyrosequencing reaction was found to be adequate for enzymatic cleanup. Insufficient APS and/or luciferin during cleanup resulted in nonspecific signal or premature reaction termination during subsequent sequencing (data not shown).

A general concern of enzyme-based PCR cleanup approaches for pyrosequencing is the overproduction of breakdown by-products that may lead to feedback inhibition of enzymes during later sequencing and may reduce the number of bases that can be sequenced. These include SO_4^{2-} , oxyluciferin, inorganic phosphate (Pi), dNMPs, and AMP. One way of limiting the pool of Pi and dNMPs is to reduce the concentration of dNTPs used in PCR. Quantitative analysis of the yield of ssDNA LATE-PCR amplicons up to 600 bases long (unpublished observations) has revealed that dNTP concentrations can routinely be lowered to 100 μM without affecting amplification

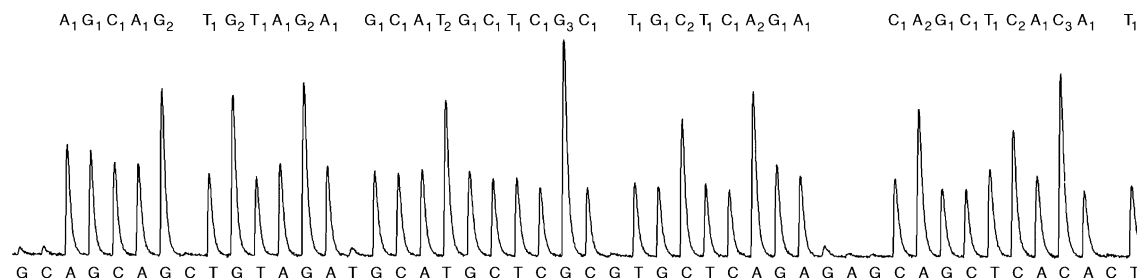


Fig. 2. Pyrogram of the LATE-PCR rs2270517 amplicon in a reaction carried out for more than 50 bases. Nucleotide dispensation order is listed below each peak, with the expected sequence indicated above.

efficiency. Under these conditions, pyrosequencing for more than 50 consecutive bases can be carried out on enzymatically prepared LATE-PCR samples (Fig. 2).

Adaptation of a traditional pyrosequencing SNP assay to a LATE-PCR format

The rs858521 locus is one of several hundred SNP sites on human chromosome 17 in close proximity to the p53 tumor suppressor gene. When heterozygous, these sites provide useful markers to detect allelic loss in neoplastic tissue [19], a clinically useful biomarker for cancer development [39]. To efficiently genotype individuals at the rs858521 locus, a traditional strand separation pyrosequencing assay was adapted to a LATE-PCR format. Symmetric PCR primers were modified to meet LATE-PCR primer design criteria as described in Materials and methods and Table 2. The biotin-labeled symmetric PCR primer was modified to become the LATE-PCR excess primer and thus maintain the same pyrosequencing template strand. This allowed retention of the sequencing primer and base dispensation order from the original symmetric PCR-based assay. Primer adjustments were made at the 5' ends to minimize the introduction of new 3' mispriming possibilities. The underlined sequences in Table 2 indicate added bases. Fig. 3 shows the pyrograms from LATE-PCR-amplified, enzyme-prepared samples obtained from individuals with each of the three possible rs858521 genotypes. Sample cleanup was completed in less than 1 min.

LATE-PCR-through-pyrosequencing on single cells

An A-to-G mutation at nucleotide 110 of the human β -globin gene is the most frequent cause of β -thalassemia in Eastern Mediterranean populations [40]. Because the lives faced by those afflicted with the disease are short, painful, and a heavy burden on health care systems, reliable and inexpensive techniques for preimplantation and prenatal diagnostics are extremely desirable. To genotype single cells at the IVS 110 site, a robust, PCR-compatible lysis buffer was used to prepare single lymphoblasts in the same tube as used for subsequent LATE-PCR amplification. Using this one-tube method, we previously reported single-cell amplification allele dropout rates of less than 4%

[30]. LATE-PCR-amplified material was then enzymatically prepared as above and pyrosequenced. Pyrograms corresponding to the complementary strand of the mutation obtained from homozygous wild-type, homozygous mutant, and heterozygous single cells are presented in Fig. 4.

Discussion

As more is learned about the role of genetic variants in complex diseases and therapeutic responses in the emerging age of genomics-based medicine, robust and efficient platforms for routine genetic screening will be increasingly needed. Pyrosequencing has emerged over the past decade as a powerful new tool for quantitative genotyping, although this has remained largely a research platform. For realistic use in a clinical or field setting, genetic assays will need to be not only reliable but also high throughput, low cost, automatable, and limited in transfer steps that risk contamination of a facility with PCR products. Here we have presented an integrated method for sample preparation and amplification of single-stranded template for pyrosequencing that improves on each of these features.

The sample throughput rate of commercially available pyrosequencing assays currently is limited by a relatively labor-intensive sample preparation step where single-stranded template is prepared from double-stranded PCR products. In contrast to the standard multitube template preparation method, the LATE-PCR-based approach eliminates the need for time-consuming solid-phase workup and can be accomplished in a single reaction vessel without expensive biotinylated primers, streptavidin beads, or purification equipment. A reduced need for material handling decreases the sample-to-sequence time of an assay and reduces the risk of laboratory contamination.

Our LATE-PCR-based approach proved to be significantly more efficient than previous one-tube methods for simplifying sample preparation through the use of dsDNA templates. In a side-by-side comparison of pyrosequencing results obtained using different PCR methods, dsDNA produced less than one-tenth the average signal of predominantly single-stranded LATE-PCR products. Furthermore, only one of three symmetric PCR amplicons yielded sufficient

Table 2
Modification of standard rs858521 PCR primers to meet LATE-PCR criteria

Primer	Sequence	Estimated T_m ^a
Standard forward	5'-CCAGCCGGTGCATTTTC-3'	52.8
Standard reverse	5' [Biotin]-CCTTGACCTGTTGTGGAGAGAA-3'	<u>58.5</u>
		$\Delta T_m = -5.7$
LATE-PCR forward (limiting)	5'- <u>TCCCCAGAGCCC</u> AGCCGGTGCATTTTC-3'	66.5
LATE-PCR reverse (excess)	5'- <u>CAATCCCTTGACCTGTTGTGGAGAGAA</u> -3'	<u>61.8</u>
		$\Delta T_m = +4.7$

^a Estimated T_m values for all primers are based on the asymmetric concentrations used in LATE-PCR (in this case, 100 nM for the forward and 1000 nM for the reverse).

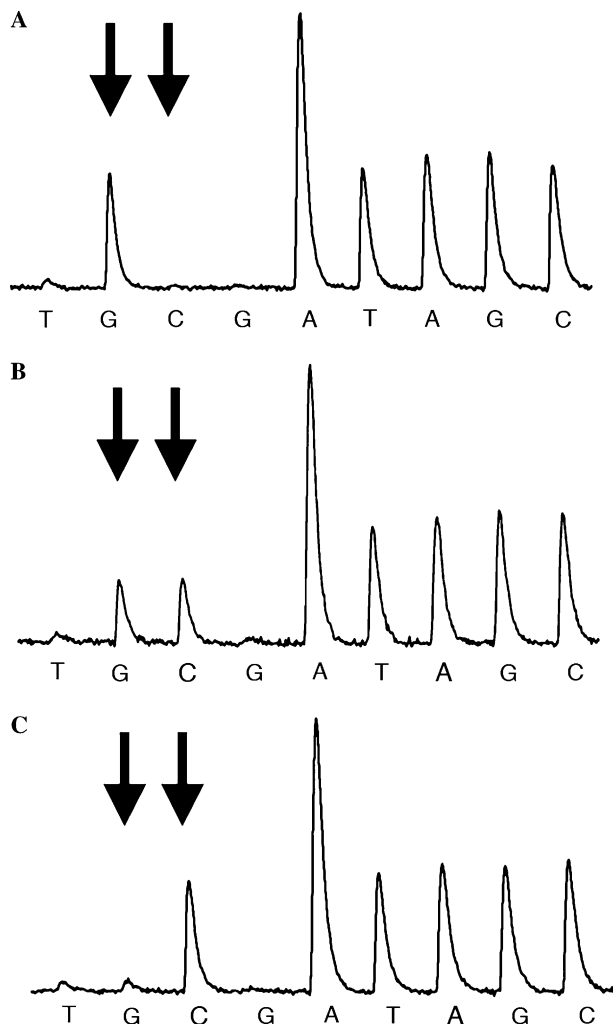


Fig. 3. Pyrograms from a LATE-PCR version of the rs858521 amplicon. (A–C) were generated from patient samples with each of the three possible rs858521 SNP genotypes (GG, GC, CC) as indicated by arrows.

signal to be used reliably in a genotyping assay. In contrast, all three sequences gave robust signals when amplified in a LATE-PCR format. In addition, it is necessary to remove or block residual primers when pyrosequencing standard PCR products, whereas this is unnecessary when using LATE-PCR material. Previously reported methods of non-separation-based primer removal for pyrosequencing have included exonuclease digestion [24,25] and blockage of primers or primer binding sites with 3' phosphorylated comple-

mentary sequences [26]. In the first method, the limited catalytic activity of exonuclease I necessitates a 20-min incubation step, and the latter approach adds complexity to initial optimization as well as oligonucleotide cost. LATE-PCR products, on the other hand, were able to be prepared in less than 1 min using only the pyrosequencing reagent mixes.

It is not difficult to envision a pyrosequencing machine with expanded fluid-handling capabilities and an integrated thermocycler where sample preparation, amplification, primer annealing, and sequencing could be fully automated using one-tube LATE-PCR chemistry. Perhaps the best format would be in a microfluidic or mesofluidic device where many parallel assays could be rapidly accomplished on a single throwaway wafer or cassette. A limited amount of work has already been done using pyrosequencing chemistry on bead-immobilized DNA in microfabricated chips [41–43]. Although this technique of micro flow-through pyrosequencing is appealing because it flushes away reaction by-products and prevents reaction dilution, enzymes need to be continually replenished because they are also dispersed at each cycle and light signal can be delocalized. A strategy for immobilizing pyrosequencing enzymes by fusion with DNA binding proteins, so far published only at the macroscale level [44], offers a potential solution but at the expense of signal strength. Proprietary technology owned by 454 Life Sciences [45] uses an elegant method of flow-through pyrosequencing on a massively parallel scale where both target DNA and enzymes are bead immobilized, but it is structured for genome-wide sequencing rather than targeted sequencing. In instances where low-cost, region-specific sequencing is needed, such as in many diagnostic, forensic, and biodetection applications, a single-chamber, liquid-phase chemistry may be easiest to incorporate into an automated and cost-effective device.

In summary, we have demonstrated an improved method of preparing and amplifying genetic material for pyrosequencing using LATE-PCR. The simple, reduced-step approach decreases the time, cost, and labor currently associated with pyrosequencing, yet it does not add complexity to assay design. With only simple modifications of PCR primers, existing assays are easily converted to a LATE-PCR form. The unique dynamics of LATE-PCR enables a single-chamber, liquid-phase integration of all steps needed for pyrosequencing and presents a novel tool for developing future high-throughput genetic assays and equipment.

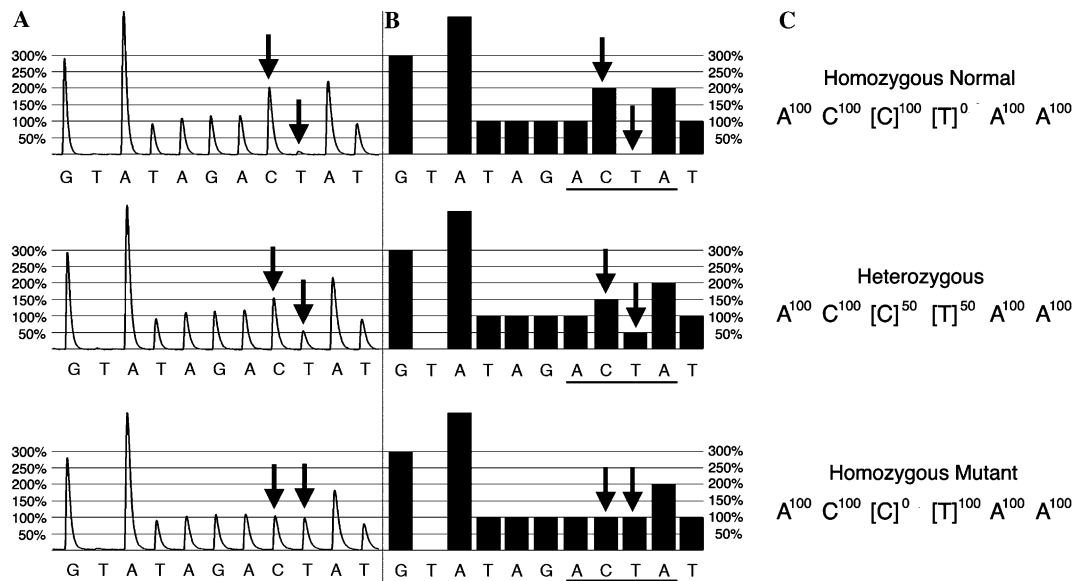


Fig. 4. Pyrograms generated from single-cell LATE-PCR amplifications of a portion of the human β -globin gene. (A) Top to bottom: Pyrograms from cells homozygous mutant (GG), heterozygous (AG), and homozygous normal (AA) at the IVS 110 site. The arrows point to the peaks reflecting the polymorphic site. In this experiment, the sense strand was sequenced, resulting in peaks of a complementary genotype (CC, CT, and TT, respectively). (B) An idealized program for each genotype indicating the relative heights of peaks in (A). (C) Sequences underlined in (B), with superscripts corresponding to fractions of nucleotide positions in the samples. For interpretation of (C), it is important to take into account that the nonvarying C base immediately flanking C polymorphism contributes to the height of the C allele peak. Thus, for the homozygous normal sample with 100% C allele and 0% T allele, the addition of the adjacent nonvarying C base makes the C peak 200%. For the heterozygous genotype with 50% C allele and 50% T allele, the C peak height becomes 150% and the T peak becomes 50%. For the homozygous mutant composed of 0% C allele and 100% T allele, the addition of the nonvarying C base makes the C peak 100%. In this sample, the T peak is also 100%.

Acknowledgments

We kindly recognize help provided by Karen Eldin, Ruth Ann Luna, and James Versalovic at Texas Children's Hospital (Houston, TX, USA) as well as Rene Myers and Liying Yan at Biotage (Westborough, MA, USA). Funding for this project was provided by Brandeis University, with additional support being provided by NIH PO1 CA91955 and NIH P50 HG02360.

References

- [1] M. Ronaghi, S. Karamohamed, B. Pettersson, M. Uhlen, P. Nyrén, Real-time DNA sequencing using detection of pyrophosphate release, *Anal. Biochem.* 242 (1996) 84–89.
- [2] M. Ronaghi, M. Uhlen, P. Nyrén, A sequencing method based on real-time pyrophosphate, *Science* 281 (1998) 363–365.
- [3] M. Ronaghi, Improved performance of pyrosequencing using single-stranded DNA-binding protein, *Anal. Biochem.* 286 (2000) 282–288.
- [4] B. Gharizadeh, T. Nordström, A. Ahmadian, M. Ronaghi, P. Nyrén, Long-read pyrosequencing using pure 2'-deoxyadenosine-5'-O'-(1-thiotriphosphate) Sp-isomer, *Anal. Biochem.* 301 (2002) 82–90.
- [5] A. Svantesson, P.O. Westermark, J.H. Kotaleski, B. Gharizadeh, A. Lansner, P. Nyrén, A mathematical model of the pyrosequencing reaction system, *Biophys. Chem.* 110 (2004) 129–145.
- [6] A. Agah, M. Aghajan, F. Mashayekhi, S. Amini, R.W. Davis, J.D. Plummer, M. Ronaghi, P.B. Griffin, A multi-enzyme model for pyrosequencing, *Nucleic Acids Res.* 32 (2004) e166.
- [7] N. Pourmand, E. Elahi, R.W. Davis, M. Ronaghi, Multiplex pyrosequencing, *Nucleic Acids Res.* 30 (2002) e31.
- [8] J. Dunker, U. Larsson, D. Petersson, J. Forsell, A.L. Schiller, A. Alderborn, L.M. Berg, Parallel DNA template preparation using a vacuum filtration sample transfer device, *BioTechniques* 34 (2003) 862–868.
- [9] H. Monstein, S. Nikpour-Badr, J. Jonasson, Rapid molecular identification and subtyping of *Helicobacter pylori* by pyrosequencing of the 16S rDNA variable V1 and V3 regions, *FEMS Microbiol. Lett.* 199 (2001) 103–107.
- [10] B. Gharizadeh, M. Kalantari, A.C. Garcia, B. Johansson, P. Nyrén, Typing of human papillomavirus by pyrosequencing, *Lab. Invest.* 81 (2001) 673–679.
- [11] B. Gharizadeh, E. Norberg, J. Loffler, S. Jalal, J. Tollemar, H. Einsele, L. Klingspor, P. Nyrén, Identification of medically important fungi by the pyrosequencing technology, *Mycoses* 47 (2004) 29–33.
- [12] S. Hjalmarsson, A. Alderborn, C. Fock, I. Muldin, H. Kling, M. Uhlen, L. Engstrand, Rapid combined characterization of microorganism and host genotypes using a single technology, *Helicobacter* 9 (2004) 138–145.
- [13] C.A. Garcia, A. Ahmadian, B. Gharizadeh, J. Lundeberg, M. Ronaghi, P. Nyrén, Mutation detection by pyrosequencing: sequencing of exons 5–8 of the p53 tumor suppressor gene, *Gene* 253 (2000) 249–257.
- [14] A. Goriely, G.A. McVean, M. Rojmyr, B. Ingemarsson, A.O. Wilkie, Evidence for selective advantage of pathogenic FGFR2 mutations in the male germ line, *Science* 301 (2003) 643–646.
- [15] A. Ahmadian, B. Gharizadeh, A.C. Gustafsson, F. Sterky, P. Nyrén, M. Uhlen, J. Lundeberg, Single-nucleotide polymorphism analysis by pyrosequencing, *Anal. Biochem.* 280 (2000) 103–110.
- [16] A. Alderborn, A. Kristofferson, U. Hammerling, Determination of single-nucleotide polymorphisms by real-time pyrophosphate DNA sequencing, *Genome Res.* 10 (2000) 1249–1258.
- [17] J.D. Gruber, P.B. Colligan, J.K. Wolford, Estimation of single nucleotide polymorphism allele frequency in DNA pools using pyrosequencing, *Hum. Genet.* 110 (2002) 395–401.
- [18] B. Neve, P. Froguel, L. Corset, E. Vaillant, V. Vatin, P. Boutin, Rapid SNP allele frequency determination in genomic DNA pools by pyrosequencing, *BioTechniques* 32 (2002) 1138–1142.

- [19] T. Tuziak, J. Jeong, T. Majewski, M.S. Kim, J. Steinberg, Z. Wang, D.S. Yoon, T.C. Kuang, K. Baggerly, D. Johnston, B. Czerniak, High-resolution whole-organ mapping with SNPs and its significance to early events of carcinogenesis, *Lab. Invest.* 85 (2005) 689–701.
- [20] E. Söderbäck, A.L. Zackrisson, B. Lindblom, A. Alderborn, Determination of CYP2D6 gene copy number by pyrosequencing, *Clin. Chem.* 51 (2005) 522–531.
- [21] J. Tost, J. Dunker, I.G. Gut, Analysis and quantification of multiple methylation variable positions in CpG islands by pyrosequencing, *BioTechniques* 35 (2003) 152–156.
- [22] S. Colella, L. Shen, K.A. Baggerly, J.P. Issa, R. Krahe, Sensitive and quantitative universal pyrosequencing methylation analysis of CpG sites, *BioTechniques* 35 (2003) 146–150.
- [23] J.M. Dupont, J. Tost, H. Jammes, I.G. Gut, De novo quantitative bisulfite sequencing using the pyrosequencing technology, *Anal. Biochem.* 333 (2004) 119–127.
- [24] T. Nordström, K. Nourizad, M. Ronaghi, P. Nyrén, Method enabling pyrosequencing on double-stranded DNA, *Anal. Biochem.* 282 (2000) 186–193.
- [25] T. Nordström, M. Ronaghi, L. Forsberg, U. de Faire, R. Morgenstern, P. Nyrén, Direct analysis of single-nucleotide polymorphism on double-stranded DNA by pyrosequencing, *Biotechnol. Appl. Biochem.* 31 (2000) 107–112.
- [26] T. Nordström, A. Alderborn, P. Nyrén, Method for one-step preparation of double-stranded DNA template applicable for use with pyrosequencing technology, *J. Biochem. Biophys. Methods* 52 (2002) 71–82.
- [27] T. Pacey-Miller, R. Henry, Single-nucleotide polymorphism detection in plants using a single-stranded pyrosequencing protocol with a universal biotinylated primer, *Anal. Biochem.* 317 (2003) 166–170.
- [28] M.A. Diggle, S.C. Clarke, A novel method for preparing single-stranded DNA for pyrosequencing, *Mol. Biotechnol.* 24 (2003) 221–224.
- [29] J.A. Sanchez, K.E. Pierce, J.E. Rice, L.J. Wangh, Linear-after-the-exponential (LATE)-PCR: an advanced method of asymmetric PCR and its uses in quantitative real-time analysis, *Proc. Natl. Acad. Sci. USA* 101 (2004) 1933–1938.
- [30] K.E. Pierce, J.E. Rice, J.A. Sanchez, L.J. Wangh, Detection of cystic fibrosis alleles from single cells using molecular beacons and a novel method of asymmetric real-time PCR, *Mol. Hum. Reprod.* 9 (2003) 815–820.
- [31] U.B. Gyllensten, H.A. Erlich, Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus, *Proc. Natl. Acad. Sci. USA* 85 (1988) 7652–7656.
- [32] K.E. Pierce, J.A. Sanchez, J.E. Rice, L.J. Wangh, Linear-after-the-exponential (LATE)-PCR: primer design criteria for high yields of specific single-stranded DNA and improved real-time detection, *Proc. Natl. Acad. Sci. USA* 102 (2005) 8609–8614.
- [33] K.E. Pierce, J.E. Rice, J.A. Sanchez, C. Brenner, L.J. Wangh, Real time PCR using molecular beacons for accurate detection of the Y chromosome in single human blastomeres, *Mol. Hum. Reprod.* 6 (2000) 1155–1164.
- [34] K.E. Pierce, J.E. Rice, J.A. Sanchez, L.J. Wangh, QuantiLyse: reliable DNA amplification from single cells, *BioTechniques* 32 (2002) 1106–1111.
- [35] T.G. Paulson, P.C. Galipeau, B.J. Reid, Loss of heterozygosity analysis using whole genome amplification, cell sorting, and fluorescence-based PCR, *Genome Res.* 9 (1999) 482–491.
- [36] H.T. Allawi, J. SantaLucia, Thermodynamics and NMR of internal G–T mismatches in DNA, *Biochemistry* 36 (1997) 10581–10594.
- [37] R. Owczarzy, Y. You, B.G. Moreira, J.A. Manthey, L. Huang, M.A. Behlke, J.A. Walder, Effects of sodium ions on DNA duplex oligomers: improved predictions of melting temperatures, *Biochemistry* 43 (2004) 3537–3554.
- [38] M. Rina, C. Pozidis, K. Mavromatis, M. Tzanodaskalaki, M. Kokkinidis, V. Bouriotis, Alkaline phosphatase from the Antarctic strain TAB5: properties and psychrophilic adaptations, *Eur. J. Biochem.* 267 (2000) 1230–1238.
- [39] B.J. Reid, L.J. Prevo, P.C. Galipeau, C.A. Sanchez, G. Longton, D.S. Levine, P.L. Blount, P.S. Rabinovitch, Predictors of progression in Barrett's esophagus II: baseline 17p (p53) loss of heterozygosity identifies a patient subset at increased risk for neoplastic progression, *Am. J. Gastroenterol.* 96 (2001) 2839–2848.
- [40] J.M. Old, Prenatal diagnosis of the hemoglobinopathies, in: A. Milunsky (Ed.), *Genetic Disorders and the Fetus: Diagnosis, Prevention, and Treatment*, Johns Hopkins University Press, Baltimore, MD, 2004, pp. 663–699.
- [41] A. Russom, N. Tooke, H. Andersson, G. Stemme, Single nucleotide polymorphism analysis by allele-specific primer extension with real-time bioluminescence detection in a microfluidic device, *J. Chromatogr. A* 1014 (2003) 37–45.
- [42] H. Andersson, W. van der Wijngaart, G. Stemme, Micromachined filter-chamber array with passive valves for biochemical assays on beads, *Electrophoresis* 22 (2001) 249–257.
- [43] M. Ehn, N. Nourizad, K. Bergstrom, A. Ahmadian, P. Nyrén, J. Lundeberg, S. Hober, Toward pyrosequencing on surface-attached genetic material by use of DNA-binding luciferase fusion proteins, *Anal. Biochem.* 329 (2004) 11–20.
- [44] A. Russom, N. Tooke, H. Andersson, G. Stemme, Pyrosequencing in a microfluidic flow-through device, *Anal. Chem.* 77 (2005) 7505–7511.
- [45] M. Margulies, M. Egholm, W.E. Altman, S. Attiya, J.S. Bader, L.A. Bemben, et al., Genome sequencing in microfabricated high-density picolitre reactors, *Nature* 437 (2005) 376–380.