Published in final edited form as:

Electrophoresis. 2010 April ; 31(8): 1322–1329. doi:10.1002/elps.200900631.

Template tailoring: Accurate determination of heterozygous alleles using peptide nucleic acid and dideoxyNTP

Muhammad Akram Tariq¹ and Nader Pourmand^{1,2}

¹Biomolecular Engineering, University of California Santa Cruz, Santa Cruz, CA, USA

²Stanford Genome Technology Center, Stanford University, Palo Alto, CA, USA

Abstract

Measurement of the length of DNA fragments plays a pivotal role in genetic mapping, disease diagnostics, human identification and forensic applications. PCR followed by electrophoresis is used for DNA length measurement of STRs, a process that requires labeled primers and allelic ladders as standards to avoid machine error. Sequencing-based approaches can be used for STR analysis to eliminate the requirement of labeled primers and allelic ladder. However, the limiting factor with this approach is unsynchronized polymerization in heterozygous sample analysis, in which alleles with different lengths can lead to imbalanced heterozygote peak height ratios. We have developed a rapid DNA length measurement method using peptide nucleic acid and dideoxy dNTPs to “tailor” DNA templates for accurate sequencing to overcome this hurdle. We also devised an accelerated “dyad” pyrosequencing strategy, such that the combined approach can be used as a faster, more accurate alternative to *de novo* sequencing. Dyad sequencing interrogates two bases at a time by allowing the polymerase to incorporate two nucleotides to DNA template, cutting the analysis time in half. In addition, for the first time, we show the effect of peptide nucleic acid as a blocking probe to stop polymerization, which is essential to analyze the heterozygous samples by sequencing. This approach provides a new platform for rapid and cost-effective DNA length measurement for STRs and resequencing of small DNA fragments.

Keywords

Dyad pyrosequencing; Peptide nucleic acid; Pyrosequencing; STRs; Template tailoring

1 Introduction

DNA sequencing has revolutionized bioscience and its use has been critical for a number of important medical discoveries. Technological advances now allow many groups to conduct whole genome sequencing both for comparisons between species, as well as for differentiating individuals, a process that is useful in forensics and which also forms the foundation of personalized medicine. One of the critical issues in DNA sequencing of whole genomes is the problem of relatively short lengths of DNA-reads using most approaches. Short read-lengths complicate the informatics needed to analyze and place those sequences in the context of the whole genome. In the realm of forensic DNA analysis and mapping of genes implicated in diseases, DNA length determinations of STRs are performed using PCR fragment size

measurements. This approach is based on an electrophoretic technique, and involves the use of dye labeled primers, requires very careful data analysis and is limited by the occurrence of technology artifacts [1].

Sequencing technologies have evolved over the years. One of the first of these, the Sanger method, has provided an elegant sequencing method, which was widely used for the last three decades [2]. A more recently developed technology, pyrosequencing, is an alternative method for sequencing DNA fragments with a number of advantages [3]. Pyrosequencing is a sequencing-by-synthesis method that monitors the polymerase activity coupled to two enzymes to generate a detectable light response: (i) ATP-Sulfurylase converts inorganic pyrophosphate to ATP generated by polymerase during nucleotide incorporation and (ii) luciferase that uses ATP as a source of energy to generate light [4]. Pyrosequencing technology has been used in a broad range of applications such as genotyping of microbes, SNP genotyping, mutation detection and gene identifications [5]. For DNA length measurements, our group has recently develop a branch migration assay which is a microarray based technique that requires 3-step hybridization [6].

Most recently, high-throughput sequencing systems have become available to reduce DNA sequencing costs by massively parallel DNA sequencing [7,8]. To use a massively parallel sequencing system for human identification, sequencing can be performed using the same, and possibly more, loci used for current STR analysis. Notably, the accuracy requirements for these analyses are extremely high, given their use in the legal system [9]. However, because STR regions are repetitive by nature, they are among the most difficult sample types to analyze properly by these high-throughput systems due to short read lengths (and possible dephasing in the GS-FLX 454 life sequencers). Even with high degrees of coverage, assembly of these highly repetitive regions may prove very difficult [10]. On the practical front, there is risk of cross contamination from running multiple individuals simultaneously on a single instrument. However, if multiple samples could not be run simultaneously, the minimum cost *per* run and time *per* run for these high-throughput sequencers would be prohibitive [9].

Thus, despite the increasing availability of high-throughput sequencing systems, low throughput systems remain important for STR analysis. Of the other recent technologies, pyrosequencing has the potential to provide a robust method for DNA length measurement, which is time and cost competitive compared to other approaches. It is automatable and provides easily interpretable results. In pyrosequencing, there is no need for specific dye labeling of PCR products [11]. The technique has been shown to be capable of determining the sequence variants within or near repeat regions in STRs in addition to fragment length differences. This feature is particularly useful in avoiding confusion during interpretation of results for human identity testing or relationship testing [12].

Importantly, the fragment length of a particular STR for a homozygous individual can be determined by using pyrosequencing. However, measurement of STR length for a heterozygous individual has been a challenge. Here, we introduce a new technique for determining the length of heterozygous DNA called “template tailoring,” in which we use peptide nucleic acid (PNA) terminal probes to accurately determine differing allele lengths. We also debut a time-saving method, “dyad pyrosequencing,” in which two bases are dispensed and interrogated in each cycle. Together these methods provide an accurate and efficient approach for DNA profiling of STRs.

Critical to this new approach are PNA probes designed to hybridize to a region flanking the STR where the pyrosequencing is tailored to end. PNA is a nucleobase oligomer in which the entire backbone has been replaced by *N*-(2-aminoethyl) glycine units. PNA is able to recognize specific sequences of DNA, obey the Watson–Crick hydrogen bonding scheme and the hybrid

complexes exhibit extraordinary thermal stability and unique ionic strength effects [13–15]. PNA has been previously reported to be capable of inhibiting transcription as well as translation [16]. We report here the ability of such PNAs to block the polymerase from incorporating further dNTPs to a growing DNA strand, in such a way that unmatched alleles do not interfere with one another as templates.

In dyad pyrosequencing, we dispense two independent nucleotides at once to interrogate two bases at a time. The interrogation of two nucleotides at a time reduces the sequencing time to half. This allows for real-time DNA length measurements in the case of both homozygous and heterozygous samples. We used the dyad pyrosequencing approach with PNA terminal probes to analyze two STRs (GATA172D05 and GATA31E08) in 15 human DNA Samples. The available data suggest that the method described here is a reliable, rapid, information-rich and cost-effective approach for DNA length measurements and may be useful for genotyping STRs for both forensic and clinical applications.

2 Material and methods

2.1 Probe design and synthesis

Oligonucleotides for amplification of the markers GATA172D05, GATA31E08 and HPRTB were designed by using primer3 software (<http://primer3.sourceforge.net>) and were synthesized by IDT (Coralville, IA, USA). Either the forward or reverse primer was biotinylated. The sequences for oligonucleotides are shown in Table 1.

The 15-base-pairs sequence of the flanking region for each marker including GATA172D05, GATA31E08 and HPRTB were used to synthesize PNA probes in the following manner:

- i. GATA172D05: PNA sequence (N-terminus//---//C-terminus) of flanking region complementary to biotinylated strand of PCR product.
- ii. GATA31E08: PNA sequence (N-terminus//---//C-terminus), 4 base pairs of the repeat region and the remaining 11 base pairs of the flanking region complementary to biotinylated strand of PCR product.
- iii. HPRTB: PNA sequence (C-terminus//---//N-terminus) of flanking region complementary to biotinylated strand of PCR product.

PNA was synthesized and HPLC purified by Biosynthesis (Lewisville, TX, USA). PNA sequences are also shown in Table 1.

2.2 DNA samples

Fifteen blood samples were purchased from Stanford Blood Centre (Palo Alto, CA, USA). Total Genomic DNA (gDNA) was extracted using Qiagen's QiaAMP DNA Blood Maxi Kit (www.qiagen.com). Extractions were performed according to manufacturer's instructions. The quantity of DNA was determined by NanoDrop (Thermo Scientific, Wilmington, DE, USA) as *per* manufacturer's recommendation.

2.3 PCR amplification

Ten nanogram of gDNA was amplified in a single-plex PCR using biotin labeled primers in a 50- μ L reaction volume containing the following: 10 ng of genomic DNA, 0.2 μ M each forward and reverse primer for all three markers, 75 mM Tris-HCl (pH 8.0), 2.0 mM MgCl₂, 200 μ M of each dNTPs and 1.5 U of AmpliTaq Gold polymerase (Applied Biosystems, Foster City CA, USA). The amplification was performed in a Gene Amp PCR System 9700 Thermal Cycler (Applied Biosystems) under the following conditions: 95°C for 10 min, followed by 32 cycles

of denaturing at 95°C for 30 s, annealing at 57°C for 20 s, extension at 72°C for 12 s and a final extension at 72°C for 7 min.

2.4 Capillary electrophoresis

One microliter of each PCR amplified product was added to 12 μ L deionized formamide and 0.5 μ L GeneScan 500LIZ size standard (Applied Biosystems). The amplified products were separated using an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). Results were analyzed using Gene-Scan Analysis software 3.7. Genotyping was performed through comparison with DNA control reference sample 9947A (PowerPlex Y System, Promega, Madison, WI, USA), previously typed [17]. The homozygous samples for both markers GATA172D05 and GATA31E08 were also sequenced using ABI 3730 sequencer.

2.5 Sample preparation for pyrosequencing

The 50 μ L of biotinylated PCR product was immobilized onto streptavidin-coated super paramagnetic beads (Dynabeads M-280-streptavidin; Dynal AS, Oslo, Norway) by incubation at 42°C for 10 min. The immobilized PCR product was treated with 50 μ L of 20 mM NaOH for 5 min to obtain single-stranded DNA. The beads with single-stranded DNA were washed one time with 1 $\times$ annealing buffer (0.1 M Tris-acetate, pH 7.75, 200 mM magnesium acetate). The immobilized single-stranded DNA was resuspended into 20 μ L of 1 $\times$ annealing buffer along with 0.5 μ M of sequencing primer for each marker. One micromolar PNA oligonucleotide for GATA172D05 and 0.25 μ M PNA for each HPRTB and GATA31E08 were also mixed at this step as a polymerization-blocking probe. The sequencing primer and PNA were annealed to single-stranded template at 50°C for 5 min after 72°C for 1 min. The excess sequencing primer and PNA was washed after annealing step before running on pyrosequencer.

2.6 Dyad pyrosequencing

The pyrosequencing reaction was performed at 28°C in 40 μ L of reaction volume using an automated PSQ 96 system (Biotage, Uppsala, Sweden). The final volume of 40 μ L contained primed target DNA, 30 μ L of 1 $\times$ annealing buffer (0.1 M Tris-acetate, pH 7.75, 200 mM magnesium acetate), 5 μ L of standard Enzyme mix and 5 μ L of standard substrate mix provided by Biotage (www.biotage.com). The dGTP and dATP α S were mixed in equal volume for dispensing together to allow the polymerase to add two nucleotides at a time during sequencing. Likewise dTTP and dATP α S were also mixed in equal volume to dispense together to interrogate next two bases in template sequence using combinatorial approach. The terminating nucleotide (ddCTP) was purchased from USB (Cleveland, OH, USA) and used in the pyrosequencing reaction.

3 Results and discussion

3.1 The effect of dyad pyrosequencing

In the dyad pyrosequencing approach, we dispensed two nucleotides at once to incorporate two bases at a time as a repeat pattern for the known markers under investigation. As a result, we obtained a single peak with a twofold increase in emitted light signal, which denoted two bases. Because we combined two nucleotides and dispensed them together, this approach is referred as dyad pyrosequencing. This approach shortens the analysis time.

3.2 The effect of the addition of ddCTP and PNA blocking oligonucleotide

The sequencing process of STR continues until it reaches the flanking region of the shortest allele in the heterozygous sample. As above, we used PNA complementary to the flanking region as a terminal-blocking probe for polymerase. This stops the polymerization on the shorter allele while allowing polymerase processivity to continue on the longer allele in the

heterozygous genotype. Note that this results in a reduction of signal by half as polymerization is continued for the longer allele on half the number of copies of the PCR product. By implementing this strategy, we were able to use one common terminating nucleotide (ddCTP) for both markers (GATA172D05 and GATA31E08) along with PNA as proof-of-principle that sequencing of multiple STR markers is possible in one sequencing run. With this approach, flanking region signals do not interfere and resulting genotypes were interpretable as previously reported for autosomal STRs using dideoxynucleotides [18]. The Klenow fragment of polymerase has strong displacement activity for DNA oligonucleotides [19] but it stops polymerization when it reaches the flanking region in our case, failing to displace the PNA. This is not surprising, given the greater stability of the PNA–DNA duplexes over DNA–DNA duplexes as measured by T_m [16,20]. As a result, we are able to perform real-time DNA length measurement for both homozygous and heterozygous samples and analysis of multiple STR markers is possible in single runs of pyrosequencing.

3.3 Template tailoring for heterozygous samples

As shown in Fig. 1, the use of PNA terminal primers solves the problem of length determination in heterozygous samples. In Fig. 1A, we show that unsynchronized polymerization in alleles with different lengths results in uninterpretable sequences. Generally, a selected nucleotide that is not included in the repeat region of markers can be used as a terminating nucleotide to stop polymerization on the shortest allele as in, for example, GATA172D05. (Note that the repeat region sequence is TAGA or AGAT for markers included in this study and ddCTP is used as terminating nucleotide.) However, for a heterozygous sample, if the terminating nucleotide is incorporated a few bases away from the starting point of the flanking region, it results in confusion when assigning genotypes to heterozygous samples, *e.g.* GATA31E08 (Fig. 1B). Note that incorporation of ddCTP for marker GATA31E08 is at the ninth nucleotide as shown in the sequence and results in incorrect genotyping. The bases from the flanking region of the shorter allele until the incorporation of ddCTP interfere with repeat regions of the longer allele in pyrosequencing (Fig. 1B) and results in uninterpretable sequences. In Fig. 1C, PNA was used to overcome this issue in sequencing heterozygous samples and this resulted in interpretable genotypes, albeit with high background. To reduce the confounding background, the terminating dideoxynucleotide was used along with a PNA terminal probe that is designed to be complementary to the flanking region (Fig. 1D). In contrast to previous methods, PNA addition along with the common terminating nucleotide (ddCTP) made it possible for multiple markers to be sequenced in single run using the same terminating nucleotide. Note that the PNA terminal probe has 4 base pairs complementary to the repeat region as explained in the probe design section, so that exactly one repeat remains unsequenced. Therefore, the genotype for this heterozygous sample is properly reported as [7,12] instead of the apparent genotype [6,11] for this sample

3.4 Refinement of STR tailoring

We tested PNA alone or PNA with a terminating nucleotide for the ability to measure DNA length of two X-chromosome based STRs, GATA172D05 and GATA31E08. For these studies, the DNA length was determined in 15 DNA samples (nine homozygous and six heterozygous samples) for the marker GATA172D05 and 15 DNA samples (seven homozygous and eight heterozygous samples) for the marker GATA31E08.

We observed that using only PNA, it was possible to reduce signals for the flanking region and also incorporation of a terminating nucleotide (ddCTP) during the sequencing to further improve the outcome of the runs. We used PNA alone and PNA with ddCTP together for the marker GATA172D05 to generate pyrograms shown in Fig. 2. The DNA length for the marker GATA172D05 of homozygous individual samples was measured using only PNA as a blocking probe for polymerase (Fig. 2A). Note that the dispensation order of nucleotides is shown on

the bottom of the figure, while output sequences are written along with each peak (each peak represents two bases) and the full sequence of the marker is shown on the top of figure (for heterozygous samples; (A) represents the sequence of shorter allele and (B) represents the sequence of the longer allele) in each pyrogram shown in the section. The combined effect of PNA and terminating nucleotide for marker GATA172D05 of the homozygous sample is shown in Fig. 2B. The heterozygous sample of marker GATA172D05 was sequenced without PNA and ddCTP and, as expected, the results are non-interpretable (Fig. 2C). Thus, although the heterozygous sample of marker GATA172D05 was sequenced and a genotype was assigned by using only PNA (Fig. 2C), background signals from the flanking region still interfere with allele calling of the longer allele. The dyad dispensation of nucleotides increases the background signal due to the addition of many possible nucleotides. Therefore, we used ddCTP to further reduce the background signals. Hence, we suggest that full template tailoring using a combination of PNA and ddNTP is more useful in producing decisive genotypes. The combined effect of both PNA and ddCTP for heterozygous samples of marker GATA172D05 are shown in Fig. 2D. The PNA concentration of 1 μM is suitable for this marker as we found a significant decrease in signal from the blocked template at this concentration of PNA. No signal was detected in the flanking regions except for a minor signal from the first base “G,” which gives a peak height of about 20% of the last peak of the repeat region of the shorter allele as shown in Fig 2. Although this peak contributes to the last peak of the shorter allele, it is recognizably artifactual so that the genotypes remain interpretable. This artifact may be due to the “smiling” effect of pyrimidine-rich PNA as PNA–DNA duplex; having pyrimidine nucleobases in the PNA makes the duplex enthalpically disfavored in water (Table 1) [21]. Because the stability of PNA–DNA duplex is dependent on salt concentration, this artifact may be correctible by decreasing the concentration of cations in solution during sequencing [22].

For marker GATA31E08, the PNA oligonucleotide was 4 base pairs inside the repeat region to check whether Klenow may be able to displace four bases as previously reported [23]. In contrast to marker GATA172D05, a lower concentration of PNA (0.25 μM) was enough to block the polymerase for this marker. The Klenow was not able to remove four bases of the repeat region. The signal reduction was 67% on average for sequences annealed to PNA; again, this is likely to be due to the high stability of the PNA–DNA duplex [16,20]. The challenging heterozygous sample of marker GATA31E08 was sequenced and the genotype was assigned accurately by full template tailoring using ddCTP and PNA. The pyrograms of homozygous and heterozygous individual for marker GATA31E08 are shown in Fig. 3. We also used SYBR green in the reaction to further reduce signals of the flanking regions but it did not show any significant effect. As above, we added one extra repeat in reporting the genotype obtained by pyrosequencing to show the accurate genotypes for each homozygous or heterozygous sample to account for the fact that the PNA terminal probe was covering one repeat. To analyze both markers (GATA172D05 and GATA31E08) together in a single pyrosequencing run, we used a common terminating nucleotide (ddCTP). Here, we showed the advantage of combining PNA and ddCTP together to obtain reliable genotypes for marker GATA31E08 in unknown samples (Fig. 3).

To check the effect of the 5' functional group on enzyme processivity, we used a PNA oligonucleotide for marker HPRTB that was C-terminus to N-terminus, with the C-terminus facing the polymerase enzyme. This orientation of the PNA reduced the signals for the flanking region sequences, but results were not interpretable. We were not able to differentiate between homozygous and heterozygous samples (data not shown). We conclude that the N-terminus of PNA facing polymerase, N-terminus to C-terminus PNA, appeared to be the best terminal blocking probe for DNA length measurement using pyrosequencing.

The genotypes of GATA31E08 and GATA172D05 for all samples are shown in Table 2. The genotypes assigned to samples by dyad pyrosequencing were concordant with genotypes

obtained by capillary electrophoresis and Sanger sequencing for all samples including both markers GATA172D05 and GATA31E08.

In this work, we have demonstrated the feasibility of using a modified pyrosequencing method to measure the length of STRs in a rapid manner. This method employs existing pyrosequencing technology with two improvements. First, we used PNA as a terminal blocking probe to stop polymerization on the flanking region, which made it possible to use pyrosequencing for DNA length measurement in both homozygous and heterozygous samples, for varying repeat patterns of different markers. Second, we used a dyad approach for the addition of two nucleotides at a time by the polymerase which reduces the sample analysis time by half.

We observed that termination of the polymerization on the longer allele in heterozygous samples for analysis of multiple STR markers in a single reaction of pyrosequencing cannot be accomplished by the addition of a single dideoxynTP (*e.g.* ddCTP), due to the low probability of finding the same base at the end of repeat regions. Nor was the analysis of heterozygous samples optimal with PNA alone. PNA probes together with dideoxynTP provided a worthwhile strategy for STR analysis using pyrosequencing. This method can also be used for other applications of STR analysis such as evolutionary studies, linkage analysis for genetic diseases and cancer diagnostics.

Although both pyrosequencing and capillary electrophoresis allow multiple samples to be analyzed (in one plate or one run, respectively), pyrosequencing has several advantages over capillary electrophoresis and other traditional methods of DNA length measurement. First, there is no need for specific dye labeling of the PCR product in pyrosequencing. Second, there is no discrepancy of DNA length measured by pyrosequencing from run to run or machine to machine such as is found in capillary electrophoresis. Therefore, pyrosequencing eliminates the need for allelic ladders as a standard to measure the length of that particular STR.

On the other hand, pyrosequencing has limitations regarding multiplexity. Multiplex PCR amplification and multiple fragment analysis is of critical importance in forensic DNA testing due to the sometimes limited amount of DNA material obtainable from a crime scene. However, this limitation of pyrosequencing can be overcome by amplifying the multiple fragments in single PCR reaction using universal primers, after which the PCR product(s) can be sequenced using sequence-specific primers for particular markers in individual wells as previously described [24]. This dyad pyrosequencing method may be combined with multiplex amplification assays such as Molecular Inversion Probe [25], Padlock Probes [26], Connector Inversion Probe [27] and Trinucleotide Threading [28] for parallel amplification of multiple STR markers.

Having solved the problem of sequencing heterozygous samples, we believe that our PNA pyrosequencing method could be readily used for forensic DNA testing. The use of dyad interrogation/incorporation makes the method more attractive in practical terms, by reducing the analysis time.

Abbreviation

PNA peptide nucleic acid

Acknowledgments

This work was supported in part by grants from the National Institutes of Health (P01-HG000205), the National Science Foundation (DBI 0830141) and NASA NNN08ZNE002C.

References

1. Butler JM, Buel E, Crivellente F, McCord BR. Electrophoresis 2004;25:1397–1412. [PubMed: 15188225]
2. Sanger F, Nicklen S, Coulson AR. Proc. Natl. Acad. Sci. USA 1977;74:5463–5467. [PubMed: 271968]
3. Gharizadeh B, Herman ZS, Eason RG, Jejelowo O, Pourmand N. Electrophoresis 2006;27:3042–3047. [PubMed: 16800029]
4. Ronaghi M, Uhlen M, Nyren P. Science 1998;281:363–365. [PubMed: 9705713]
5. Ahmadian A, Ehn M, Hober S. Clin. Chim. Acta 2006;363:83–94. [PubMed: 16165119]
6. Pourmand N, Caramuta S, Villablanca A, Mori S, Karhanek M, Wang SX, Davis RW. Proc. Natl. Acad. Sci. USA 2007;104:6146–6151. [PubMed: 17389407]
7. Rogers YH, Venter JC. Nature 2005;437:326–327. [PubMed: 16163333]
8. Hutchison CA. Nucleic Acids Res 2007;35:6227–6237. [PubMed: 17855400]
9. Jobling MA, Gill P. Nat. Rev. Genet 2004;5:739–751. [PubMed: 15510165]
10. Hert DG, Fredlake CP, Barron AE. Electrophoresis 2008;29:4618–4626. [PubMed: 19053153]
11. Ronaghi M. Genome Res 2001;11:3–11. [PubMed: 11156611]
12. Huel RL, Basic L, Madacki-Todorovic K, Smajlovic L, Eminovic I, Berbic I, Milos A, Parsons TJ. Croat. Med. J 2007;48:494–502. [PubMed: 17696304]
13. Demidov VV. Trends Biotechnol 2003;21:4–7. [PubMed: 12480343]
14. Egholm M, Buchardt O, Christensen L, Behrens C, Freier SM, Driver DA, Berg RH, Kim SK, Norden B, Nielsen PE. Nature 1993;365:566–568. [PubMed: 7692304]
15. Pellestor F, Paulasova P, Hamamah S. Curr. Pharm. Des 2008;14:2439–2444. [PubMed: 18781993]
16. Ray A, Norden B. FASEB J 2000;14:1041–1060. [PubMed: 10834926]
17. Szibor R, Edelmann J, Hering S, Plate I, Wittig H, Roewer L, Wiegand P, Cali F, Romano V, Michael M. Forensic Sci. Int 2003;138:37–43. [PubMed: 14642717]
18. Styrman H, Divne AM, Nilsson M, Allen M. Int. Congress Ser 2006;1288:669–671.
19. Ronaghi M, Nygren M, Lundeberg J, Nyren P. Anal. Biochem 1999;267:65–71. [PubMed: 9918656]
20. Shakeel S, Karim S, Ali A. J. Chem. Technol. Biotechnol 2006;81:892–899.
21. Sen A, Nielsen PE. Biophys. J 2006;90:1329–1337. [PubMed: 16326919]
22. Tomac S, Sarkar M, Ratilainen T, Wittung P, Nielsen PE, Norden B, Graslund A. J. Am. Chem. Soc 1996;118:5544–5552.
23. Gharizadeh B, Akhras M, Nourizad N, Ghaderi M, Yasuda K, Nyren P, Pourmand N. J Biotechnol 2006;124:504–511. [PubMed: 16530870]
24. Gharizadeh B, Zheng B, Akhras M, Ghaderi M, Jejelowo O, Strander B, Nyrén P, Wallin KL, Pourmand N. Mol. Cell. Probes 2006;20:230–238. [PubMed: 16516439]
25. Hardenbol P, Baner J, Jain M, Nilsson M, Namsaraev EA, Karlin-Neumann GA, Fakhrai-Rad H, Ronaghi M, Willis TD, Landegren U, Davis RW. Nat. Biotechnol 2003;21:673–678. [PubMed: 12730666]
26. Thomas DC, Nardone GA, Randall SK. Arch. Pathol. Lab. Med 1999;123:1170–1176. [PubMed: 10583921]
27. Akhras MS, Unemo M, Thiagarajan S, Nyren P, Davis RW, Fire AZ, Pourmand N. PLoS One 2007;2:e915. [PubMed: 17878950]
28. Pettersson E, Lindskog M, Lundeberg J, Ahmadian A. Nucleic Acids Res 2006;34:e49. [PubMed: 16582098]

Figure 1.

Concept of DNA template tailoring. (A) Dyad pyrosequencing of GATA31E08 under standard conditions. As shown in panel A, it is not possible to interpret the sequence results due to the occurrence of unsynchronized polymerization of heterozygous individuals at this locus. In dyad pyrosequencing, each peak represents two bases (AG or AT) but the first peak in pyrogram, represents three bases (AGA) because third base (A) is also available for polymerization during dispensation of two nucleotides “AG” in the repeat motif of this marker (AGAT). (B) Dyad pyrosequencing of GATA31E08 with ddNTP. Addition of ddCTP does not produce accurate genotypes in heterozygous samples. (C) Partial DNA template tailoring of GATA31E08 using PNA alone. Dyad pyrosequencing using PNA to signal the ends of the separate strands produces recognizable genotypes. Note that the PNA probe includes 4 base pairs complementary to the repeat region in addition to the portion that is complementary to the flanking region. Therefore, the genotype for this heterozygous sample is properly reported as [7,12] instead of the apparent genotype 6, 11. (D) Full Template Tailoring of GATA31E08 using PNA with ddCTP. Use of PNA along with terminating nucleotides (in this case, ddCTP) reduces the signal-to-noise ratio and results in accurate determination of genotype by dyad pyrosequencing.

Figure 2.

DNA template tailoring by using PNA and ddCTP for homozygous and heterozygous samples for the marker GATA172D05. (A) Partial template tailoring of homozygous samples using PNA alone (genotype 6, 6). In dyad pyrosequencing, each peak represents two bases (TA or GA). (B) Full template tailoring of homozygous samples using PNA and ddCTP (genotype 6, 6). (C) Sequencing of heterozygous sample without template tailoring (non-interpretable genotypes). (D) Partial template tailoring of heterozygous sample using PNA alone (genotype 6, 12). (E) Full template tailoring of heterozygous samples using PNA and ddCTP (genotype 6, 12).

Figure 3.

Dyad pyrosequencing using PNA and ddCTP for homozygous and heterozygous samples for the marker GATA31E08. (A) Sample Ak09; samples were run blind. Pairs of nucleotides were pooled and dispensed for dyad pyrosequencing. Template tailoring of homozygous samples using PNA as well as ddCTP. The genotype for this sample is 7, 7 instead of apparent genotype 6, 6 as explained in Fig. 1. (B) Sample Ak08; samples were run blind. Pairs of nucleotides were pooled and dispensed for dyad pyrosequencing. Template tailoring of heterozygous samples using PNA as well as ddCTP. The genotype for this sample is [7,12] instead of apparent genotype [6,11], as mentioned in Fig. 1.

Table 1

The dye-labeling, sequences for amplification, and sequencing oligonucleotides as well as PNA probes used for pyrosequencing

Marker	Dye labeling ^{a)}	Primer sequences
Amplification primer-GATA172D05	Biotin	5'-Biotin/ataattgaaagcccgattc 5'-tagtggtgatggtgcacag
Seq. primer-GATA172D05	No labeling	5'-gtgatggtgcacagatata
PNA Probe-GATA172D05	No labeling	N-Terminus-gctatatcaatacct-C-Terminus
Amplification primer-GATA31E08	Biotin	5'-Biotin/cagctgacagacacagaga 5'-aggggagaaggctagaatga
Seq. primer-GATA31E08	No labeling	5'-agatagatgataggaggg
PNA Probe- GATA31E08	No labeling	N-Terminus-agatagagatacat-C-Terminus
Amplification primer-HPRTB	Biotin	5'-Biotin/gtctctattccatctctgtctcc 5'-ttctttctcaccctgtct
Seq. primer	No labeling	5'-gagaaggcatgaatttgctt
PNA Probe	No labeling	C-Terminus-gatagacagatagag-N-Terminus

^{a)} Forward primer is always labeled.

Table 2

DNA samples showing genotypes for the markers GATA172D05 and GATA31E08 determined by both methods (PNA-Dyad Pyrosequencing and Genetic analyzer ABI PRISM 3100)

GATA172D05				GATA31E08			
Sample #	ABI PRISM 3100 Genetic Analyzer	PNA-dyad pyrosequencing	Sample #	ABI PRISM 3100 Genetic Analyzer	PNA-dyad pyrosequencing	Sample #	ABI PRISM 3100 Genetic Analyzer
01	6, 6	6, 6	01	13, 13	13, 13	01	13, 13
02	10, 10	10, 10	02	11, 11	11, 11	02	11, 11
03	10, 10	10, 10	03	7, 7	7, 7	03	7, 7
04	8, 10	8, 10	06	9, 9	9, 9	06	9, 9
06	10, 10	10, 10	08	11, 11	11, 11	08	11, 11
08	12, 12	12, 12	09	12, 12	12, 12	09	12, 12
09	11, 11	11, 11	Ak3	10, 13	10, 13	Ak3	10, 13
10	6, 9	6, 9	Ak5	9, 12	9, 12	Ak5	9, 12
Ak2	6, 12	6, 12	Ak6	9, 12	9, 12	Ak6	9, 12
Ak4	8, 8	8, 8	Ak8	9, 12	9, 12	Ak8	9, 12
Ak5	10, 12	10, 12	Ak9	10, 10	10, 10	Ak9	10, 10
Ak8	11, 12	11, 12	Ak11	7, 12	7, 12	Ak11	7, 12
Ak9	10, 12	10, 12	Ak12	7, 7	7, 7	Ak12	7, 7
Ak10	11, 11	11, 11	Ak13	7, 11	7, 11	Ak13	7, 11
Ak15	8, 8	8, 8	Ak19	11, 12	11, 12	Ak19	11, 12