Analytical Biochemistry 391 (2009) 1–10

Contents lists available at ScienceDirect

Analytical Biochemistry

journal homepage: www.elsevier.com/locate/yabio

A single nucleotide polymorphism melt curve assay employing an intercalating dye probe fluorescence resonance energy transfer for forensic analysis

Micah D. Halpern [a,*], Jack Ballantyne [b,c]

[a] Life Sciences Division, Midwest Research Institute, Kansas City, MO 64110, USA
[b] Department of Chemistry, University of Central Florida, Orlando, FL 32816, USA
[c] National Center for Forensic Science, Orlando, FL 32816, USA

ARTICLE INFO

Article history:
Received 13 January 2009
Available online 9 May 2009

Keywords:
SNP
FRET
Dye
Probe
Multiple
Melt curve analysis
Forensics

ABSTRACT

The characterization and use of DNA sequence polymorphisms are an important aspect of forensic analysis. A number of approaches are being explored for single nucleotide polymorphism (SNP) genotyping, but current detection methods are subject to limitations that adversely impact their utility for forensic analysis. We have developed a novel method for genotyping both single and multiple SNPs that uses an intercalating dye and a probe labeled with a single fluorophore to affect a fluorescence energy transfer. Melting curve analysis is then used to distinguish true alleles from mismatched alleles. We term the new method dye probe fluorescence resonance energy transfer (dpFRET). In the current work, development proceeded at first with synthetic DNA template testing to establish proof of concept for the chemistry involved, followed by the design of polymerase chain reaction (PCR)-based genomic DNA assays to demonstrate potential forensic applications. The loci chosen for testing included both nuclear (MHC DRB) and mitochondrial DNA (cytochrome *b*) genes. A preliminary assessment of the sensitivity limits of the technology indicated that dpFRET was capable of accurately genotyping DNA from one single diploid cell equivalent. This technology could also potentially impact a wide range of nonforensic disciplines to aid in discovery, screening, and association of DNA sequence polymorphisms.

Single nucleotide polymorphisms (SNPs)[1] are defined as a nucleotide variation in chromosomal DNA at a particular site within a sequence (e.g., CCT to TCT). Analysis of this type of sequence polymorphism has become central to a number of disciplines due to the ability to use these changes to investigate a number of genetic phenomena from identification to disease. SNPs are particularly appealing for forensic applications for a number of reasons, including small polymerase chain reaction (PCR) product size for degraded DNA samples, potential for multiplexing and automation, and simplified analysis. SNP markers are generally biallelic, with two possible alleles resulting in three possible genotypes. This means that a modest amount of information can be generated per marker for SNPs. It has been estimated that approximately 50 to 100 SNP markers would be required to match the discriminatory power of 10 to 16 short tandem repeat (STR) loci [1]. This poses a problem in that it is difficult to simultaneously amplify a suitable number of SNP markers from low-DNA content samples. The ultimate solution for SNP typing would be an approach that could genotype multiple polymorphisms per reaction, thereby reducing the impact on sample consumption.

Current SNP detection methods encompass a variety of formats [2,3]. Some of the primary SNP typing methods that have been evaluated for forensic studies include minisequencing [4], Taqman [5], and pyrosequencing [6]. Development for SNP detection and screening in other fields has the potential to contribute to advancing approaches in forensic science by alleviating or avoiding issues posed by current approaches. These technologies are almost exclusively PCR based and fall under the major categories of hybridization based, enzyme based, postamplification detection, and different forms of DNA sequencing.

The goal of the research detailed in the current work focuses on improving SNP hybridization methodology. Within this category, developments aimed at discovering and identifying DNA polymorphisms can be classified under two major subcategories: (i) generic DNA intercalator techniques [7,8] and (ii) strand-specific hybridization [9]. Genotyping methods solely using intercalating dyes have shown a somewhat low level of resolution between larger amplicons with similar sequence [10,11]. Recent developments for achieving higher resolution screening have included the use of novel proprietary dyes and advances in data analysis [12]. Although somewhat limited in their ability to resolve many

* Corresponding author. Address: Life Sciences Division, Midwest Research Institute, 425 Volker Boulevard, Kansas City, MO 64110, USA. Fax: +1 816 360 1955.
E-mail address: mhalpern@mriresearch.org (M.D. Halpern).

[1] *Abbreviations used:* SNP, single nucleotide polymorphism; PCR, polymerase chain reaction; STR, short tandem repeat; FRET, fluorescence resonance energy transfer; dpFRET, dye probe fluorescence resonance energy transfer; NHS, *N*-hydroxysuccinimidyl; HPLC, high-performance liquid chromatography; NTC, no-template control; T_m, melting temperature; EDTA, ethylenediaminetetraacetic acid.

0003-2697/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2009.05.003

different types of changes in DNA between samples, the major benefit to this hybridization-based approach is the cost savings associated with minimized reagent requirements and reduced assay design constraints. The second hybridization subcategory comprises strand-specific methods that use additional nucleic acid reaction components to monitor the progress of amplification reactions, typically through fluorescence resonance energy transfer (FRET). There are two commonly used types of FRET probes; those using hydrolysis of nucleic acid probes to separate donor from acceptor (i.e., Taqman [13]) and those using hybridization to alter the spatial relationship of donor and acceptor molecules (i.e., molecular beacons [14] and dual-labeled hybridization probes [15,16]). The use of either approach requires labeling with two fluorescent molecules, thereby increasing the cost involved in using these approaches. In addition, both methods require the presence of a reasonably long stretch of known sequence so that the probe/probe pair can bind specifically in close proximity to each other. This can be a problem in some applications where the length of known sequences that can be used to design an effective probe may be relatively short. Furthermore, the use of pairs of probes involves more complex experimental design and requires careful design parameters often limited by sequence identity.

The optimal approach to discovery and screening of single and multiple polymorphisms would be to combine the reduced cost and ease of use of generic intercalating dyes with the resolution and increased sensitivity of hybridization probes. A less sophisticated version of this approach was demonstrated by genotyping with unlabeled probes post-PCR [17] as well as inclusion within the amplification reaction [18,19]. This required additional analysis and was not able to discriminate all potential alleles. An integrated system using FRET between an intercalating dye and a probe labeled with a single fluorophore was also reported previously by Howell and coworkers [20]. They demonstrated a basic application of the approach that showed a dramatic increase in signal intensity when compared with standard intercalating dye and other FRET approaches. The same technology may also be used for studying changes in DNA hybridization [21]. Takatsu and coworkers [22] described a related approach based on labeled nucleotide incorporation followed by dye/fluorophore FRET detection. However, these studies did not identify and demonstrate the true potential of a combined dye probe approach for genotyping both single and multiple SNPs with minimal sample consumption.

In this work, the method we developed for genotyping both single and multiple SNPs used an intercalating dye and a probe labeled with a single fluorophore to affect a fluorescence energy transfer (dpFRET) followed by melt curve analysis (Fig. 1). The purpose of the study was development of the dpFRET technology for SNP genotyping and preliminary determination of the limit of detection for use in forensic analysis. This approach was capable of differentiating single and multiple SNPs (up to 30% divergence) within a sample when compared with a reference sequence. Also unique to this method is the generation of an amplicon melt peak that functions as a positive amplification control. We demonstrated that this approach is robust for low copy number detection with no apparent allelic dropout. The approach also proved to be successful at genotyping both haploid and diploid loci with highly flexible probe design strategies and increased sensitivity and resolution compared with the use of an intercalating dye with unlabeled probes. The results suggest that this approach would potentially contribute to advancing the use of SNPs in forensic analysis by providing the capability to genotype multiple polymorphisms with a single assay that minimizes the amount of sample consumed and has single genomic equivalent sensitivity with reduced cost and time to results compared with other approaches.

Materials and methods

Synthetic SNP testing

Detection and analysis of single and multiple SNPs using dpFRET was first tested using two synthetic template libraries

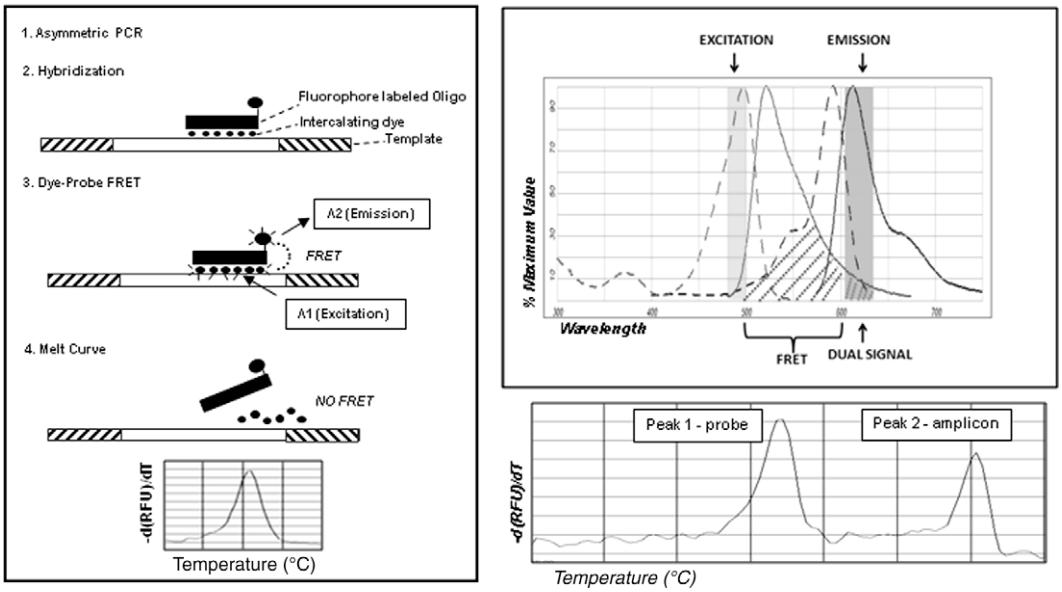


Fig. 1. dpFRET SNP genotyping strategy. The basic dpFRET protocol (left panel) entails (1) generation of template for probe hybridization by asymmetric PCR, (2) hybridization of a fluorophore-labeled probe in the presence of a DNA intercalating dye, (3) FRET detection of the interaction between the dye and probe, and (4) standard melt curve analysis for both the probe and amplicon. Excitation and emission spectra for both the dye and fluorophore attached to the probe are illustrated (top of right panel), with hash marks detailing the region where the dye emission and fluorophore excitation overlap. Gray boxes delineate the filter bandwidths used for both excitation and emission measurements. A small portion of the dye emission labeled as dual signal is detected by the emission filter and results in signal donated by the amplicon intercalated dye. Typical results (bottom of right panel) for dpFRET SNP genotyping produce both a probe and amplicon melt peak, with the amplicon peak resulting from dual signal donated by the intercalated dye.

composed of a variable human sequence and a variable animal species sequence library. Sequence corresponding to positions 14925 to 14974 of the Cambridge human mitochondrial genome (J01415) and mutated templates were synthesized and purified by standard desalting, and concentrations were standardized by a commercial source (Integrated DNA Technologies, Coralville, IA, USA). The mutated templates consisted of a mix of molecules containing every possible single point mutation within the 30-base central core region (positions 14935–14964). The variable animal species template library contained sequence corresponding to the same position of the Cambridge human mitochondrial genome from a number of animal species and was generated by the same commercial source. Nonvariable 10-base sequences flanking the variable regions were also included in each template to avoid potential problems associated with incomplete synthesis such as N-1 templates.

Both template libraries were evaluated by standard melt curve analysis with human reference probe sequences (30 bases: ACGTCTCGAGTGATGTGGGCGATTGATGAA; 21 bases: TCGAGTGAT GTGGGCGATTGA; 15 bases: GTGGGCGATTGATGA) labeled at the 3′ terminus with a Texas Red-X *N*-hydroxysuccinimidyl (NHS) ester. The fluorescent probe was commercially synthesized, high-performance liquid chromatography (HPLC) purified, and quantity standardized by a commercial source (Integrated DNA Technologies). Hybridization reactions contained 1× SYBR Green I Master Mix (Bio-Rad Laboratories, Hercules, CA, USA), 50 μM template, and 5 μM labeled probe and were subjected to the following thermal protocol on an IQ5 real-time thermal cycler (Bio-Rad Laboratories): 95 °C for 1 min, 25 °C for 1 min, and an incremental increase of 0.2 °C to a final temperature of 95 °C. A standard excitation filter of 490 nm (30 nm bandwidth) was coupled with an emission filter of 620 nm (20 nm bandwidth) placed in the appropriate corresponding position of the emission filter wheel.

SNP species identification: cytochrome b

Published sequences (NCBI) of cytochrome *b* for multiple animal species were aligned using MegAlign (DNAStar, Madison, WI, USA), and regions of conservation were used to manually design primers according to standard practice. Optimal primer sequences used for dpFRET testing were CYTB 0088F Mix (5′-TCCGCATGATGAAAyTTyGGnTC-3′) and CYTB 0438R Mix (5′-GTGGCCCCTCAGAAdGAyATyTG-3′), where y = C or T; d = A, G, or T; and n = A, G, C, or T. Previously extracted and quantitated genomic samples derived from whole blood for multiple animal species were provided by the Brookfield Zoo (Brookfield, IL, USA). Previously extracted and quantitated genomic samples for human and ferret species were provided by the National Center for Forensic Science (Orlando, FL, USA). All quantitation was verified using PicoGreen and supplier-recommended protocols (Invitrogen, Carlsbad, CA, USA). Asymmetric PCR reactions contained 1× SYBR Green Master Mix (Bio-Rad) composed of 50 mM KCl, 20 mM Tris–HCl (pH 8.4), 0.2 mM of each deoxynucleoside triphosphate (dATP, dCTP, dGTP, and dTTP), 1.25 U of iTaq DNA polymerase, 3 mM $MgCl_2$, 1× SYBR Green I, and 10 nM fluorescein supplemented with 500 nM forward primer and 15 nM reverse primer. The protocol used for asymmetric amplification included an initial denaturation at 95 °C for 3 min followed by 40 cycles of 95 °C for 10 s and 59 °C for 40 s. This was immediately followed by 40 cycles of 95 °C for 10 s and 56 °C for 40 s. Following amplification, the reaction was supplemented with 5 μM commercially synthesized 30-base human reference probe (described previously) and subjected to melt curve analysis using a 0.5 °C incremental increase in temperature on an IQ5 real-time PCR platform.

SNP individual identification: MHC DRB

Published sequences (NCBI) of MHC DRB for multiple animal species were aligned using MegAlign (DNAStar), and regions of conservation were used to manually design primers according to standard practice. Optimal primer sequences for dpFRET testing were UNIV_MHCdr_3F Mix (5′-ACGGsACsGAGCGGGTG-3′) and UNIV_MHCdr_3R (5′-CACCCCGTAGTTGTGTC-3′), where s = G or C. Previously extracted and quantitated genomic samples derived from blood and/or hair from a calf, a dam, and two suspected bulls for giraffe paternity and two families of captive Humboldt penguins (*Spheniscus humboldti*) were provided by the Brookfield Zoo. Quantitation was verified as described previously. Asymmetric PCR reactions containing 1× SYBR Green Master Mix (Bio-Rad) buffer composition as stated above, 100 nM forward primer, and 500 nM reverse primer were amplified using the following thermal protocol: initial denaturation at 95 °C for 3 min followed by 40 cycles of 95 °C for 10 s and 63 °C for 40 s. This was immediately followed by 40 cycles of 95 °C for 10 s and 59 °C for 40 s. Following amplification, the reactions were supplemented with 5 μM commercially synthesized Texas Red fluorescently labeled probe as well as unlabeled probes (giraffe samples) designed against a human or penguin reference sequence as follows: human probe 2 (ATAACCAAGAGGAGTCCGTGCGCTTCGACAGCGA/3′TR), human probe 3 (5′TR/AGCGACGTGGGGGAGTACCGGGCGGTGACGGAGCTGGG), human probe 4 (GGGCGGCCTGATGCCGAGTACTGGAACAGCCAGAAGGA/3′TR), human probe 5 (CAGAAGGACCTCCTGGAGCAGAGGCGGGCCGCGGTGGA/3′TR), penguin probe 4 (GGCTGAGGTGGACACGTACTGCCGA/3′TR), and penguin probe 5 (CACAACTACGGGGTGGTGACCCCTTTCACT/3′TR). The giraffe samples were not tested with penguin reference probes. Reactions were subjected to melt curve analysis using a 0.5 °C incremental increase in temperature on an IQ5 real-time PCR platform (Bio-Rad). Amplicons generated for dpFRET testing were also sequenced using standard dideoxy sequencing according to the manufacturer's protocol (Applied Biosystems, Foster City, CA, USA) for comparison with dpFRET results.

Sensitivity and allelic dropout: MHC DRB

Human genomic samples were used to determine assay sensitivity and potential for allelic dropout. Both homozygote and heterozygote samples were tested using protocols described previously for MHC DRB. Samples were requantitated using PicoGreen and the manufacturer's protocol (Invitrogen) and were diluted 10-fold from 50.0 pg (∼ 10 genomic equivalent copies) to 0.5 pg (∼0.1 genomic equivalent copies) in water. Seven replicates for the homozygote, heterozygote, and no-template control (NTC) were used to demonstrate detection limit reproducibility. Amplification and melt curve analysis were performed as described previously for the MHC DRB locus.

Results

As illustrated in Fig. 1 (left panel), dpFRET uses asymmetric PCR for template generation followed by FRET-based melt curve analysis between an intercalating dye and a fluorophore-labeled probe. Although different dye/fluorophore combinations are possible, we employ here SYBR Green I as the donor dye (excitation 490 and emission 520) and Texas Red as the acceptor fluorophore (excitation 590 and emission 620). The dye is excited at a wavelength of 490 nm and emits at a wavelength of 520 nm, which is then transferred via FRET to the fluorophore on the labeled probe and reemitted at 620 nm. Fig. 1 (top of right panel) shows excitation and emission wavelengths for SYBR Green I (light gray) and Texas Red (dark gray), the region of FRET between the two molecules (dark hashed), and the dual emission signal generated by both dyes

(light hashed). Filter bandwidths used to provide the excitation signal (light gray box) and measure the emission signal (dark gray box) are also labeled based on filters provided with the real-time PCR system used (IQ5, Bio-Rad).

Also illustrated in Fig. 1 (bottom of right panel) is one of the uniquely beneficial outcomes for the dpFRET approach, namely, the generation of two melt peaks. The peak at the lower melt temperature is a result of the signal from the FRET probe, and the peak at the higher temperature is a result from the melting of the amplicon itself. The amplicon melt peak is generated by fluorescence of intercalated SYBR Green I at the tail end of the SYBR Green I emission spectrum (dark gray hashed region in Fig. 1, top of right panel). This secondary melt peak provides a positive signal for amplification of specific product and can be used to distinguish nonspecific signal occasionally generated by the probe for higher cycle amplification reactions.

Synthetic DNA single mitochondrial DNA SNP testing: template variation

Results for the synthetic variable human sequence template library testing using a 30-base probe are shown in Fig. 2. The top panel of the figure shows the melt temperature for every mutation at each position within the template tested with a fluorophore-labeled probe. Error bars of ±0.4 °C are labeled for each data point to account for thermal block variation. The range for an exact match (reference template) is highlighted across the graph. The bottom panel represents similar testing with an unlabeled probe (standard intercalating dye melt analysis) to ascertain the fluorophore effect on melting temperature. The 30-base fluorophore-labeled probe resulted in discrimination of any change at any position except for mutations in the template complementary to probe nucleotides 30, 29, and 1. In contrast, the unlabeled probe was unable to discriminate mutations at multiple positions both distal and internal within the template (probe nucleotides 26, 22, 13, and 1). Also important to note is the similarity between the overall labeled and unlabeled probe melt points, with the labeled probe points displaying more significant variation from the reference template for most mutations.

To understand the effect of probe size, 21- and 15-base fluorophore-labeled probes were also tested and showed similar results with finer resolution at the ends of the template using the dpFRET approach (data not shown). The fluorophore-labeled 21-base probe was indistinguishable from the reference for template mutations complementary to probe positions 21 and 1 and showed no effect due to template mutation in flanking sequence. Similar melting protocols using an unlabeled probe resulted in melt temperatures indistinguishable from the reference for mutations at multiple positions (probe nucleotides 17, 8, 4, 3, 2, and 1). In addition, an effect was seen for mutations in upstream sequence flanking the unlabeled probe (probe nucleotides +1, +3, and +4). The fluorophore-labeled 15-base probe resulted in differential melt temperatures from the reference for all mutations except probe nucleotide 15, with a minor difference due to a flanking mutation (probe nucleotide –12). An unlabeled 15-base probe was not tested.

Synthetic multiple mitochondrial DNA SNP testing: species variation

All mitochondrial DNA SNP synthetic animal species templates showed reduced melt temperatures compared with the human reference sequence when hybridized with a human probe sequence (Fig. 3). A few templates are listed with the numbers of SNPs in parentheses in Fig. 3 to illustrate the range of sequence divergence. In general, an increased number of SNPs within the template tended to reduce the melt temperature. Four species templates (skate, aardvark, dogfish, and dugong) did not produce melt curves when tested with a human probe sequence due to increased numbers (10–12) of SNPs. It should also be noted that the closely re-

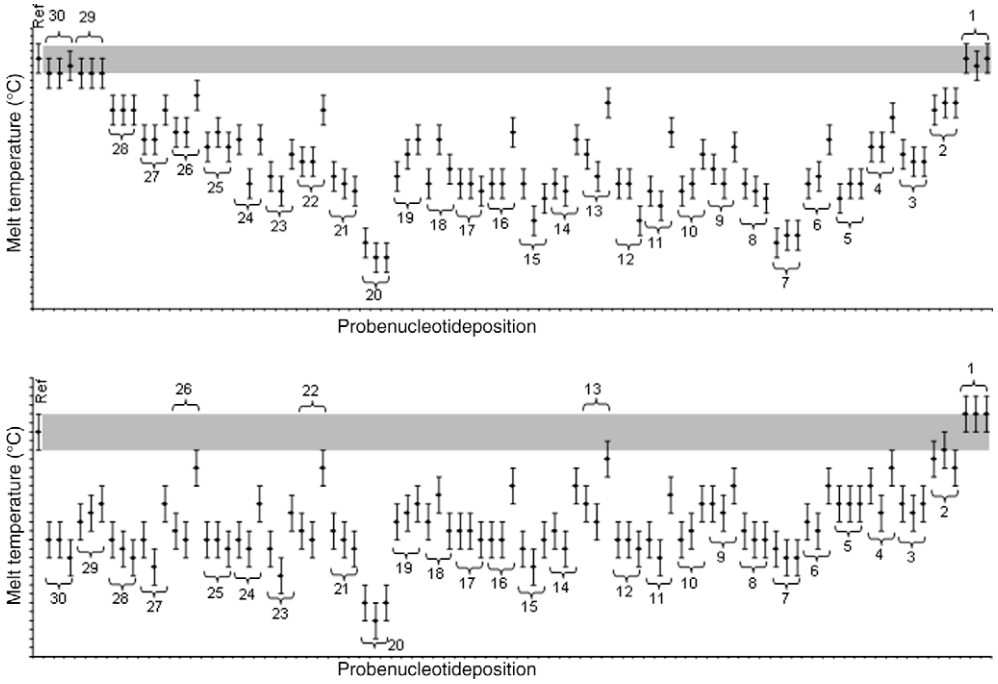


Fig. 2. Synthetic templates: Single SNP detection. Melt curve results for testing of all potential mutations within human cytochrome *b* with 30-base fluorophore-labeled (top panel) and unlabeled (bottom panel) probes are shown. The reference sequence temperature range for a template lacking an SNP (gray bar) is highlighted (error bars = ±0.4 °C). dpFRET resulted in discrimination of any change at any position except for mutations in the distal ends. The unlabeled probe was unable to discriminate mutations at multiple positions both distal and internal within the template and resulted in reduced melt peak temperature variation from the reference template for most mutations.

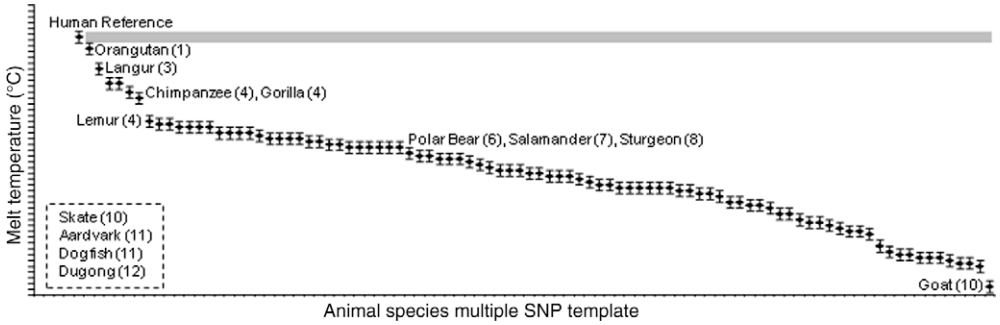


Fig. 3. Synthetic templates: Multiple SNP detection. Melt curve results for testing of variable numbers of SNPs within a single template based on various animal species cytochrome *b* sequences are shown. Select species are highlighted to illustrate divergence, with numbers of SNPs relative to the human reference sequence shown in parentheses. Four templates did not generate a melt curve: skate, aardvark, dogfish, and dugong. The human reference sequence temperature range for a template lacking an SNP (gray bar) is highlighted (error bars = ±1.0 °C). Results indicated the ability to detect as many as nine collective mutations within a 30-base sequence, with no melt peak produced for additional (more than nine) mutations. dpFRET was able to differentiate human species from any other species but was unable to differentiate among all species tested.

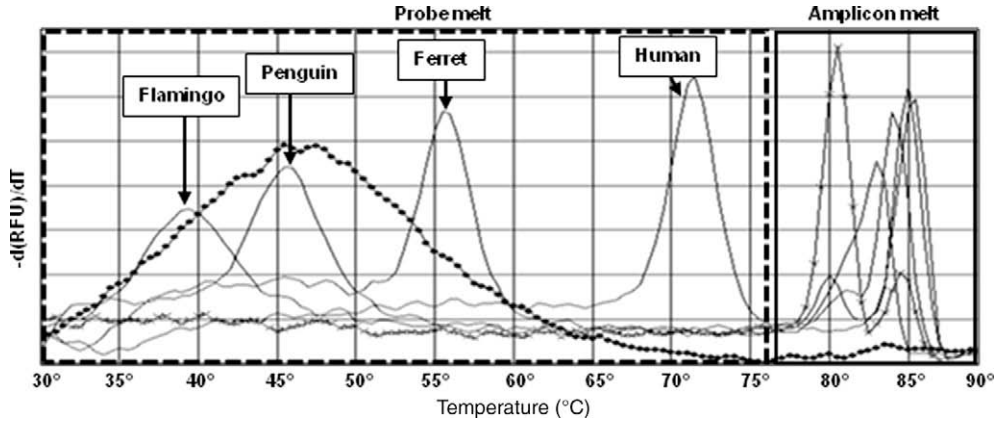


Fig. 4. dpFRET cytochrome *b* SNP real-world species testing. Four of five species tested resulted in both probe and amplicon peaks (smooth) and are labeled with species designations. Python (cross) was unable to generate a probe peak due to increased sequence variation but resulted in an amplicon peak, indicating specific amplification. The NTC (diamond) resulted in a broad probe peak but did not produce an amplicon peak, indicating nonspecific amplification and/or probe template interaction.

lated orangutan sequence showed a differential melt temperature and contained only a single SNP. Unlabeled probe was not tested against the animal species library.

Haploid locus SNP testing: cytochrome b

Results for cytochrome *b* melt curve species testing for genomic DNA extracts from living animals are shown in Fig. 4, with probe melt temperatures listed in Table 1. All species tested resulted in two melt peaks, indicative of both amplification (SYBR amplicon peak) and probe hybridization, except python, which produced only an amplification peak. Published python species sequence differs by more than 10 nt from the human reference probe sequence used for testing. The NTC resulted in a broad nonspecific probe melt peak but did not exhibit any significant amplification, as evidenced by the lack of an amplicon peak

Table 1
Probe and amplicon melt temperatures for dpFRET cytochrome B Species testing

ID	Probe (°C)	Amplicon (°C)
Human	71.5	85.5
Penguin	45.5	85.5
Flamingo	39.5	84.0
Ferret	55.5	83.0
Python	–	81.0
Negative	47.0	–

Diploid locus SNP testing: MHC DRB

Paternity results for real-world testing of one suspect giraffe family and two known Humboldt penguin (*S. humboldti*) families are shown in Figs. 5 and 6. A sequence alignment for the amplification products produced using the universal MHC DRB PCR assay is listed at top of each figure. Giraffe samples (Fig. 5) were genotyped with both labeled probes (left graph in each panel) and unlabeled probes (right graph in each panel). Labeled probes resulted in both higher signal intensity and increased resolution, particularly for probes 2 and 5. In Fig. 6, melt curves generated by dpFRET analysis were converted to allele designations of either A, B, or C for presentation purposes. Paternity results previously established by the Brookfield Zoo through both Southern blot analysis and zookeeper records for the two penguin families are shown for comparison. Previously established penguin paternity agreed with results generated by dpFRET analysis.

SNP assay sensitivity and allelic dropout

The reproducible limit of detection using dpFRET for MHC SNP analysis based on seven replicates was 5.0 pg (∼ 1 diploid cell equivalent) for both homozygote and heterozygote samples

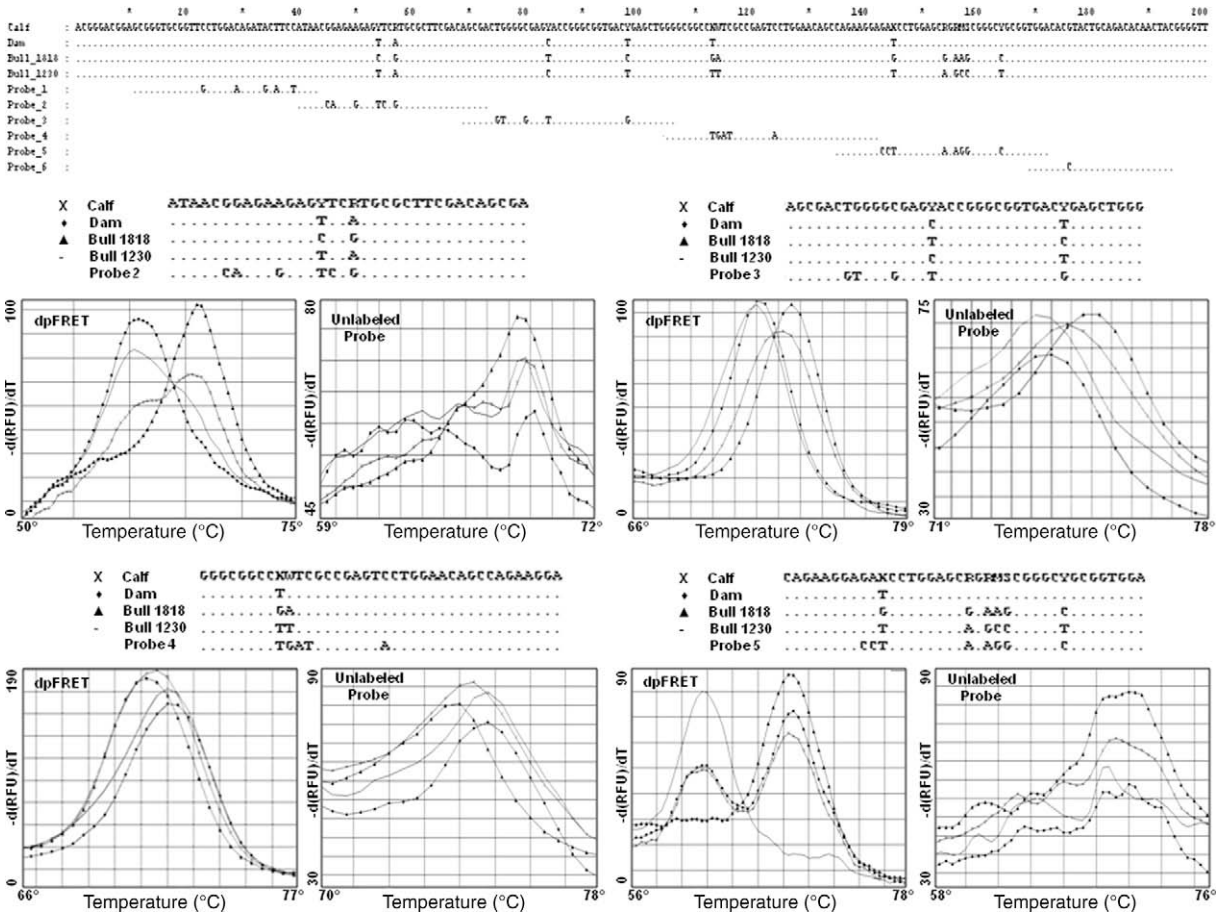
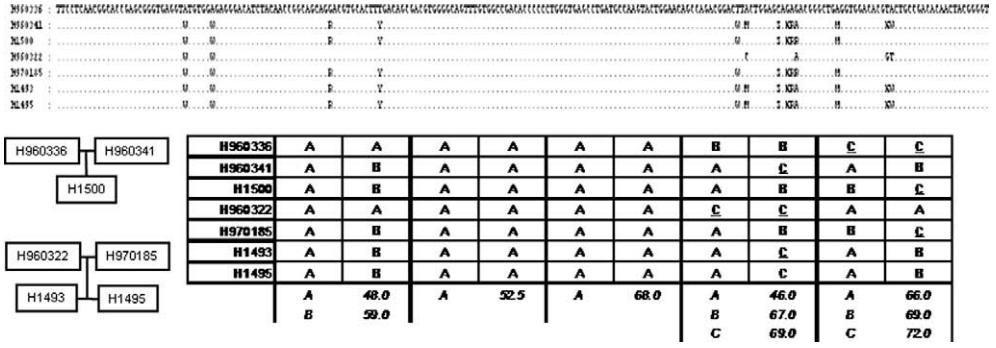


Fig. 5. dpFRET MHC DRB SNP giraffe paternity testing. A sequence alignment for all samples tested is shown at the top of the figure with standard IUB codes for mixed base positions. Each sample was tested with both labeled dpFRET probes (left graph in each panel) and unlabeled probes (right graph in each panel). dpFRET resulted in higher signal intensity for all probes and markedly increased peak resolution for probes 2 and 5. Genotyping for unlabeled probes 2 and 5 was not possible due to inadequate melt peak generation and resolution. All probe sequences are aligned against the giraffe calf sequence, with differences from listed giraffe sequences shown.



H960336	A	A	A	A	A	A	B	B	C	C
H960341	A	B	A	A	A	A	A	C	A	B
H1500	A	B	A	A	A	A	A	B	B	C
H960322	A	A	A	A	A	A	C	C	A	A
H970185	A	B	A	A	A	A	A	B	B	C
H1493	A	B	A	A	A	A	A	C	A	B
H1495	A	B	A	A	A	A	A	C	A	B
	A	48.0	A	52.5	A	68.0	A	46.0	A	66.0
	B	59.0					B	67.0	B	69.0
							C	69.0	C	72.0

Fig. 6. dpFRET MHC DRB SNP penguin paternity testing. A sequence alignment for all samples tested is shown at the top of the figure with standard IUB codes for mixed base positions. Melt curves generated by dpFRET analysis were converted to allele designations of either A, B, or C, with coinciding melt peak temperatures listed for each probe (lower right panel). Paternity results previously established by the Brookfield Zoo through both Southern blot analysis and zookeeper records for the two penguin families are shown for comparison (lower left panel).

(Fig. 7). Fluorescent signal showed no decrease for 50.0 and 5.0 pg, and no allelic dropout was observed for the heterozygote at these concentrations. At a concentration of 0.5 pg (~ 0.1 genomic equivalent), replicates resulted in no detectable amplicon peak, allelic dropout for the heterozygote, and/or generally decreased signal for most samples. NTC samples resulted in the absence of a characteristic MHC DRB amplicon peak but did exhibit nonspecific amplification or probe hybridization.

Discussion

The goal of this study was to investigate a novel technology for improved genotyping of single and multiple SNPs and to explore its potential use in forensic analysis. Synthetic DNA template testing was used initially to establish proof of concept for the chemistry, thereby removing the necessity to account for variables upstream in the process (e.g., PCR artifacts). Once the chemistry itself proved to be suffi-

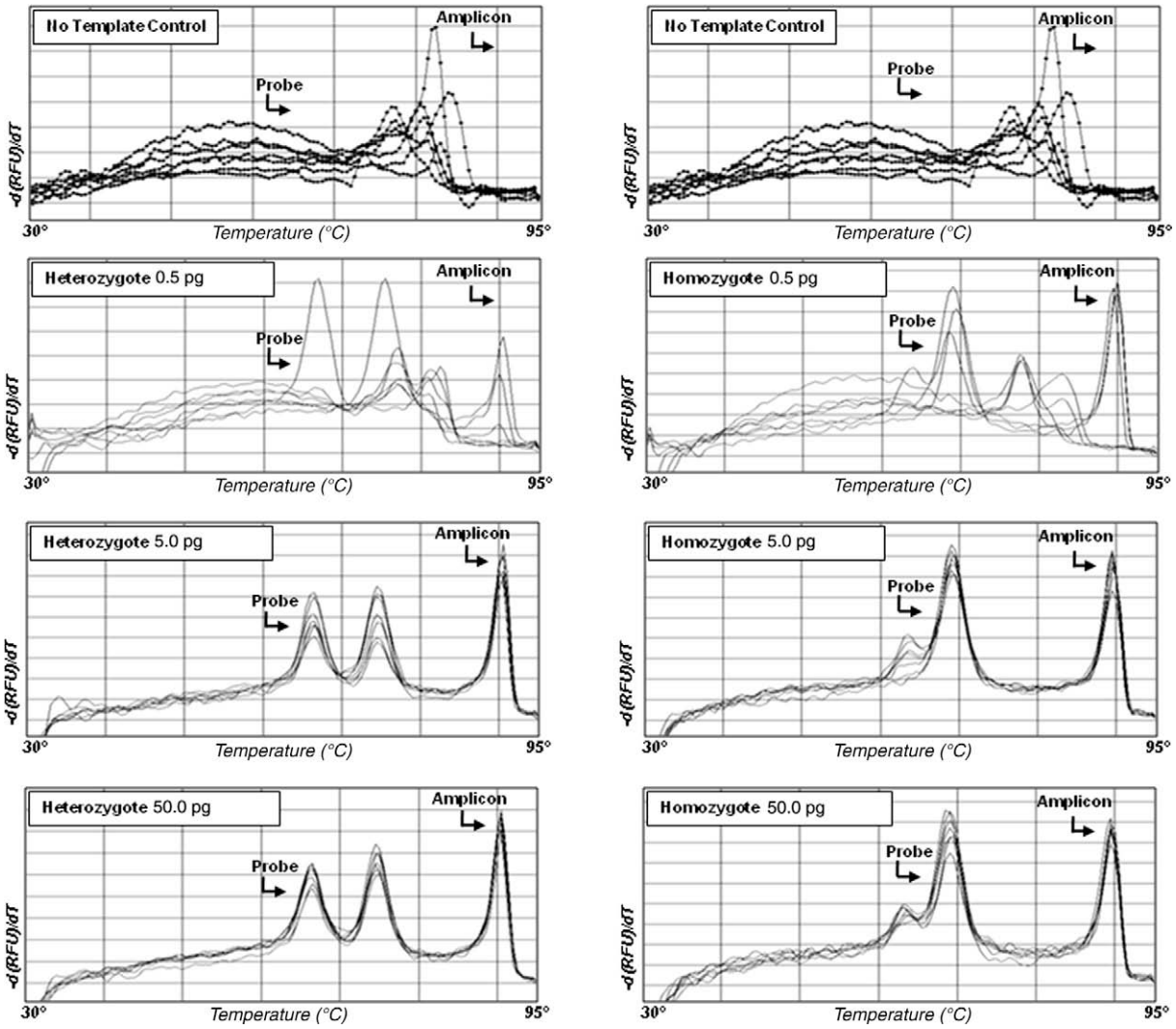


Fig. 7. dpFRET SNP sensitivity and lack of allelic dropout. Seven replicates for a human heterozygote (left panels) and a homozygote (right panels) were assayed at 0.5, 5.0, and 50.0 pg including no-template controls (NTCs). NTCs resulted in nonspecific amplification and/or probe hybridization, as indicated by the lack of an MHC DRB characteristic amplicon peak. Both the heterozygote and homozygote samples resulted in reproducible detection at 50.0 and 5.0 pg, with unreliable detection at 0.5 pg, as indicated by lack of an amplicon peak, allele dropout for the heterozygote, or overall reduced signal for the amplicon and/or probe.

ciently robust, the next step was to develop PCR-based assays that could be used for general screening purposes but, more important, could test real-world application of the approach. Following development of the complete protocol, it was necessary to establish the boundaries within which the technology was applicable.

Hybridization-based genotyping of DNA polymorphisms depends on oligonucleotide melting temperature (T_m). The T_m of duplex DNA is defined as the temperature where one-half of the nucleotides are Watson–Crick paired and one-half are unpaired [9]. T_m can be predicted using a variety of formulas, with the most accurate being the thermodynamic nearest neighbor model [23]. The nearest neighbor model is based on the assumption that probe hybridization energy can be calculated from the enthalpy and entropy of all nearest neighbor pairs, including a contribution from each dangling end [24]. Dangling ends (also known as end effects and end fraying) account for the effects seen when a shorter probe is bound to a target with flanking sequence [25,26]. Various interactions contribute to probe/template stability, but it has been demonstrated that melting of the complex is initiated at the ends of the duplex [27]. The results of this work suggest that this dangling end effect provides dpFRET with a higher level of resolution as compared with an intercalating dye.

Synthetic template single SNP resolution

Results for all probe sizes tested (30, 21, and 15 bases) demonstrated that dpFRET is not currently capable of assigning a unique melt temperature to every potential change (position and nucleotide mutation). This is evidenced by different mutations at different positions sharing the same melt temperature. However, the capability of producing a differential melt temperature relative to a perfect match with the probe sequence was demonstrated. Thus, a mutation at two different locations within the sequence can potentially produce the same melt temperature, but that temperature is nearly always different from a perfect match between the probe and reference sequence. Mutations at the ends (5′ and 3′) of the template were indistinguishable from the reference sequence for larger (30 and 21 bases) probes, most likely due to inadequate end effects related to the size of the probe. A reduction in probe size to 15 bases produced a differential melt temperature from the reference sequence for all mutations at all positions, excluding template sequence complementary at the 3′ end of the probe. In its current state, dpFRET can be applied for SNP discovery with follow-up sequencing for determination of the exact position and mutation. For purposes of SNP screening, it may be necessary

to take into account assay design considerations for discrimination of certain targeted changes. These data suggest that probe size should be limited to 15 to 30 bases depending on application and desired level of resolution.

For both the 30- and 21-base probes, dpFRET showed higher resolution for internal template changes than did SYBR Green I (intercalating dye) alone. This result lends credibility to the hypothesized end effects theory. It appears that when an intercalating dye is used, internal mismatches are averaged out across the template as it melts. Any single mismatch is averaged with all matching nucleotides across a template, producing a lower signal-to-noise ratio due to increased noise. By localization of the differential melting signal to the end of the hybrid complex (fluorophore-labeled end), the effect is more significant because FRET can occur only across a limited distance commonly known as the Förster distance. Therefore, signal differences contributed by the mismatch remain constant, but the noise produced by dye intercalated at a distance is minimized. This would have the potential to increase the signal-to-noise ratio, providing a higher level of resolution. Although data have been generated for one particular 30-base sequence and appear to support the described approach, additional testing with a range of synthetic template sequences should be undertaken to lend further support for this hypothesis.

Synthetic template multiple SNP resolution

One of the benefits of dpFRET that could contribute to solving the SNP multiplexing requirement of forensic analysis is the ability to detect multiple changes within one template with a single probe design. In an effort to test the limits of this approach, a template library was synthetically generated and encompassed 1 to 12 SNPs in varying configurations based on a region of cytochrome *b* sequence for a number of animal species. The reference and complementary probe sequence were based on human cytochrome *b* with the intended application for animal species genotyping to determine human versus nonhuman sample source attribution. Results indicated the ability to detect as many as nine collective mutations within a 30-base sequence. Beyond approximately 9 base pairs, the probe and template were unable to hybridize in a manner sufficient to intercalate dye and transfer signal to the fluorophore probe for genotyping. Hence, even with nearly 30% divergence between the probe and template, a melt signal was still generated. Similar to probe size testing on the variable human sequence library, all probe/template complexes showed a reduced melt temperature compared with the reference human sequence but were unable to classify all templates as unique due to insufficient resolution. Thus, it was possible to tell human species from any other species, but the approach could not differentiate among all species. This is most likely due to the fact that multiple mutations at variable positions can have the same destabilizing effect on the DNA duplex. Similar to single mutation testing, additional synthetic sequence testing would also provide support for this hypothesis.

Cytochrome b haploid locus species identification

The next stage required development of a comprehensive protocol that incorporated PCR amplification of target sequences. Initial development focused on a haploid marker (mitochondrial cytochrome *b*) to minimize melt curve complexity. This was followed by development of diploid marker (nuclear MHC DRB) testing to explore the ability of the assay to discriminate two different alleles within the same individual. Both marker assays used original universal primer designs based on alignments of published sequence for multiple species. Results to date have shown the assays to be successful for amplification of multiple species with potential utility in a number of fields, including forensics, for species and individual identification.

Most notable for haploid cytochrome *b* species testing was a lower resolution of amplicon melt peaks (range = 81.0–85.5 °C) as compared with the melt peaks generated by a dpFRET probe (range = 39.5–71.5 °C). Unique identification of all species was possible through analysis of probe melt peaks in contrast to melt peaks generated from the PCR amplicons themselves, which were unable to resolve all species identifications. Similar to synthetic probe size testing results, the reduction in size for the signal generating duplex from amplicon to probe resulted in much higher resolution. Also similar to synthetic species testing, the real-world python sample showed no probe peak due to sequence divergence beyond the 10-base or 30% limit. It should also be noted that the NTC resulted in a broad probe melt peak but displayed no amplicon peak. This phenomenon was reproduced in follow-up development, and it was found that that the source of this peak is due primarily to excess probe concentration and/or nonspecific probe self-hybridization. It is hypothesized that the probe forms a probe/probe dimer that produces a signal at a significantly reduced melt temperature and characteristically results in no amplicon peak. Optimization of probe concentration was shown to partially alleviate this effect. Two melt peaks (probe and amplicon) is a unique characteristic of the dpFRET approach and provides an internal amplification control that can be used to further qualify positive results.

The detection of an amplicon peak measured at higher wavelengths (620 nm as opposed to 510 nm) is due to the strong fluorescent signal generated by SYBR Green I, whose emission tail end falls within this range. Thus, not only does the probe/template hybrid duplex contribute signal from FRET, but also fluorescent signal is donated by the intercalation of dye by the amplicon. It is this additional amplicon signal that can be used as a qualification of positive amplification in a manner similar to standard intercalating dye melt curve analysis. It was this additional information provided by the amplicon peak that led to the classification of the probe/probe dimer signal as noise. It is also important to note that the amplicon melt peaks produced early in development for species testing included a small shoulder peak. It was discovered through follow-up testing and optimization that this shoulder was due to a minor population of unlabeled probe resulting from incomplete synthesis that participated in amplification and production of a secondary specific product. Testing used to validate this hypothesis consisted of both reductions in probe concentrations and the addition of small amounts of ethylenediaminetetraacetic acid (EDTA) with the probe postamplification (data not shown). Both approaches were successful by limiting reagents necessary to produce the anomalous result. The method of reducing probe concentration lowered the minor population of unlabeled probe below the point at which it could contribute to amplification. The addition of EDTA reduced by chelation the amount of free magnesium required by the polymerase for activity. Optimization and development of the cytochrome *b* species testing assay provided information on important variables that was subsequently used in development of a bioassay for a more complex diploid locus.

MHC DRB diploid locus individual identification

Preliminary dpFRET testing of the MHC DRB marker used human sequence as a reference for probe designs and convenience samples from an unknown giraffe family and known Humboldt penguin families for testing. Paternity results for the penguin families had already been established by the Brookfield Zoo using Southern blot hybridization and AJ Jeffreys' VNTR probes (personal communication) and confirmed by zookeeper observations

(personal communication). Southern blot analysis is a labor-intensive process that requires sufficient amounts of genomic material (not available for giraffe paternity testing), prompting efforts for development of dpFRET as an alternative. An optimal approach would be to use a single probe set for SNP genotyping designed against a single reference. In this case, probes designed against human sequence were used to test samples from two other species. It was logical to assume that so long as divergence remained below approximately 30% as demonstrated by synthetic testing, this approach might be feasible. Results from human MHC DRB probe testing showed that probes were able to hybridize to two different species sequences (giraffe and penguin) and still largely maintained the ability to resolve differences between individuals. In other words, although there existed differential SNPs from the human design that were conserved among all alternative species samples tested, differences between both giraffe and penguin individuals could still be resolved for most probes. For regions of the amplicon that were more highly divergent from human, species-specific probes were designed against a Humboldt penguin reference sequence that showed better resolution than human probe sequences for heterozygote individuals.

Results from the MHC DRB experiments provided four important pieces of information that helped to further define the limit and application of this approach. First, it was demonstrated through synthetic templates as well as giraffe MHC DRB testing that dpFRET shows sufficiently superior results to standard intercalating dye genotyping with an unlabeled probe. This was exhibited by increased signal intensity and, more important, increased peak resolution for both data sets. Decreased confidence in the correct determination of giraffe paternity would have resulted if solely unlabeled probes had been used for testing. Second, the capability to resolve multiple alleles (heterozygote) within a single sample with a single assay was demonstrated. As opposed to other approaches (e.g., Taqman, SNPlex) that require single or multiple assays per SNP and, hence, would require large numbers of assays to genotype multiple SNPs within a sequence, dpFRET has the potential to significantly reduce the amount of sample consumed for testing through the ability to genotype multiple SNPs within a sequence with one assay. Third, by cross-species application of one set of probe sequences, the flexibility of both assay design and broad application was also demonstrated. This is particularly important in fields such as conservation biology, where current studies require assays specific to each and every animal tested. Finally, it was demonstrated that data equivalent to current approaches (i.e., DNA sequencing and Southern blot analysis) could be generated using dpFRET, which requires significantly less resources with faster time to results. Future development will include testing on additional species that could impact a number of fields, including conservation and population biology as well as forensics.

Assay sensitivity

The need to define the limits of sensitivity of a genotyping approach is important in many fields, especially forensics. Results for human dpFRET SNP genotyping demonstrated a preliminary established detection limit of a single genomic equivalent for both homozygote and heterozygote samples based on seven replicates. More important, no allelic dropout was observed for all replicates of a heterozygote at the level of a single genomic equivalent. Although this result will need to be confirmed with additional testing, it is not surprising based on the amplification protocol. The current version of dpFRET uses 50 to 80 cycles of amplification depending on the intended approach. Unpublished claims of reliable single copy detection have been made using similar numbers of cycles involving other chemistries. It is hypothesized that low levels of sensitivity are possible due to probe-based detection of specific amplicon that is essentially "blind" to generation of nonspecific product that can significantly impede size-based analysis. Signal generation produced by probe hybridization in dpFRET is capable of capitalizing on this same approach. Only nonspecific product with less than 30% divergence will produce a probe melt peak with dpFRET analysis. This was demonstrated by amplification products generated for cytochrome *b* dpFRET testing that were analyzed by agarose electrophoresis (data not shown). In addition to the specific product, multiple smaller nonspecific products that showed no signal during previous genotyping with dpFRET were also amplified.

Conclusions

The potential benefits of using dpFRET for discovering and screening polymorphisms in DNA are numerous. Although not as inexpensive as melt curve analysis based solely on the use of an intercalating dye, it is less costly than many other approaches for SNP genotyping. dpFRET uses a probe with a single fluorophore as opposed to other approaches (e.g., Taqman, dual hybridization probes, HyBeacons) that require multiple fluorophores that increase the cost per assay. Although dpFRET cannot distinguish all changes at all positions within a sequence, the approach is capable of detecting SNPs compared with a reference sequence and resulted in higher signal and peak resolution as compared with an intercalating dye and unlabeled probes. No additional equipment is required beyond a real-time thermal cycler with melt curve analysis as opposed to other SNP genotyping strategies (e.g., sequencing, SNPlex, allele-specific extension) that necessitate the use of costly electrophoresis equipment to separate and/or detect alleles. dpFRET probe design is extremely flexible, and preliminary results suggest that it is somewhat sequence independent as opposed to other approaches (e.g., dual hybridization probes, Taqman) that define specific strategies for designing probes. Sample manipulation following amplification necessary for non-melt-based SNP genotyping methods is not required for dpFRET, resulting in significantly shorter time to results and reduced labor time. It was demonstrated that the equivalent of a single genomic copy is reproducibly detected and genotyped, and this is equivalent to or more sensitive than other approaches. The described method can also be adapted to a number of platforms, including microfluidic and high throughput, for automated and point-of-care applications. The advantages of the dpFRET methodology are numerous and, more important, have the potential to alleviate some of the shortcomings of current forensic approaches to SNP detection and can be applied to any other field requiring SNP analysis.

Acknowledgments

The authors gratefully acknowledge the Midwest Research Institute for funding and support of this project. We also thank Jean Dubach and the Brookfield Zoo for providing both samples and helpful discussions.

References

[1] P. Gill, D.J. Werrett, B. Budowle, R. Guerrieri, An assessment of whether SNPs will replace STRs in national DNA databases: joint considerations of the DNA Working Group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDAM), Sci. Justice 44 (2004) 51–53.
[2] I.G. Gut, Automation in genotyping of single nucleotide polymorphisms, Hum. Mutat. 17 (2001) 475–492.
[3] P.Y. Kwok, Methods for genotyping single nucleotide polymorphisms, Annu. Rev. Genomics Hum. Genet. 2 (2001) 235–258.

[4] G. Tully, K.M. Sullivan, P. Nixon, R.E. Stones, P. Gill, Rapid detection of mitochondrial sequence polymorphisms using multiplex solid-phase fluorescent minisequencing, Genomics 34 (1996) 107–113.
[5] M. Lareu, J. Puente, B. Sobrino, B. Quintáns, M. Brión, A. Carracedo, The use of the LightCycler for the detection of Y chromosome SNPs, Forensic Sci. Intl. 118 (2001) 163–168.
[6] H. Andréasson, A. Asp, A. Alderborn, U. Gyllensten, M. Allen, Mitochondrial sequence analysis for forensic identification using pyrosequencing technology, BioTechniques 32 (2002) 124–133.
[7] K.M. Ririe, R.P. Rasmussen, C.T. Wittwer, Product differentiation by analysis of DNA melting curves during the polymerase chain reaction, Anal. Biochem. 245 (1997) 154–160.
[8] M. Liew, R. Pryor, R. Palais, C. Meadows, M. Erali, E. Lyon, C. Wittwer, Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons, Clin. Chem. 50 (2004) 1156–1164.
[9] J.G. Wetmur, DNA probes: applications of the principles of nucleic acid hybridization, Crit. Rev. Biochem. Mol. Biol. 26 (1991) 227–259.
[10] M.G. Herrmann, J.D. Durtschi, L.K. Bromley, C.T. Wittwer, K.V. Voelkerding, Amplicon DNA melting analysis for mutation scanning and genotyping: cross-platform comparison of instruments and dyes, Clin. Chem. 52 (2006) 494–503.
[11] M. Liew, M. Seipp, J. Durtschi, R.L. Margraf, S. Dames, M. Erali, K. Voelkerding, C. Wittwer, Closed-tube SNP genotyping without labeled probes: a comparison between unlabeled probe and amplicon melting, Am. J. Clin. Pathol. 127 (2007) 341–348.
[12] G.H. Reed, C.T. Wittwer, Sensitivity and specificity of single-nucleotide polymorphism scanning by high-resolution melting analysis, Clin. Chem. 50 (2004) 1748–1754.
[13] P.M. Holland, R.D. Abramson, R. Watson, D.H. Gelfand, Detection of specific polymerase chain reaction product by utilizing the 5′–3′ exonuclease activity of *Thermus aquaticus* DNA polymerase, Proc. Natl. Acad. Sci. USA 88 (1991) 7276–7280.
[14] S.A. Marras, F.R. Kramer, S. Tyagi, Genotyping SNPs with molecular beacons, Methods Mol. Biol. 212 (2003) 111–128.
[15] P.S. Bernard, R.S. Ajioka, J.P. Kushner, C.T. Wittwer, Homogeneous multiplex genotyping of hemochromatosis mutations with fluorescent hybridization probes, Am. J. Pathol. 153 (1998) 1055–1061.
[16] J.A. Nicklas, E. Buel, A real-time multiplex SNP melting assay to discriminate individuals, J. Forensic Sci. 53 (2008) 1316–1324.
[17] L. Zhou, L. Wang, R. Palais, R. Pryor, C.T. Wittwer, High-resolution DNA melting analysis for simultaneous mutation scanning and genotyping in solution, Clin. Chem. 51 (2005) 1770–1777.
[18] L. Zhou, A.N. Myers, J.G. Vandersteen, L. Wang, C.T. Wittwer, Closed-tube genotyping with unlabeled oligonucleotide probes and a saturating DNA dye, Clin. Chem. 50 (2004) 1328–1335.
[19] M. Erali, R. Palais, C. Wittwer, SNP genotyping by unlabeled probe melting analysis, Methods Mol. Biol. 429 (2008) 199–206.
[20] W.M. Howell, M. Jobs, A.J. Brookes, IFRET: an improved fluorescence system for DNA-melting analysis, Genome Res. 12 (2002) 1401–1407.
[21] W.M. Howell, Detection of DNA hybridization using induced fluorescence resonance energy transfer, Methods Mol. Biol. 335 (2006) 33–41.
[22] K. Takatsu, T. Yokomaku, S. Kurata, T. Kanagawa, A FRET-based analysis of SNPs without fluorescent probes, Nucleic Acids Res 32 (2004) e156.
[23] J. SantaLucia Jr., A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics, Proc. Natl. Acad. Sci. USA 95 (1998) 1460–1465.
[24] S. Bommarito, N. Peyret, J. SantaLucia Jr., Thermodynamic parameters for DNA sequences with dangling ends, Nucleic Acids Res. 28 (2000) 1929–1934.
[25] F.N. Hayes, E.H. Lilly, R.L. Ratliff, D.A. Smith, D.L. Williams, Thermal transitions in mixtures of polydeoxyribodinucleotides, Biopolymers 9 (1970) 1105–1117.
[26] M.J. Lane, T. Paner, I. Kashin, B.D. Faldasz, B. Li, F.J. Gallo, A.S. Benight, The thermodynamic advantage of DNA oligonucleotide "stacking hybridization" reactions: energetics of a DNA nick, Nucleic Acids Res. 25 (1997) 611–617.
[27] M.J. Doktycz, M.D. Morris, S.J. Dormady, K.L. Beattie, K.B. Jacobson, Optical melting of 128 octamer DNA duplexes: effects of base pair location and nearest neighbors on thermal stability, J. Biol. Chem. 270 (1995) 8439–8445.