

# Use of extremely short Förster resonance energy transfer probes in real-time polymerase chain reaction

Igor V. Kutuyavin\*

Perpetual Genomics, Woodinville, WA 98077, USA

Received April 4, 2013; Revised August 7, 2013; Accepted August 12, 2013

## ABSTRACT

Described in the article is a new approach for the sequence-specific detection of nucleic acids in real-time polymerase chain reaction (PCR) using fluorescently labeled oligonucleotide probes. The method is based on the production of PCR amplicons, which fold into dumbbell-like secondary structures carrying a specially designed 'probeluring' sequence at their 5' ends. Hybridization of this sequence to a complementary 'anchoring' tail introduced at the 3' end of a fluorescent probe enables the probe to bind to its target during PCR, and the subsequent probe cleavage results in the fluorescence signal. As it has been shown in the study, this amplicon-endorsed and guided formation of the probe-target duplex allows the use of extremely short oligonucleotide probes, up to tetranucleotides in length. In particular, the short length of the fluorescent probes makes possible the development of a 'universal' probe inventory that is relatively small in size but represents all possible sequence variations. The unparalleled cost-effectiveness of the inventory approach is discussed. Despite the short length of the probes, this new method, named Angler real-time PCR, remains highly sequence specific, and the results of the study indicate that it can be effectively used for quantitative PCR and the detection of polymorphic variations.

## INTRODUCTION

The amounts of nucleic acids in test samples are often limited, which makes their direct detection difficult, thus necessitating an initial nucleic acid amplification step. Polymerase chain reaction (PCR) is the most commonly used DNA amplification technique that is capable of multiplying a sequence of interest to a detectable level starting

from as little as a single copy of DNA or RNA (1,2). Fluorimetric detection of PCR products made possible real-time techniques, wherein the amplification is continuously monitored (3,4). The detection results can be obtained in <10 min by applying an optimized PCR protocol and instruments with fast temperature ramping (5). Although the use of fluorescent dyes like SYBR Green is a less expensive and more straightforward way to track the PCR products in real time, this approach is not sequence specific. The sequence-specific detection is provided by the methods using fluorescently labeled oligonucleotide probes. The probes are designed to bind exclusively to a target amplicon and, in most cases, synthesized to carry two dyes that exhibit the Förster Resonance Energy Transfer (FRET) interaction (6). When FRET occurs, emission of one of the fluorescent dyes called the 'reporter' is extinguished by another dye called the 'quencher'. Disruption of FRET by the probe cleavage (7) or distancing effect (8,9) results in a fluorescent signal, and this effect is widely used for nucleic acid detection (10–12).

Despite the considerable progress achieved to date (7–12), the high cost of the FRET probes presently deters the widespread use of these highly specific and accurate molecular tools in real-time PCR. Perhaps, one of the most effective solutions would be the establishment of a complete inventory of FRET probes representing all sequence variations (13), which could be used at on-site genomic centers and universities. The universal library approach allows the distribution of the originally high detection cost across numerous research projects until each individual probe of the library is eventually consumed. However, the preparation of an inventory is only feasible for short oligonucleotide probes. For example, the universal library of all possible tetranucleotide variations would consist of as little as 256 unique sequences. It would not be difficult or expensive to establish and maintain such a relatively small probe inventory, especially with the vision of its unprecedented cost-effectiveness. However, this approach faces two serious problems. First, the short tetranucleotide probes have poor hybridization properties to perform at the elevated

\*To whom correspondence should be addressed. Tel: +1 425 478 2645; Email: perpetualgenomics@comcast.net

temperatures of PCR (>55–60°C). Although the well-known and powerful duplex-stabilizing methods like the conjugated minor groove binders (14,15) and locked nucleic acids (16,17) can reduce the probe length to 10–12, and in rare cases to 8–9-mer oligonucleotides, these technologies are not really helpful in the cases of shorter probes. Second, and perhaps equally challenging, the tetranucleotide probes have to be sequence specific, i.e. bind and generate the signal nowhere else but at the desired amplicon site. Even though any technical solution might seem presently impossible, the method described in this manuscript and named Angler PCR effectively addresses both problems. The preliminary evaluation study indicates that the technology can be used for both key PCR applications, namely, quantitative measurements of nucleic acids (qPCR) and detection of polymorphic variations.

## MATERIALS AND METHODS

### PCR reaction components

JumpStart™ Taq DNA polymerase, an antibody-inactivated ‘hot start’ enzyme, and dNTPs were purchased from Sigma-Aldrich®. All primers and FRET probes were prepared by Cepheid Corporation (Sunnyvale, CA, USA) using reagents and phosphoramidites of the protected nucleosides and dyes such as 6-fluorescein (FAM, cat. No 10-1964), 6-(hexachloro) fluorescein (HEX, cat. No 10-5902) and Epoch Yakima Yellow™ (YY, cat. No 10-5921) from Glen Research (Sterling, VA, USA). A phosphoramidite of Black Hole Quencher (BHQ1, cat. No BNS-5051) was purchased from Biosearch Technologies, Inc (Novato, CA, USA). According to analytical ion exchange high-performance liquid chromatography, the primers and probes used in this study were at least 90% pure. The oligonucleotide 2 μM stock solutions were prepared and stored refrigerated in 5 mM sodium cacodylate, 5 mM Tris–HCl (pH 8.0).

### Real-time instruments, reaction compositions and PCR protocols

A SmartCycler® instrument (Cepheid, Sunnyvale, CA, USA) was used in the proof-of-principle studies and qPCR experiments (Figures 3 and 4). Single nucleotide polymorphism (SNP) genotyping (Figure 5 and Supplementary Figures S9–S11) was conducted on a plate-based Mx3000P QPCR System of Agilent Technologies (Santa Rosa, CA, USA). To ensure reproducibility of the data, the experiments were repeated at least twice. The real-time curves shown in Figures 3 and 4A represent an average of 4–5 identical reactions performed simultaneously. In all studies, background fluorescence was subtracted using instrumental software, and the data, i.e. fluorescence versus PCR cycle, were transferred to an Excel file (Microsoft Office) for further processing and threshold calculations.

### Real-time PCR

The PCR reaction mixtures were prepared by mixing concentrated stock solutions of the reagents at room

temperature. All assays were conducted in an aqueous buffer of 20 mM Tris–HCl (pH 8.0) containing 50 mM KCl and a mixture of all four dNTPs (200 μM each). The FRET probes (Taqman and Angler) were always applied at 200 nM concentration. The concentrations of other key reaction components such as forward and reverse primers, magnesium chloride, JumpStart™ Taq DNA polymerase varied depending on the study and method used. These variations including the PCR cycle profiles are listed in Table 1. Human genomic DNA (GenScript, Piscataway, NJ, USA) was used as a target for detection of the β2-microglobulin gene (GenBank accession #NM004048) in the amounts as indicated in each particular experiment (Figures 3 and 4). Taqman PCR was conducted using the reaction composition and cycle profile previously optimized for this assay (18). Sequences of the primers and probes used in the proof-of-principle (Figure 3) and qPCR experiments (Figure 4) are shown in Figure 1. The final reaction volume in these experiments (SmartCycler) was 25 μl.

### SNP-genotyping study

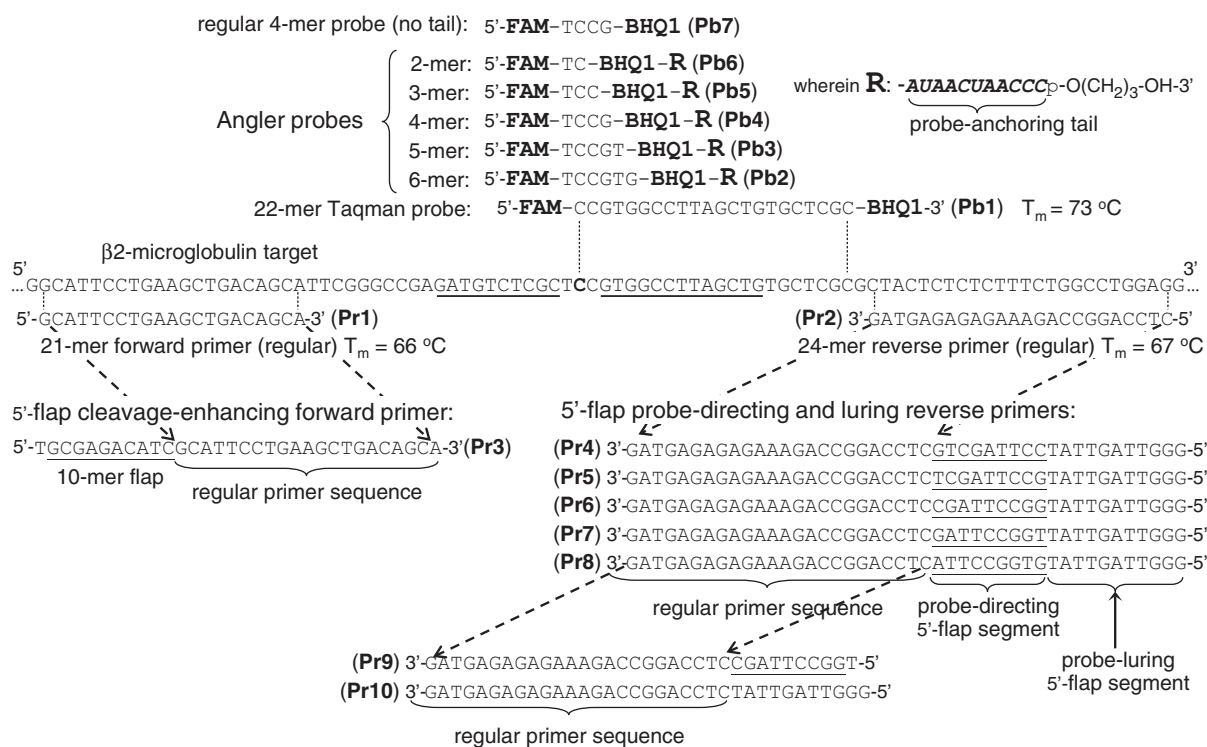
Information about the ten human SNP variations detected in this study and its surrounding sequences (Supplementary Figure S12, Supplementary Material) was taken from the Internet site of the National Cancer Institute (Bethesda, MD, USA) (19). The same source provided information about these SNP allelic frequencies in various sample panels and the genotyping methods applied (Supplementary Table S4). The present study was performed using a Human Random Control DNA Panel from Sigma-Aldrich® (Panel 1 in 96-well plate format, cat. No HRC1-1EA) comprising genomic DNAs from 96 UK Caucasian blood donors. The original DNA solutions were diluted with 10 mM Tris–HCl buffer (pH 8) providing the DNA loads of 33.3 ng/μl and stored in a freezer at –20°C. One microliter of each individual DNA sample along with other PCR components was dispensed correspondingly to a MULTIMAX 96-Well Non-Skirted PCR Plate (cat. No T-3069-1) from BioExpress (Kaysville, UT, USA) to a final volume of 15 μl. Each of the reaction mixtures incorporated a pair of forward and reverse primers, and two Angler FRET probes (200 nM each) labeled with the instrument-distinguishable dyes, i.e. FAM/HEX or FAM/YY. Sequences of these oligonucleotides are shown in Table 2 and Supplementary Table S3. The entire genotyping study (Figure 5 and Supplementary Figures S10 and S11) was performed using the reaction composition and time/temperature profile of the asymmetric Angler method used in the qPCR study (Figure 4). However, taking into account the scale of the intended study, the concentration of the DNA polymerase, which is one of the most expensive PCR components, was reduced to 0.067 U/μl (Table 1). The wells were sealed using 8x Strip Optical Caps from Agilent Technologies (cat. No 401425), briefly agitated, centrifuged (1 min, 160×g) and the plate was placed into the heating block of Mx3000P QPCR System (Agilent Technologies). Real-time curves for each studied SNP are shown in Supplementary Figure S11 (Supplementary Material). At the end of PCR (cycle 40), the fluorescence

**Table 1.** Concentrations of reaction components and PCR cycle profiles used in the project assays

Study	Method <sup>a</sup>	F primer nM	R primer nM	MgCl <sub>2</sub> mM	DNA pol. U/μl	PCR cycle profile <sup>b</sup> of 40 cycles total
Figure 3A and B	sym Angler	100	100	5	0.04	(95°10''→56°40''→ <u>72°15''</u> )
Figure 4A	Taqman	200	200	2	0.04	(95°10''→64°45'')
	sym Angler	100	100	5	0.04	(95°10''→56°40''→ <u>72°15''</u> )
	asym Angler	100	400	5	0.2	(95°10''→56°40''→ <u>72°15''</u> )
Figure 5, Supplementary Figure S9	asym Angler	100	400	5	0.067	(95°10''→56°40''→ <u>72°15''</u> )

<sup>a</sup>The Angler assays were conducted in symmetric (sym) and asymmetric (asym) formats. In the asymmetric format, the reverse (R) primer was applied in 4-fold excess over the forward (F) primer.

<sup>b</sup>In all methods, to activate the antibody-blocked DNA polymerase, the reaction mixtures were incubated at 95°C for 2 min before the temperature cycling (this step is not shown in the cycle profile abbreviation). The Taqman assay was performed using two-step PCR. A three-step cycle profile was applied in all Angler assays. The abbreviation (95°10''→56°40''→72°15'') means that the reaction was incubated at 95°C for 10s, then at 56°C for 40s followed by incubation at 72°C for 15s. Underlined is the step at which the reaction fluorescence was measured.



**Figure 1.** Shown are PCR primers, FRET probes and the detected β2-microglobulin gene sequence used in the experiments of Figures 3 and 4. Wherever it is possible, the oligonucleotides are aligned with the target sequence in 5'→3' orientation as indicated. The primers' flap sequences participating in the amplicon folding and their binding sites in the detected DNA are underlined. All nucleosides in the probe-anchoring sequence **R** were modified with 2-amino deoxyadenosine, 5-propynyl deoxyuridine and 5-propynyl deoxycytidine. **FAM** is 6-fluorescein and **BHQ1** is a Black Hole Quencher 1.

of the individual wells was measured in both dyes' channels providing 2D diagrams, which were consecutively used to determine the haplotypes of all 96 individual DNA samples (Figure 5 and Supplementary Figure S10). Regardless of the dyes used, the photomultiplier of the instrument was always set at X1. Fluorescence of the Yakima Yellow dye was measured in the Cy3 channel.

### Physical measurements and calculations

Oligonucleotide extinction coefficients were calculated using CalcExt2.8 software (Cepheid). Unless otherwise indicated, melting temperatures of unmodified primers

and probes were calculated using the 'nearest-neighbour' approach (20,21) for perfect-match duplexes at 200 nM with adjustment for the PCR buffer used in this study. The effect of dyes on duplex stability was disregarded in calculations.

### PCR threshold calculations

The real-time data (fluorescence versus PCR cycle) collected from the instruments were transformed into a semi-log format. Using a Microsoft Office option, a linear trendline equation was obtained for the first four data points with the logarithm values >2, i.e.  $\log(F) = aCt + b$  rearranged to  $Ct = [\log(F) + b]/a$ . Based

**Table 2.** Structures of primers and probes used in the SNP-genotyping study of Figure 5

SNP	PCR function	Oligonucleotide sequence and modifications <sup>a</sup>
rs4680	F primer	5'-TCCAGCGAAATACCCAGCGGATGGTGGA-3'
	R primer	5'-CGGCGGCGATATAGGACAAGGTCAACGGGTCAGGCATGCA-3'
	Probe 1	5'-(FAM)-C <u>ATG</u> -(BHQ1)-ATATCGCCGCCGp-O(CH <sub>2</sub> ) <sub>3</sub> -OH-3'
	Probe 2	5'-(HEX)-C <u>GTG</u> -(BHQ1)-ATATCGCCGCCGp-O(CH <sub>2</sub> ) <sub>3</sub> -OH-3'
rs9643226	F primer	5'-ATTTGTGATCTGAGTCATAGACCTGGGGGCA-3'
	R primer	5'-CGGCGGCGTTTATTCTAGGTCATAGCTCAGAGTAAAAGCCAAAGTCTCA-3'
	Probe 1	5'-(FAM)-GCGT-(BHQ1)-TAAACGCCGCCGp-O(CH <sub>2</sub> ) <sub>3</sub> -OH-3'
	Probe 2	5'-(YY)-CGGT-(BHQ1)-TAAACGCCGCCGp-O(CH <sub>2</sub> ) <sub>3</sub> -OH-3'
rs10505477	F primer	5'-AGGAGATGAAGTCAATACTGACTTTGCCCTTTTCTAAAT-3'
	R primer	5'-CGGCGGCGTTAAAGGCTTCCTGGTCATGTGTACCACCTGTCTATCAA-3'
	Probe 1	5'-(FAM)-A <u>ATT</u> -(BHQ1)-TTAACGCCGCCGp-O(CH <sub>2</sub> ) <sub>3</sub> -OH-3'
	Probe 2	5'-(YY)-A <u>GT</u> -(BHQ1)-TTAACGCCGCCGp-O(CH <sub>2</sub> ) <sub>3</sub> -OH-3'

<sup>a</sup>The cleavage-enhancing flap sequences in forward (F) primers and the probe-directing sequences in reverse (R) primers are underlined. The 2-amino deoxyadenosine and 5-propynyl deoxyuridine duplex-stabilizing modifications used in design of the A/T-rich Angler probes are marked by bold italic font. **FAM** is 6-fluorescein, **HEX** is 6-(4, 7, 2', 4', 5', 7'-hexachloro) fluorescein and **YY** is Yakima Yellow fluorescent dye (Glen Research). **BHQ1** is Black Hole Quencher 1 from Biosearch Technologies, Inc. In all cases, the second nucleotide from the 5' end of the FRET probes (underlined) was the SNP detecting one. All probes incorporated a 1,3-propanediol moiety conjugated to the 3'-terminal phosphate to prevent the accidental extension during PCR.

on this linear trendline equation, the fluorescent curve threshold value (Ct) was determined as the cycle number at which the logarithm of fluorescence  $\log(F) = 2$ . This corresponds to 100 fluorescence units in the conventional format as shown in Figure 4A.

### PCR yield

An average PCR cycle yield  $Y$  (%) for the studied assays was calculated using the equation  $Y = 100 \cdot [\text{antilog}(1/a) - 1]\%$ , wherein  $a$  is the slope coefficient (positive value) taken from the corresponding linear trendline equations obtained for each of the assays after the target-titration study shown in Figure 4B. The reported yield values were rounded to an integer.

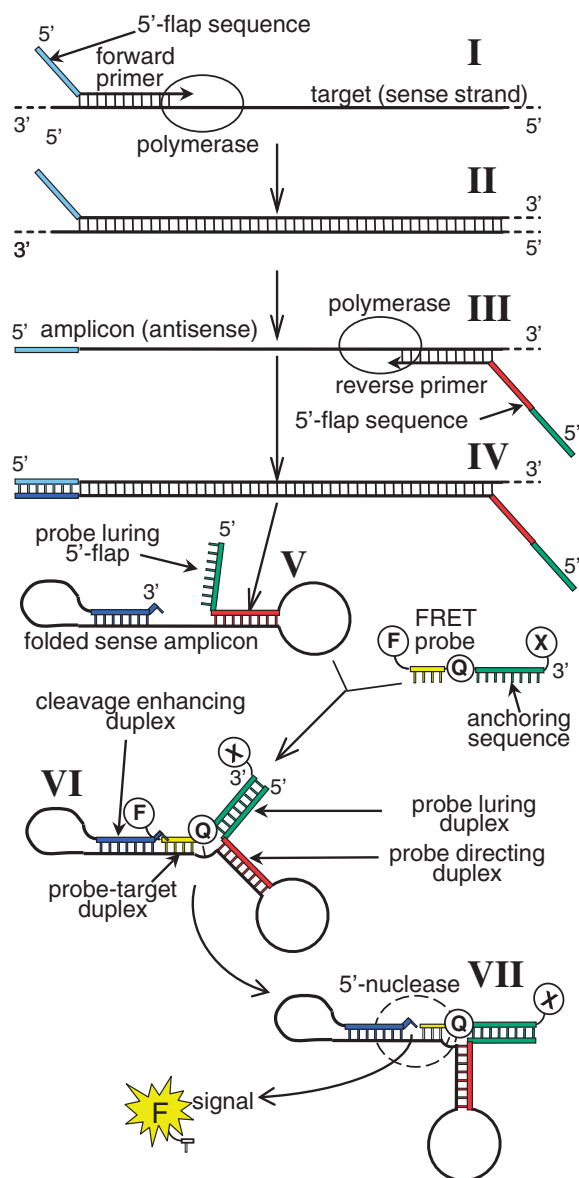
## RESULTS

### Angler real-time PCR: the mechanism of action

All oligonucleotide components in Angler PCR, the primers and probes, carry special tail sequences that are critical for the assay performance. Figure 2 illustrates how the Angler system functions. Similar to the Snake technology (18), the forward PCR primer contains a 5'-flap sequence (light blue) called the 'cleavage-enhancing' flap, which is complementary, save for the 5'-terminal base, to the product of this primer extension, downstream from the original primer sequence. Extension of this forward primer in stage **I** results in the synthesis of an antisense strand, providing a double-stranded amplicon **II**. After strand separation (95°C), a reverse primer hybridizes to the antisense strand and DNA polymerase extends the complex (stage **III**), resulting in yet another double-stranded amplicon **IV**. The extension of the reverse primer generates a replica of the 5'-flap sequence of the forward primer at the 3' end of the sense strand (dark blue). The reverse primer also incorporates a 5'-flap sequence. This flap sequence comprises two segments. The first segment (red), named 'probe-directing', is

designed to fold the sense amplicon into a stem-loop structure. The second segment (green), named 'probe-luring', is an artificial nucleotide sequence that is designed to be complementary to a corresponding anchoring sequence (also green) conjugated to the 3' end of a FRET probe. After strand separation of the double-stranded amplicon **IV**, the sense amplicon folds into a dumbbell-like secondary structure **V**. The short probe (yellow), except its anchoring 3' tail, is complementary to the target sequence of the sense amplicon located between the cleavage-enhancing and probe-directing duplexes. Hybridization of the probe to the folded sense amplicon results in the formation of a three-way DNA junction **VI**, or Y-structure for short. The flap and tail sequences in the primers and probe are selected in length and base composition to make all three key duplexes in the structure **VI** relatively stable at the PCR detection temperature.

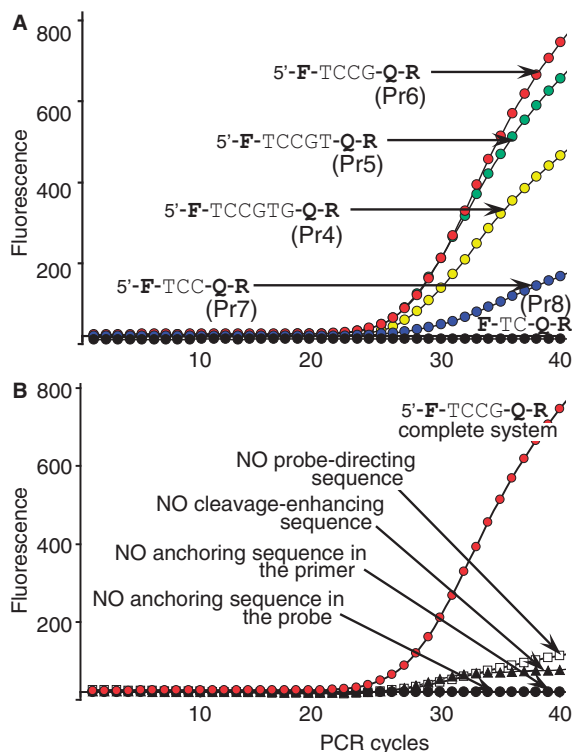
Especially important is the stability of the probe-luring duplex because it controls the first step of the detection reaction when the folded amplicon 'attracts' the probe for its initial binding. On one hand, the amplicon-luring flaps with low hybridization properties would not serve their main purpose. On the other hand, if the luring duplex is too stable, it would negatively affect cycling of the probe hybridization and cleavage reactions, resulting in reduction of the assay signal. Ideally, the probe-luring duplexes should have melting temperatures close to the PCR detection temperature. The formation of the luring duplex in the first step makes the second step of the reaction, the probe binding to its target, an intramolecular process and results in the rapid confinement of the probe by the amplicon. The hybridized and entrapped probe is contained between the amplicon duplex 'arms'. However, a single nucleotide mismatch within the probe binding site helps the probe to escape the molecular trap. This molecular mechanism was found to be so powerful that it enables the use of extremely short tetranucleotide probes in real-time PCR. In this manuscript, the length of the probes is



**Figure 2.** Mechanism of the signal generation in Angler PCR. The method is based on the production of PCR amplicons, which fold into dumbbell-like secondary structures carrying a specially designed 'probe-luring' sequence at their 5' ends. Hybridization of this sequence to a complementary 'anchoring' tail introduced at the 3' end of a fluorescent probe enables the short probe to bind to its target during PCR, and the subsequent probe cleavage results in the fluorescence signal. The important structural elements of the system design are individually colored. The mechanism is discussed in the text.

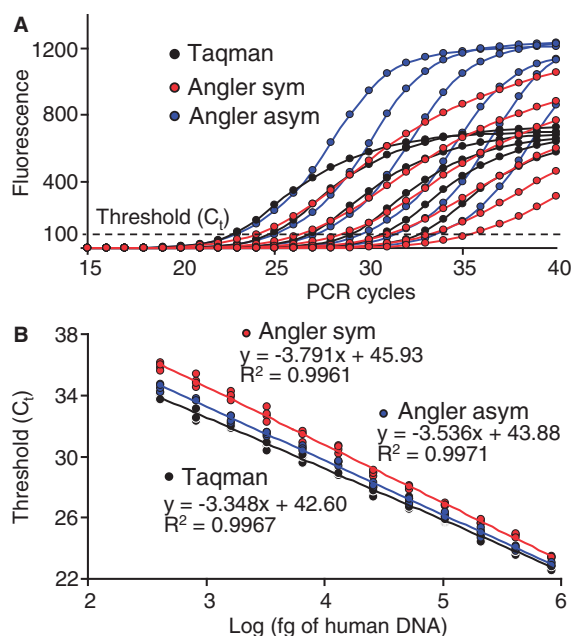
defined by the nucleotide sequence between the FRET dyes; the 3'-nucleotide tail is not counted.

The cleavage-enhancing duplex of the folded amplicon and the probe-target duplex have to be in a coaxial stacking interaction to ensure the formation of an optimal cleavage structure for the Taq 5'-nuclease domain. The 5' nuclease actually requires the cleaved downstream duplex to be  $\geq 10$  bp (22). For this reason, in the case of short FRET probes, the Y-structure may not be the optimal substrate for the 5' nuclease, but it can rearrange into a T-like structure **VII** wherein the probe



**Figure 3.** Results of the Angler feasibility study. (A) Shows the performance of the Angler probes of variable length. Each of the investigated probes Pb2-Pb6 was applied in the reaction with the 5'-flapped forward primer Pr3, which was paired with one of the reverse primers Pr4-Pr8. The reverse primers were designed individually for each probe to form the Y-structure **VI** as illustrated in Figure 2. The primer number shown in parentheses next to the probe is the one used in the particular assay. The results of diagram (B) emphasize the importance of the special flap and tail sequences in the Angler primers and probes. The PCR experiments were performed using the tetranucleotide Angler probe (complete system, Pb4+Pr3+Pr6). However, when a component like the cleavage-enhancing flap in the forward primer (Pr1), probe-luring (Pr9) or probe-directing (Pr10) sequence in the reverse primer or probe-anchoring 3' tail in the probe (Pb7) is removed from the complete system, the real-time fluorescence curves showed a weak or lacked signal. Sequences of all primers, probes and DNA target are shown in Figure 1. The reaction mixtures contained 105 ng of human genomic DNA (GenScript). The rest of the experimental details are provided in the 'Materials and Methods' section and Table 1.

and luring duplexes are coaxially oriented, simulating, for the enzyme, a fairly long 5'-cleavage substrate. Once the 5' nuclease recognizes either structure **VI** or **VII**, it cleaves the probe between the first and second nucleotides from the 5' end (18,22). The probe cleavage releases a reporting dye **F** from the FRET interaction with a quencher **Q**, generating a detectable fluorescence signal. The FRET probe carries a small moiety **X** at the 3' end of the anchoring sequence. It can be of any chemical structure as long as it prevents the probe extension during PCR. To release structural tension at the Y-structure (23,24), the Angler assays were designed such that the catalytic sense amplicon has one unpaired nucleotide between the probe-target and probe-directing duplexes. This rule had been observed throughout the study, although this may not be necessary because the quencher moiety itself, by

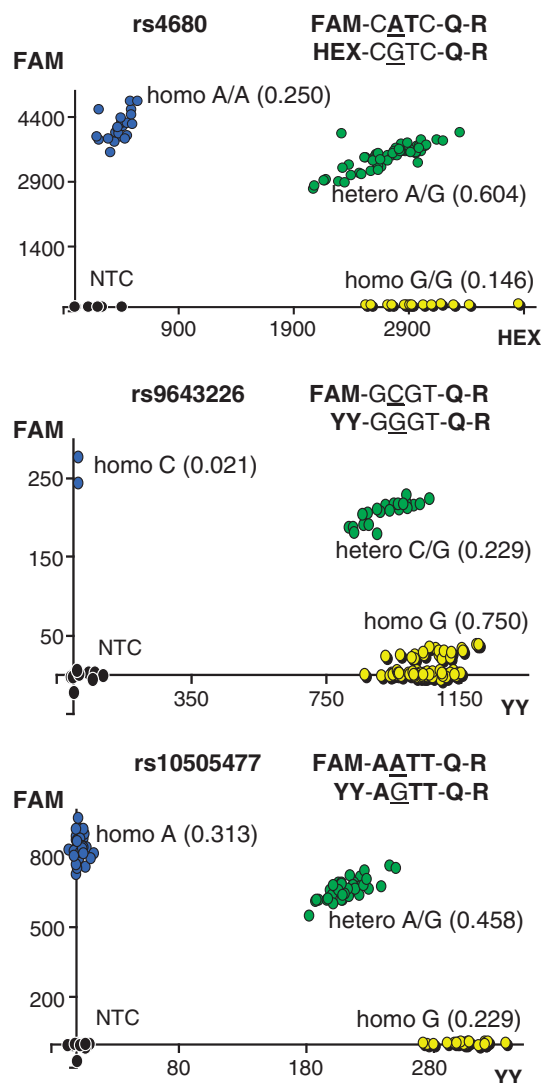


**Figure 4.** Validation of the Angler assay for quantitative PCR (qPCR). (A) Shows the fluorescence curves obtained for the Angler assay (4-mer FRET probe Pb4) in symmetric (red, Pr3/Pr6 100/100 nM) and asymmetric (blue, Pr3/Pr6 100/400 nM) formats. The black curves are for conventional Taqman (22-mer FRET probe Pb1, Pr1+Pr2). The human DNA amounts in the reactions ranged from 840 (left curve) to 0.41 ng (right curve for each assay) with 2-fold dilutions, but shown in A are only 4-fold dilutions. The threshold values ( $C_t$ ) were determined for each of the curves at 100 fluorescence units (dashed line in A) and plotted versus the logarithm of the target loads in (B). The target dilution experiments were performed in 4–5 paralleled reactions. (B) Shows the linear trendline equations and coefficients of determination ( $R^2$ ) calculated for each of the studied assays. Sequences of the detected human  $\beta$ 2-microglobulin gene, primers and probes are shown in Figure 1. The rest of the experimental details are provided in the ‘Materials and Methods’ section and Table 1.

its location, is anticipated to be an effective tension-releasing factor. Actually, the preliminary results of Supplementary Figure S7 discussed in P1 (Supplementary Material) point to a structural flexibility in the design of the Angler three-way junction.

### Angler technology feasibility study

The proof-of-principle experiments were performed by detecting a fragment of the human  $\beta$ 2-microglobulin gene. The structures of all primers and probes used in this study are shown in Figure 1. All three duplexes that form the Y-structure VI (Figure 2) contribute to its thermal stability, but these duplexes do not need to be equally stable. For example, if the probe-directing and probe-luring duplexes are designed to sustain PCR temperatures, then the probe-target duplex can be relatively short and unstable, but the Y-structure would still form. The experiments shown in Figure 3A reveal that the Angler probe can be short. Actually, the 4-mer FRET probe (Pb4) demonstrates the best performance as the fluorescence signal begins to decline with the increase in the probe length to 5–6-mers (Pb3 and Pb2). A weak, but still detectable, signal was generated using the trinucleotide probe Pb5, whereas the



**Figure 5.** Detection of SNP variations using tetranucleotide-long Angler probes. Shown are the scatter plots for three SNPs found in human genomic DNA. The polymorphism numbers as well as the abbreviated probe structures are listed on the top of each diagram. The SNP-discriminating nucleotides are underlined. FAM, HEX and YY are abbreviations of the fluorescent dyes used in the study. Q is a quenching dye, whereas R is a 3'-anchoring sequence. Full sequences of the probes and the primers used in each SNP case are listed in Table 2 and also illustrated in Supplementary Figure S12. No target controls (NTC), the homozygous (homo) and heterozygous (hetero) signal groups are marked by a specific color. All experimental details can be found in ‘Materials and Methods’ section and Table 1.

dinucleotide probe Pb6 is apparently too short to perform in the Angler assay.

Use of the extremely short FRET probes raises a concern about the sequence specificity of the detection. Reduction of the oligonucleotide length to essentially a few nucleotides greatly increases the chances of the probe binding somewhere else other than the targeted site. This unintended probe hybridization may jeopardize the assay detection specificity. However, the results of Figure 3B indicate that this is unlikely to be the case. In Angler PCR, the detected target sequence is actually longer than the probe. In addition to the probe binding

site (4-mer), it incorporates the binding sites of the cleavage-enhancing flap (~8–11-mer) and the probe-directing sequence (~8–11-mer) resulting overall in a fairly long ~20–26-mer sequence, comparable in length with those of the conventional assays (7–11). For example, in the case of the 4-mer probe Pb4, lack of the cleavage-enhancing flap (Pr1) or the probe-directing sequence (Pr10) in the primer designs almost completely eliminate the assay signal response. But even in these cases, the specifically detected sequences remain reasonably long (13–14-mers). The absence of the anchoring sequence in the probe (Pb7) or the luring flap sequence in the reverse primer (Pr9) completely abolishes the signal.

Because of the relatively short amplicons (~100–200 bp), many conventional assays like Taqman are commonly performed using two-step PCR wherein the annealing and extension stages are combined into one step. However, this is not an option for the Angler assay. Folding of the PCR amplicons slows down the primer hybridization (25,26), and this was shown to negatively affect PCR even in a less complex system like Snake (18,27). For this reason, the PCR profile used in this study comprised three steps. The amplicon denaturation step (95°C, 10 s) was followed by a detection step (56°C, 40 s) at which the amplicons are folded and catalyze the probe cleavage, and then an extension step (72°C, 15 s) at which the amplicons unfold, enabling the primers to readily hybridize and get extended by DNA polymerase. The Taq polymerase expresses its maximum activity at temperatures >70°C (28,29), and 15 s was found to be sufficient time to complete the PCR replication. Accordingly, throughout this study, the primers were designed to address such an elevated temperature. This was not really challenging because of the additional duplex stabilization provided by the 5'-flap sequences present in both primers.

### Angler quantitative PCR

With regard to the quantitative PCR application, the Angler technology was anticipated to maintain the effectiveness and accuracy provided by all other conventional assays. To confirm this, the Angler and Taqman assays were compared head-to-head in a genomic DNA dilution study (Figure 4). Reflecting the system's complexity, the Angler assay detection is delayed by a few cycles (Figure 4B), but most importantly it upholds the linear dependence between the threshold values and the logarithm of the target loads ( $R^2 > 0.99$ ). The Angler predecessor, the Snake system was found to positively respond to an excess of the reverse primer (18,27) in asymmetric PCR (30). The Angler assay apparently follows the same trend. The calculations based on the linear slope coefficients (Figure 4B) show improvement of the average PCR yield from 84% for symmetric Angler to 92% for asymmetric Angler, which still falls short of the Taqman efficiency of 99%. Actually, the PCR yield improvement in the asymmetric condition narrowed the sensitivity gap between the Angler and Taqman assays to roughly one cycle. Moreover, the asymmetry in the primer concentrations greatly enhanced the Angler fluorescence

productivity. At the same concentration of the FRET probes (200 nM), the Angler assay produced twice as much signal in comparison with Taqman (Figure 4A).

### Detection of polymorphic variations in angler PCR

In the most common cases of two polymorphic variants, the assays for detection of the sequence variations comprise two PCR primers and two FRET probes labeled with different reporting dyes. The probes are designed to incorporate the polymorphic site and they are usually homologous in sequence, with the exception of the nucleotides discriminating the polymorphism. Use of the Angler assay imposes an additional requirement on the system design. Both probes have to be of the same length and carry identical 3'-anchoring tails. Examples of the Taqman and Angler system designs for detection of the polymorphic variations are shown in Supplementary Figure S8 and discussed in P2 (Supplementary Material).

There are a number of factors that influence the ability of any probe-based PCR to identify polymorphic variations, particularly those as small as single nucleotides. These factors are discussed in P3 and illustrated in Supplementary Figure S9 (Supplementary Material). The probe length is one of the most influential factors: SNP discrimination improves with the reduction of probe length. In Angler, the luring sequences are separated from the FRET probes by the quencher moiety, and therefore have little, if any, effect on the probes' ability to discriminate between matched and mismatched binding sites. As the Angler probes are the shortest ones ever tested, no problem with the detection specificity was anticipated. Selection of the SNP variations for the test trial was conducted based on the following criteria. First, preference was given to the polymorphisms that had already been genotyped by other techniques. This helped to avoid extremely rare SNPs (frequencies <0.05). Second, the SNP selection aimed to cover the broad range of the nucleotide sequences surrounding the polymorphisms, including highly A/T- and G/C-rich sites.

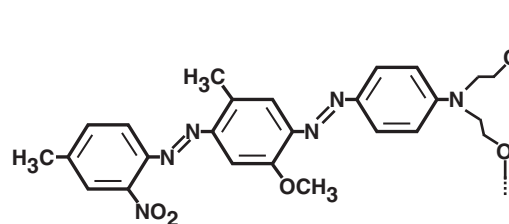
The difference between the hybridization properties of the A/T- and G/C-rich oligonucleotides rapidly grows with the oligonucleotide length reduction. Owing to this problem, all Angler probes comprising  $\geq 2$  A and T nucleotides were base-modified using the duplex-stabilizing 2-amino deoxyadenosine and 5-propynyl deoxyuridine nucleoside derivatives. The Taq 5' nuclease was found to be especially susceptible to the mismatches located at the second nucleotide from the 5' end of the cleaved strand (18). Snake and Angler are related technologies, and this determined the choice of the probe design throughout the study. However, at such short tetranucleotide probe lengths, other mismatch locations could be also highly discriminatory. In an attempt to promote the probe cleavage cycling, the corresponding anchoring and luring sequences were designed to form the duplexes with melting temperatures close to the detection temperature of 56°C. The probe sequences are likely to contribute to the Y- and T-structure stabilities (Figure 2), but this was ignored. Ten SNPs found in human genomic DNA were selected for the

test trial and successfully genotyped. The results obtained from two stand-alone experiments including the individual sample genotypes are shown in Supplementary Figure S10 (Supplementary Material). A brief summary of the results for three polymorphic variations is also shown in Figure 5. Although the fluorescence monitoring was performed during all PCR cycles (Supplementary Figure S11), the real-time method is not necessary in genotyping. Therefore, the fluorescence was measured at the end of PCR (cycle 40) and plotted as a function of the signals in two reporting dye channels for every one of the 96 reactions. As anticipated, the tetranucleotide-long Angler probes discriminated the studied SNPs in an absolute 'yes-or-no' fashion with the homozygous signals located at or close to the corresponding axis. Three groups identifying the allelic combinations are well clustered and separated with minimum outliers to ensure unambiguous genotyping. The SNP allelic frequencies determined in this study correlate well with those from other surveys (Supplementary Table S4 in Supplementary Material).

## DISCUSSION

The Angler probes are relatively short and simple molecules. They can be manufactured as easily as any other FRET probes for conventional assays. An increase in the primers' length, because of the 5'-flap sequences, also does not represent a technical challenge these days. The present data indicate that the technology has successfully passed its initial development and has proven its effectiveness for major real-time applications such as qPCR and the detection of polymorphic variations. Although it may seem that no adjustments are needed, the system has not yet been completely optimized, and it can be improved in a number of ways. For example, the assay signal productivity can be increased by supplementing the PCR reaction with a flap endonuclease. The same effect can be reached by optimizing the probe hybridization properties to promote rapid cycling in probe binding and cleavage reactions. The most influential factor in the probes' hybridization properties is the anchoring tail, and therefore the anchoring tail design needs to be finalized. The use of the expensive base modifications in the tails like the one shown in Figure 1 is absolutely unnecessary. An improved tail design was applied in the SNP genotyping study (Tables 2 and Supplementary Table S3). However, the best anchoring tails are yet to be tested, and the corresponding proposal is discussed in the Supplementary Material of the manuscript (see P4, Supplementary Table S5 in Supplementary Material).

Perhaps the most promising improvement of the Angler assay is in the redesign of the FRET probes. The positive signal response observed for the TCC-trinucleotide probe Pb5 in the experiment of Figure 3A looks encouraging. Intercalating drugs are known for their strong binding to the structural duplex abnormalities and, in particular, the junctions of branched DNAs (31). The presently used BHQ1 quencher moiety is located right at the DNA junction (structure VI, Figure 2) and it certainly helps to release its structural tension. However, this quencher, the



**Figure 6.** Molecular structure of BHQ1 quencher used in the design of the Angler FRET probes.

structure of which is shown in Figure 6, does not look like an intercalator, but rather a minor groove binder. In this aspect, linking the FRET probe to its anchoring tail through a strong intercalating moiety may further stabilize and promote the T-structure formation (structure VII, Figure 2). The expanded polyaromatic systems covering an entire DNA base pair (32) and positively charged intercalators like, e.g. N-(2-hydroxyethyl)phenazinium derivatives (33) represent a special interest. If the desired stabilization effect is achieved, it may be possible to reduce the probe length to trinucleotides, and/or avoid using the expensive base modifications in the design of the A/T-rich probes. In either case, it would be a substantial improvement to the technology. The quencher can be located at the 3' end of the anchoring tail, wherein it would also serve the purpose of the presently used 1,3-propanediol moiety, i.e. blocking the polymerase extension during PCR.

The main advantage of the Angler technology is not in its signal productivity, SNP-discrimination capability or any other real-time characteristics. Unlike all known conventional assays, the Angler method is approaching, if it has not already done so, the universal library format. A complete inventory of all possible tetranucleotide variations comprises 256 sequences. However, the real library can be smaller, possibly much smaller. For example, similar to Roche's inventory (P5, Supplementary Material), it can comprise just a few dozen 4-mer FRET probes, if it is solely designated for the qPCR detection of reasonably long target sequences, which provide an ample opportunity for the discovery of a corresponding binding site for one of the inventory probes. However, to be a true universal library, the inventory must also cover the detection of polymorphic variations. Unfortunately, in this case there is no flexibility in the probe selection because the probe-binding site must incorporate the polymorphic nucleotides. Nevertheless, the practically 'universal' inventory of the 4-mer probes can be still cut in half to 128 sequence variations because the detecting probes can be addressed to either one of two DNA duplex strands. If the probe redesign to the trinucleotide sequences proves to be effective (see earlier in the text), the corresponding estimates would translate to as few as 32 members in the library. Actually, all of these library size estimates look too small and affordable in light of the presently established and marketed FRET probe inventory for qPCR analysis of nearly 200 000 human genes (Life Technologies, see P5 in the Supplementary Material).

The universal library approach has little, if any, advantage for industrial molecular diagnostics wherein the



probes are completely consumed during the numerous tests. As for the research community, the situation is different. A majority of the researchers need to analyze from a few to a few hundred samples using a small fraction of the expensive FRET probes. This is the case wherein the unprecedented cost-effectiveness of the universal library approach makes a major impact. Although considerable progress has been achieved in the performance of real-time PCR, the high cost of the FRET probes represents a serious issue. Academic research budgets are practically always under financial stress, especially during the present tough economic times. This drives many researchers to use unreliable molecular methods like the SYBR Green detection for real-time PCR. SYBR Green is less expensive but not sequence specific. The market price of the sequence-specific FRET probes presently ranges from ~\$230–260 (Taqman) to \$375–450 per probe (Molecular Beacons and Scorpions). Actually, this price range is well justified and is mainly defined by the high manufacturing cost of the FRET probes compared with the PCR primers, which can be obtained for ~\$12–15 per pair. Synthesis scale reduction below the presently provided minimum of ~5–10 nmole would not affect much the probe prices because, unlike the PCR primers, the FRET probes require thorough purification and this is the most cost-defining procedure.

Calculations show that a 10 nmole amount of a FRET probe is sufficient to run ~5000 detection reactions in 10  $\mu$ l volume with 200 nM probe concentration. However, the majority of researchers 'do not need' to run more than few hundred reactions with no need to detect the same sequences in the foreseeable future. This means that ~90–98% of the original probe amount is going to be left unused or simply wasted. In contrast to the inexpensive primers, routine wasting of the expensive FRET probes is a serious problem. If the manufacturer could sell FRET probes fractionally, in amounts satisfying the majority of the researchers, the actual cost of the detection components could drop by an order of magnitude or more, making the FRET probes as inexpensive as the amplification components, the PCR primers. Unfortunately, this is only possible in the cases of commonly used targets. Examples of the presently established market inventories enabling the fractional sale are discussed in P5 (Supplementary Material). In a majority of the cases, a fractional sale is not an option and the manufacturer must recover the entire cost of a probe at the time of the sale. It is impossible to estimate the number and amounts of the FRET probes that have been disposed of or are presently 'decaying' in laboratory freezers. The numbers must be considerable because the probe-based PCR methods remain the most reliable and accurate research tools for the detection and analysis of nucleic acids. Unlike the presently established inventories (P5), a universal probe library would solve the problem completely, enabling the fractional sale or use of the probes for all possible PCR applications and target sequence varieties until each probe of the library is completely consumed. The probes themselves remain expensive, but their presently high cost gets distributed across thousands of PCR reactions. The sequence-specific and reliable

probe-based detection would become as affordable as the SYBR Green-like detection assays.

If the Angler assay is going to eventually outcompete the traditional technologies, it will be specifically because of its economics. The estimated increase in the Angler primers' cost by ~\$6–8 looks negligible in comparison with the significant savings provided by the complete FRET probes inventory. It is needless to say that the popularization of this new method would require an assay-designated computer program, which would choose the DNA strand to be detected, design and select the FRET probes with their individual anchoring tails, add the necessary flap sequences to the primers, avoid amplicon secondary structures that are strong enough to compete with the Angler specific folding, etc. It should not be difficult to add all these assay-specific design capabilities to the computer programs that are presently used for the conventional assays like Taqman and Molecular Beacon. In the meantime, the PCR primers can be selected using any available software and then reinforced with the 5'-flap sequences following the simple and straightforward rules described in P6 and exemplified in Supplementary Table S6 of Supplementary Material.

When compared with the conventional assays in its presently published form, the Angler method has a number of disadvantages. This includes relatively complex system design, longer PCR primers and use of the modified nucleotides in A/T-rich probes. The three-step PCR cycle versus two-step one commonly applied in other assays can be hardly called a disadvantage, as the overall cycle time remains close to 1 min. Further development of the method including the Angler probe redesign and assay-designated software (see earlier in the text) may help to solve many of these, rather minor, issues.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

The author thanks Cepheid's employees, particularly Alexander Gall, Sergey Lokhov and Ekaterina Viazovkina for interest, technical help and support. The assistance of William Brabant (Cepheid Corporation) and Vassily Kutyavin (Fred Hutchinson Cancer Research Center, WA, USA) in the editing and revision of the manuscript is gratefully acknowledged.

## FUNDING

National Institute of Health [SBIR grant 1R43GM093446-01]; Cepheid Corporation (Sunnyvale, CA, USA). The company provided all reagents for the research and access to the necessary facilities and instrumentation. Funding for open access charge: Personal funds.

*Conflict of interest statement.* None declared.

## REFERENCES

- Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, **230**, 1350–1354.
- Saiki, R.K., Bugawan, T.L., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1986) Analysis of enzymatically amplified beta-globin and HLA-DQ alpha DNA with allele-specific oligonucleotide probes. *Nature*, **324**, 163–166.
- Higuchi, R., Dollinger, G., Walsh, P.S. and Griffith, R. (1992) Simultaneous amplification and detection of specific DNA sequences. *Biotechnology*, **10**, 413–417.
- Higuchi, R., Fockler, C., Dollinger, G. and Watson, R. (1993) Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology*, **11**, 1026–1030.
- Belgrader, P., Benett, W., Hadley, D., Richards, J., Stratton, P., Mariella, R. Jr and Milanovich, F. (1999) PCR detection of bacteria in seven minutes. *Science*, **284**, 449–450.
- Clegg, R.M. (1992) Fluorescence resonance energy transfer and nucleic acids. *Methods Enzymol.*, **211**, 353–388.
- Lie, Y.S. and Petropoulos, C.J. (1998) Advances in quantitative PCR technology: 5' nuclease assays. *Curr. Opin. Biotechnol.*, **9**, 43–48.
- Tyagi, S. and Kramer, F.R. (1996) Molecular beacons: probes that fluoresce upon hybridization. *Nat. Biotechnol.*, **14**, 303–308.
- Whitcombe, D., Theaker, J., Guy, S.P., Brown, T. and Little, S. (1999) Detection of PCR products using self-probing amplicons and fluorescence. *Nat. Biotechnol.*, **17**, 804–807.
- Didenko, V.V. (2001) DNA probes using fluorescence resonance energy transfer (FRET): designs and applications. *Biotechniques*, **31**, 1106–1121.
- Walker, N.J. (2002) A technique whose time has come. *Science*, **296**, 557–559.
- Ranasinghe, R.T. and Brown, T. (2005) Fluorescence based strategies for genetic analysis. *Chem. Commun. (Camb.)*, **44**, 5487–5502.
- Drmanac, R., Strezoska, Z., Labat, I., Drmanac, S. and Crkvenjakov, R. (1990) Reliable hybridization of oligonucleotides as short as six nucleotides. *DNA Cell Biol.*, **9**, 527–534.
- Sinyakov, A.N., Lokhov, S.G., Kutyavin, I.V., Gamper, H.B. and Meyer, R.B. (1995) Exceptional and selective stabilization of A-T Rich DNA/DNA duplexes by N-methylpyrrole carboxamide peptides conjugated to oligodeoxynucleotides. *J. Am. Chem. Soc.*, **117**, 4995–4996.
- Lukhtanov, E.A., Kutyavin, I.V., Gamper, H.B. and Meyer, R.B. Jr (1995) Oligonucleotides with conjugated dihydropyrroloindole oligopeptides: preparation and hybridization properties. *Bioconjugate Chem.*, **6**, 418–426.
- Koshkin, A.A., Nielsen, P., Meldgaard, M., Rajwansi, V.K., Singh, S.K. and Wengel, J. (1998) LNA (Locked Nucleic Acid): an RNA mimic forming exceedingly stable LNA:LNA duplexes. *J. Am. Chem. Soc.*, **120**, 13252–13253.
- Gamper, H.B. Jr, Arar, K., Gewirtz, A. and Hou, Y.M. (2005) Unrestricted accessibility of short oligonucleotides to RNA. *RNA*, **11**, 1441–1447.
- Kutyavin, I.V. (2010) New approach to real-time nucleic acids detection: folding polymerase chain reaction amplicons into a secondary structure to improve cleavage of Forster resonance energy transfer probes in 5'-nuclease assays. *Nucleic Acids Res.*, **38**, e29.
- <http://variantgps.nci.nih.gov/cgfseq/pages/sequenceSubmit.do?method=sequence&regionId=1>.
- Breslauer, K.J., Frank, R., Blocker, H. and Marky, L.A. (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl Acad. Sci. USA*, **83**, 3746–3750.
- SantaLucia, J. Jr (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl Acad. Sci. USA*, **95**, 1460–1465.
- Kaiser, M.W., Lyamicheva, N., Ma, W., Miller, C., Neri, B., Fors, L. and Lyamichev, V.I. (1999) A comparison of eubacterial and archaeal structure-specific 5'-exonucleases. *J. Biol. Chem.*, **274**, 21387–21394.
- Leontis, N.B., Kwok, W. and Newman, J.S. (1991) Stability and structure of three-way DNA junctions containing unpaired nucleotides. *Nucleic Acids Res.*, **25**, 759–766.
- Shlyakhtenko, L.S., Appella, E., Harrington, R.E., Kutyavin, I. and Lyubchenko, Y.L. (1994) Structure of three-way DNA junctions. 2. Effect of extra bases and mismatches. *J. Biomol. Struct. Dyn.*, **12**, 131–143.
- Uhlenbeck, O.C. (1972) Complementary oligonucleotide binding to transfer RNA. *J. Mol. Biol.*, **65**, 25–41.
- Lima, W.F., Monia, B.P., Ecker, D.J. and Freier, S.M. (1992) Implication of RNA structure on antisense oligonucleotide hybridization kinetics. *Biochemistry*, **31**, 12055–12061.
- Kutyavin, I.V. (2011) Use of base modifications in primers and amplicons to improve nucleic acids detection in the real-time snake polymerase chain reaction. *Assay Drug Dev. Technol.*, **9**, 58–68.
- Innis, M.A., Myambo, K.B., Gelfand, D.H. and Brow, M.A. (1988) DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc. Natl Acad. Sci. USA*, **85**, 9436–9440.
- Takagi, M., Nishioka, M., Kakihara, H., Kitabayashi, M., Inoue, H., Kawakami, B., Oka, M. and Imanaka, T. (1997) Characterization of DNA polymerase from *Pyrococcus* sp. Strain KOD1 and its application to PCR. *Appl. Environ. Microbiol.*, **63**, 4504–4510.
- Gyllenstein, U.B. and Erlich, H.A. (1988) 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**, 7652–7656.
- Guo, Q., Seeman, N.C. and Kallenbach, N.R. (1989) Site-specific interaction of intercalating drugs with a branched DNA molecule. *Biochemistry*, **28**, 2355–2359.
- Vinogradov, S., Roig, V., Sergueeva, Z., Nguyen, C.H., Arimondo, P., Thuong, N.T., Bisagni, E., Sun, J.S., Helene, C. and Asseline, U. (2003) Synthesis and binding properties of oligo-2'-deoxyribonucleotides conjugated with triple-helix-specific intercalators: benzo[e] and benzo[g] pyridindoles. *Bioconjugate Chem.*, **14**, 120–135.
- Lokhov, S.G., Podyminogin, M.A., Sergeev, D.S., Silnikov, V.N., Kutyavin, I.V., Shishkin, G.V. and Zarytova, V.F. (1992) Synthesis and high stability of complementary complexes of N-(2-hydroxyethyl)phenazinium derivatives of oligonucleotides. *Bioconjugate Chem.*, **3**, 414–419.