Research article

# High-throughput genotyping of single nucleotide polymorphisms with rolling circle amplification

Fawad A Faruqi[1], Seiyu Hosono[1], Mark D Driscoll[1], Frank B Dean[1], Osama Alsmadi[1], Rajanikanta Bandaru[1], Gyanendra Kumar[1], Brian Grimwade[1], Qiuling Zong[1], Zhenyu Sun[1], Yuefen Du[1], Stephen Kingsmore[1], Tim Knott[2] and Roger S Lasken*[1]

Address: [1]Molecular Staging Inc., 300 George St., Suite 701, New Haven, CT 06511, USA and [2]Amersham Pharmacia Biotech, Amersham Laboratories, White Lion Road, Amersham, Buckinghamshire, HP7 9LL, England

E-mail: Fawad A Faruqi - fawadf@molecularstaging.com; Seiyu Hosono - seiyuh@molecularstaging.com; Mark D Driscoll - markd@molecularstaging.com; Frank B Dean - frankd@molecularstaging.com; Osama Alsmadi - osamaa@molecularstaging.com; Rajanikanta Bandaru - rajb@molecularstaging.com; Gyanendra Kumar - gyank@molecularstaging.com; Brian Grimwade - briang@molecularstaging.com; Qiuling Zong - qiulingz@molecularstaging.com; Zhenyu Sun - zhenyus@molecularstaging.com; Yuefen Du - yuefend@molecularstaging.com; Stephen Kingsmore - stephenk@molecularstaging.com; Tim Knott - tim.knott@eu.apbiotech.com; Roger S Lasken* - rogerl@molecularstaging.com

*Corresponding author

Published: 1 August 2001

*BMC Genomics* 2001, **2**:4

Received: 30 May 2001
Accepted: 1 August 2001

This article is available from: http://www.biomedcentral.com/1471-2164/2/4

© 2001 Faruqi et al; licensee BioMed Central Ltd. Verbatim copying and redistribution of this article are permitted in any medium for any non-commercial purpose, provided this notice is preserved along with the article's original URL. For commercial use, contact info@biomedcentral.com

## Abstract

**Background:** Single nucleotide polymorphisms (SNPs) are the foundation of powerful complex trait and pharmacogenomic analyses. The availability of large SNP databases, however, has emphasized a need for inexpensive SNP genotyping methods of commensurate simplicity, robustness, and scalability.

We describe a solution-based, microtiter plate method for SNP genotyping of human genomic DNA. The method is based upon allele discrimination by ligation of open circle probes followed by rolling circle amplification of the signal using fluorescent primers. Only the probe with a 3' base complementary to the SNP is circularized by ligation.

**Results:** SNP scoring by ligation was optimized to a 100,000 fold discrimination against probe mismatched to the SNP. The assay was used to genotype 10 SNPs from a set of 192 genomic DNA samples in a high-throughput format. Assay directly from genomic DNA eliminates the need to preamplify the target as done for many other genotyping methods. The sensitivity of the assay was demonstrated by genotyping from 1 ng of genomic DNA. We demonstrate that the assay can detect a single molecule of the circularized probe.

**Conclusions:** Compatibility with homogeneous formats and the ability to assay small amounts of genomic DNA meets the exacting requirements of automated, high-throughput SNP scoring.

## Background

Sequencing studies of human transcriptomes and genomes have identified hundreds of thousands of single nucleotide polymorphisms (SNPs), the most common

type of human genetic variation [1]. Human SNPs are being assembled into extremely high-density genetic maps that are anticipated to considerably expand the remit of genetic analyses [2]. Broad availability of SNPs for candidate genes will enhance pharmacogenomic and complex trait association studies. Furthermore, extremely dense SNP maps have the potential to make possible genome-wide association studies for complex traits that bypass shortcomings of current genetic linkage analyses [3,4].

The emerging era of multiplexed SNP-based genetic analysis has underscored a need for simple and accurate genotyping methods that can accommodate thousands of loci with economy of cost and consumption of sample DNA. In general, current methods require pre-amplification of genomic DNA, typically by Polymerase Chain Reaction (PCR), followed by SNP genotyping with an allele discrimination method, such as DNA cleavage, ligation, single base extension or hybridization [5]. Current methods are limited either by expense, inaccuracy, consumption of sample DNA, or lack of scalability.

Recently a new method for SNP detection from genomic DNA based on DNA ligase-mediated single nucleotide discrimination and signal amplification by Rolling Circle Amplification (RCA) has been described [6–10]. An oligonucleotide Open Circle Probe (OCP) anneals to the target SNP such that the 5' and 3' ends of the OCP can be ligated together forming a circle topologically linked to the target. A base-pair match between the 3' end of the OCP and the SNP allows DNA ligase to circularize the OCP. A mismatch between the OCP and the SNP prevents ligation and circularization. In this manner, single base selectivity is achieved not only by the specific hybridization of the OCP ends to target sequences adjacent to the SNP, but also by the highly discriminative nick closure activity of the thermostable DNA ligase toward a perfectly matched substrate. Upon OCP circularization, an isothermal exponential RCA (ERCA) reaction involving an exonuclease (-) DNA polymerase with strand-displacement activity and two primers rapidly amplifies the signal by as much as $10^{12}$-fold, allowing for direct SNP genotyping from small quantities of DNA target [6,11,12].

We describe here a simple, scalable assay for SNP genotyping directly from human genomic DNA that uses a 96-well plate format and fluorescent primers called Amplifluors™ [13–15,12,16]. Ten different SNPs have been characterized, each for DNA samples from 192 different individuals, using this reporter assay system. We also report that the ERCA allows genotyping with as few as ~ 300 copies of the target sequence.

## Results

### *Experimental strategy*

Two OCPs were designed for each SNP screened, one complementary to each of the two possible bases of a bi-allelic SNP. If the OCP is complementary to the SNP, it can be circularized by DNA ligase (Fig. 1). The OCPs were ~ 90 nucleotides in length, and are designed such that regions at both the 5' and 3' ends annealed to contiguous segments of the SNP-containing target sequence. The OCP backbone sequence connecting the target complementary 5' and 3' regions is unique for the two allele-specific OCPs. The 5' target-specific region of the OCP is 20–30 bp in length, while the 3' target-specific region is 12–20 bp [12]. The Tm of the 5' target specific region is ~ 5°C above the ligation temperature, while that of the 3' target specific region is ~ 15°C below the ligation temperature. Allele specificity of circularization is dependent upon two factors: 1) favorable hybridization of the matching OCP because of the correct base pair formed at the 3' end

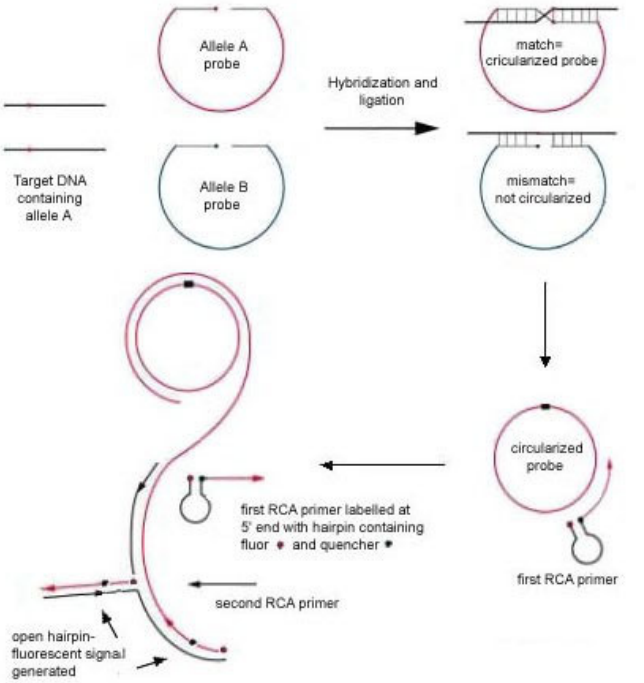


**Figure 1**
Schematic of RCA for SNP identification. An allele specific OCP anneals to the target sequence so that the 5' and 3' arms are adjacent to each other. The 3' terminal nucleotide of the OCP is at the SNP position. If it is complementary to the target, the OCP ends are covalently linked by a DNA ligase to form circles. The circles are then amplified in an ERCA reaction containing two primers, one of which is an Amplifluor containing a fluorophore and a non-fluorescent quencher as indicated. Incorporation of the primer into DNA product opens the primer hairpin giving a fluorescent signal.

and also because of base stacking between the 3' and 5' annealed ends, and 2) specificity of the ligase for correctly base paired ends.

Upon ligation, the DNA circle creates an effective template for an exponential, or hyperbranching, RCA reaction [6,11]. The region of the OCP between the target-specific ends, which we have designated the OCP backbone, provides a binding site for a complementary primer (P1). In addition to P1, a reverse primer (P2) is also present to achieve amplification with exponential kinetics (Fig. 1). P1 is an "Amplifluor™"-primer, with a 5' hairpin and loop structure with a fluorophore and quencher at the base of the hairpin arms [13,12,16]. The two allele-specific OCPs have different backbone sequences complementary either to a FAM- or TET- labeled P1 allowing discrimination of the polymorphic base. Both P1s have an internal DABCYL quencher. Amplifluor primers are designed such that, at the ambient temperature of the assay, the fluorophore and the quencher are in close proximity, preventing fluorescence. Incorporation of Amplifluor into a double-stranded ERCA product opens the hairpin separating the fluorophore and quencher, resulting in increased fluorescence. Two different allele-specific P2s are also used in the assay. They are partly target- and partly backbone-specific, corresponding to the 3' target-specific region of the OCP, and 7–12 nucleotides of the adjacent backbone sequence. The backbone-specific 5' region of the allele-specific P2s not only increases the Tm of the primer to that desired for the assay, but also confers OCP specificity. P2s have two abasic residues at the 5' end to reduce primer-dimer formation (unpublished data).

### *Analysis of allele discrimination during the OCP ligation step of the assay*

The first step in the genotyping assay involves SNP interrogation by ligation of the OCP using the thermostable ligase, Ampligase™, which is reported to have excellent discrimination between a matched and mismatched 3' terminal nucleotide [17]. In order to directly investigate the efficiency and specificity of the ligation step, allele discrimination of OCP ligation was analyzed using 5' $^{32}$P-labeled OCP. The oligonucleotide target contained the sequence for the M1101K locus of the CFTR gene, a disease-causing mutation. An electrophoretic mobility shift for circularized OCP allowed quantification of ligation kinetics. Use of 40 nM OCP and 100 nM oligonucleotide target allowed sufficient signal for detection without the ERCA step that is required for genomic DNA target. 10–15% of the OCP matching the SNP were ligated in 15 sec (Fig. 2a). For mismatched OCP, less than 1% was ligated after 6 hours (Fig. 2b). The ability of Ampligase to discriminate against a mismatch was measured by comparison of initial ligation rates. Allele Discrimination = (percent match ligation/percent mismatch ligation) ($time_{mismatch}/time_{match}$). Ligation of matched OCP occurred at a 100,000-fold greater rate than for mismatched OCP. Greater temperature and shorter 3' OCP arm length both tended to give increased allele discrimination (Fig. 3). The $T_m$ of the 5' OCP arm is greater than the reaction temperature and so the OCP tightly anneals to the target. However, the shorter 3' arm, having a $T_m$ below the reaction temperature, is in equilibrium between the annealed and melted state. A component of the allele discrimination is apparently derived from differential hybridization of the matched and mismatched 3' arms. Presumably, increasing reaction temperature has the same effect as shortening the 3' OCP arm.

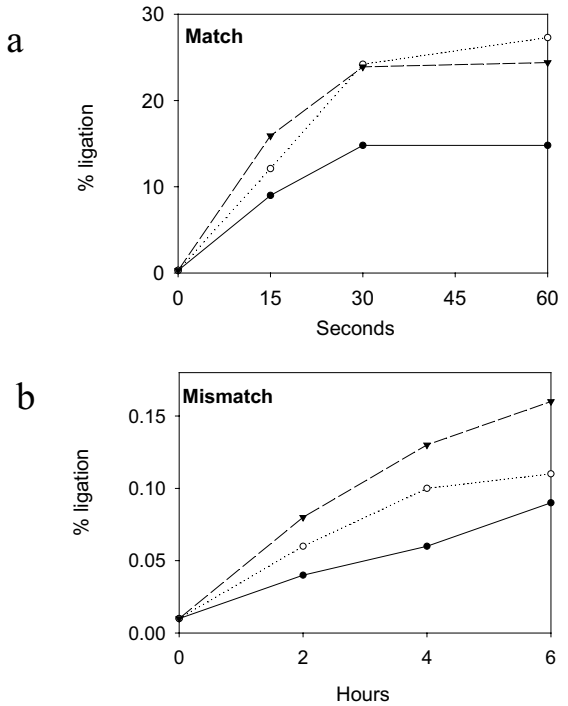


**Figure 2**
Kinetics of the ligation reaction. All reactions contained an oligonucleotide target with the wild type T at the SNP position. Reactions labeled "Match" had the corresponding 3' A on the OCP (a). Reactions labeled "Mismatch" had T on the OCP 3' end (b). Reactions contained 20 unit Ampligase/reaction (closed circles), 40 unit Ampligase/reaction (open circles), or 80 unit Ampligase/reaction (triangles). OCP was labeled on the 5' end with $\gamma$-$^{32}$P-dATP and T4 polynucleotide kinase. Ligation was quantified as percent conversion to circular form on a 6% DNA sequencing gel.

### *Analysis of the ERCA signal amplification step of the assay*

In reactions where the OCP matches the SNP, an estimated 100–30,000 OCPs are circularized by ligation de-

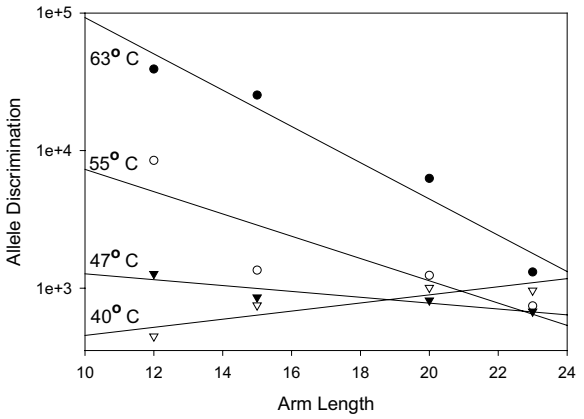


**Figure 3**
Effect of 3' arm length and ligation temperature on allele discrimination. The 3' arm of the matching OCPs formed 12, 15, 20, or 23 basepairs, as indicated, when annealed to the oligonucleotide target. Ligation reaction temperatures were as indicated.

pending on the amount of input DNA target. Detection of this small number of circles depends on ERCA in the second stage of the assay. ERCA is capable of $10^{12}$-fold signal amplification [6]. In order to investigate the efficiency of the ERCA step, OCP was circularized by ligation on a synthetic oligonucleotide target, and gel-purified to remove any linear DNA. A known number of the purified circles were added to reactions allowing direct analysis of ERCA signal amplification. The circles were formed using OCP-1822T for the SNP G1822A using an oligonucleotide target (Table 1I). A 10-fold serial dilution of the circles ranging from $10^6$ to approximately one circle was used in the ERCA reactions. An Amplifluor primer was used to measure the kinetics of the reaction in a real-time instrument. The results confirmed that the assay is capable of detecting < 10 molecules of circular template (Fig. 4a). The time at which fluorescent signal is first detected ($C_T$) is proportional to the number of starting templates over the six orders of magnitude tested (Fig. 4b). This demonstrates the quantitative potential of ERCA and that it is capable of detecting and amplifying the small number of circles formed by ligation on genomic DNA.

### *SNP genotyping on human genomic DNA*

The G1822A SNP on chromosome 13q32 was used for genotyping with OCPs using the SNP assay (B. Grimwade, unpublished). Human genomic DNA was digested with the restriction endonuclease Alu I and assayed in two different tubes, each containing one of the two allele specific OCP. Reactions contained 0.5 pM OCP, and 100 ng of the genomic DNA, corresponding to a gene copy number of approximately 30,000. OCP ligation was performed with Ampligase for 20 min at 60°C, approximately 15°C above the Tm for the OCP 3' arm, in order to maximize specificity of ligation. Since the 5' arm of the OCP has a Tm above that of the ligation temperature, the 5' arm hybridizes to its target in a stable manner, and SNP specificity is achieved via the 3' arm.

The isothermal ERCA reactions were performed at 60°C using Bst DNA polymerase. The reactions contained the appropriate OCP-specific P1 Amplifluor primer and the corresponding allele-specific P2 (1 uM each). When the genomic DNA was homozygous for the G allele (Fig. 5a, top panel) a fluorescent signal was detected only for the OCP with a matching C at the 3' end and the corresponding primer pair P1inFAM and P2inC. Similarly, when the reaction contained genomic DNA homozygous for the A allele (Fig. 5a, center panel) a fluorescence signal was detected only with the OCP containing a matching T at the 3' end and the corresponding primer pair P1ocTET and P2ocT. Only background fluorescence was detected for the mismatched OCPs. As expected, when the target DNA was heterozygous, a fluorescence signal was observed with both OCP/primer combinations (Fig. 5a, lower panel). When the reaction end-point products were analyzed on an agarose gel, an ERCA product ladder was observed only for reactions where the OCP correctly matched the SNP base (Fig 5b). ERCA reactions produce a characteristic ladder of discrete bands because the product consists of double stranded, tandemly linked multiples of the circle sequence[6,11]. As expected, the ladder band spacing was consistent with the length of the OCP.

### *Validation of the assay for ten SNPs for 192 individuals*

The accuracy of the SNP genotyping assay was demonstrated on two sets of human genomic DNA samples, each containing DNA from 96 different individuals. Ten different SNPs representing five of the six possible types of single nucleotide substitutions were assayed (Table 2II). Results were compared to the known genotypes determined by restriction fragment length polymorphism or DNA sequencing of both strands by the single base extension method. The SNP assay gave correct genotypes an average of 93 % of the time (Table 2II). When misscored samples were repeated in triplicate the genotyping accuracy was above 99% (data not shown). A scatter plot of the data for SNP G1822A shows the clear distinction between each of the three possible genotypes (Fig. 5c). The protocol uses a microtiter plate format providing a high throughput method of genotyping.

**Table I: Sequence of oligonucleotides used in the study.**

| Oligonucleotide | Sequence |
|---|---|
| OCP-1822inC | 5'GAAGAACTGGACAGATTTACTACGTGACTCGTCATGTCTCAGCTCTAGTA CGCTGATCTTAGTGTCAGGATACGGAGATGTTCTGCTTTGTC |
| OCP-1822ocT | 5'GAAGAACTGGACAGATTTACTACGTATGTTGACTGGTCACACGTCGTTCT AGTACGCTTCTACTCCCTCTTGAGATGTTCTGCTTTGTT |
| P1-1822inC-FAM | 5'FAM-TCGATGACTGACGGTCATCG(DABCYL-dT)ACTAGAGCTGAGACA TGACGAGTC |
| P1-1822ocT-TET | 5'TET-TCGATGACTGACGGTCATCG(DABCYL-dT)ACGACGTGTGACCAG TCAACAT |
| P2-1822inC | 5'aba-aba-GATACGGAGATGTTCTGCTTTGTC |
| P2-1822ocT | 5'aba-aba-CCTCTTGAGATGTTCTGCTTTGTT |
| OCP-1822C | 5'GAAGAACTGGACAGATTTACTACGTATTCTGACTCGTCATGTCTCAGCTT TAGTTTAATACGACTCACATATCCCAGATGTTCTGCTTTGTC |
| OCP-1822T | 5'GAAGAACTGGACAGATTTACTACGTATTCTGACTCGTCATGTCTCAGCTT TAGTTTAATACGACTCACTATAGGGAGATGTTCTGCTTTGTT |
| P2-1822C-TET | 5'TET-TCGATGACTGACGGTCATCG(DABCYL-dT)ATATCCCAGATGTTC TGCTTTGTC |
| P2-1822T-FAM | 5'FAM-TCGATGACTGACGGTCATCG(DABCYL-dT)TATAGGGAGATGTTC TGCTTTGTT |
| P1 | 5'aba-aba-CTAAAGCTGAGACATGACGAGTC |
| | Oligonucleotides used in the studies of ligation kinetics, affect of OCP 3' arm length, and ligation temperature |
| CFTR M1101K target | 5'CTTGTACCTGTCAACACTGCGCTGGTTCCAAAXGAGAATAGAAATGATTT T TGTCATCTT |
| | x = T in wild type and A in mutant |
| OCP-M1101K23 | 5'TTTGGAACCAGCGCAGTGTTGACACTCTGACTCGTCATGTCTCAGCTTTA GTTTAATACGACTCACTATAGGGGACAAAAATCATTTCTATTCTCY |
| OCP-M1101K20 | 5'TTTGGAACCAGCGCAGTGTTGACACTCTGACTCGTCATGTCTCAGCTTTA GTTTAATACGACTCACTATAGGGAAAAATCATTTCTATTCTCY |
| OCP-M1101K15 | 5'TTTGGAACCAGCGCAGTGTTGACACTCTGACTCGTCATGTCTCAGCTTTA GTTTAATACGACTCACTATAGGGTCATTTCTATTCTCY |
| OCP-M1101K23 | 5'TTTGGAACCAGCGCAGTGTTGACACTCTGACTCGTCATGTCTCAGCTTTAG TTTAATACGACTCACTATAGGGTTTCTATTCTCY |
| | Y = A in wild type and T in mutant. The number 12, 15, 20 or 23 refers to the number of base pairs the 3' end forms with the target. |

### *Single-tube assay for SNP genotyping using a low copy number of targets*

The above experiments involved a two-tube assay with 100 ng of genomic DNA and one OCP/P2 combination per tube. Therefore, a total of 200 ng genomic DNA was used per DNA sample to be investigated. One of the objectives for high-throughput screening of SNPs is the ability to accurately detect a SNP from a small amount of genomic DNA. Previously, it was shown that 20 ng of genomic DNA, equivalent to ~ 6000 copies of the gene, was sufficient to detect a C/T polymorphism using the serial invasive signal amplification reaction (SISAR) [18]. In order to investigate the sensitivity of ERCA for detecting SNPs, 1 ng of genomic DNA, equivalent to ~ 300 copies of the gene, was used as target for circle formation with OCPs for the SNP G1822A.

The method was also simplified so that instead of one tube each for the two allele-specific OCPs, a single-tube for both OCPs was used. Both OCP systems could be combined in the same tube because the seven 5' bases of P2 are complementary to the unique backbone region of the OCP. The P2 for one allele could only effectively prime one OCP and the P2 for the other allele could only prime the other OCP at the ERCA reaction temperature. The P1 primer was the same for the two OCPs and detection and specificity was achieved by two allele-specific Amplifluor P2s carrying either a FAM or a TET fluorophore. The annealing/ ligation reaction therefore includ-

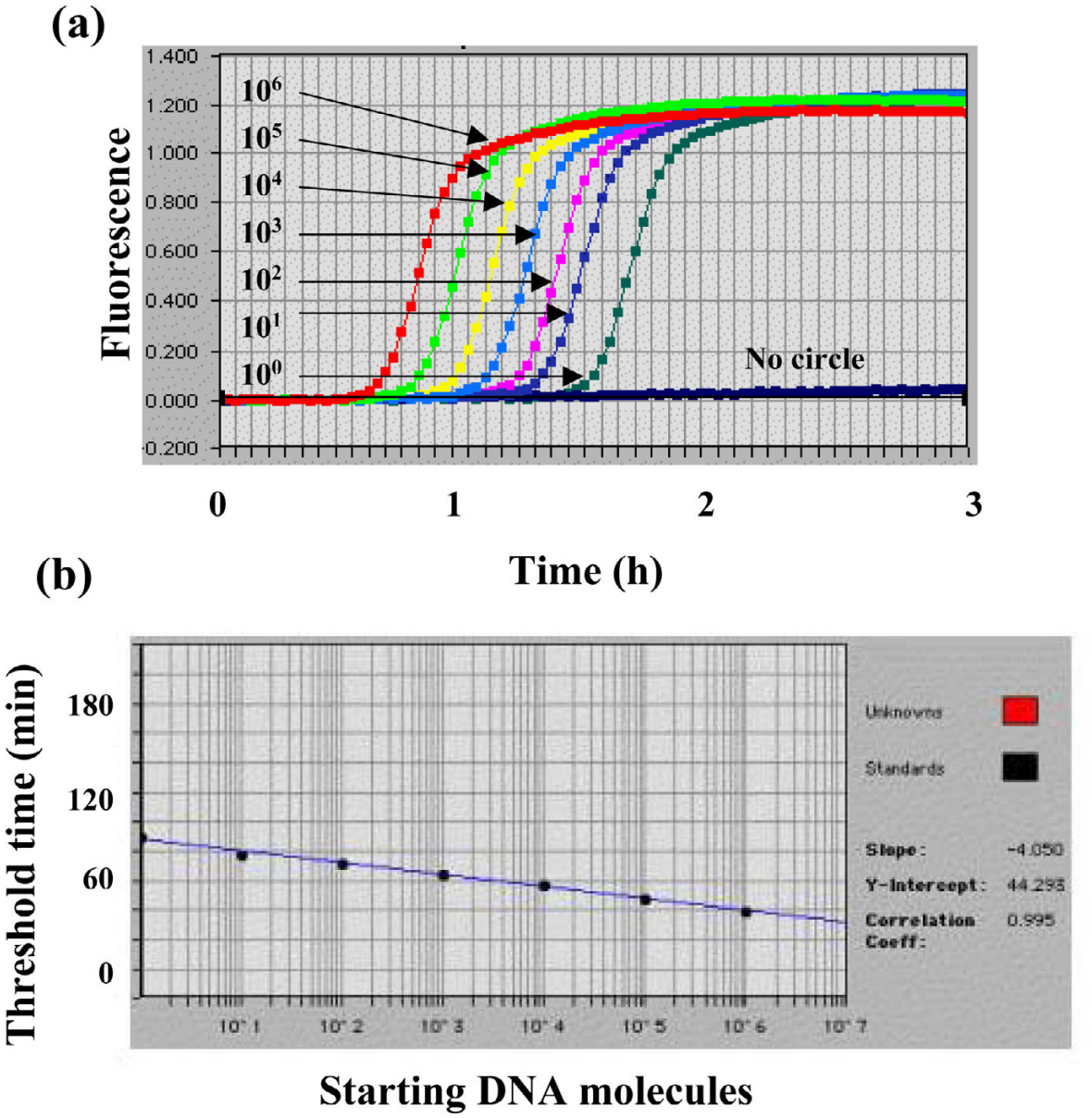


**Figure 4**

Detection of purified circles with ERCA. A 10-fold serial dilution of preformed circles ranging from $10^6$ to about 1 copy was used as substrate in an ERCA reaction. (a) Real-time profile of the amplification signal obtained with each of the circle dilutions. (b) A standard curve of threshold cycle ($C_T$) vs. circle number shows that the reaction is linear over six logs of the target.

ed both of the allele-specific OCPs (OCP-1822C and OCP-1822T, Table 1I), and the ERCA reaction was performed in the presence of both of the allele-specific P2 Amplifluor primers (P2-1822C-TET and P2-1822T-FAM). This assay detected and differentiated all three possible genotypes for the G1822A SNP (Fig. 6). The background signal observed was slightly higher than that obtained with 100 ng genomic DNA in a two-tube system, with a signal/noise ratio of 5:1. However, correct SNP genotyping was observed with 48 different human genomic DNA samples (data not shown).

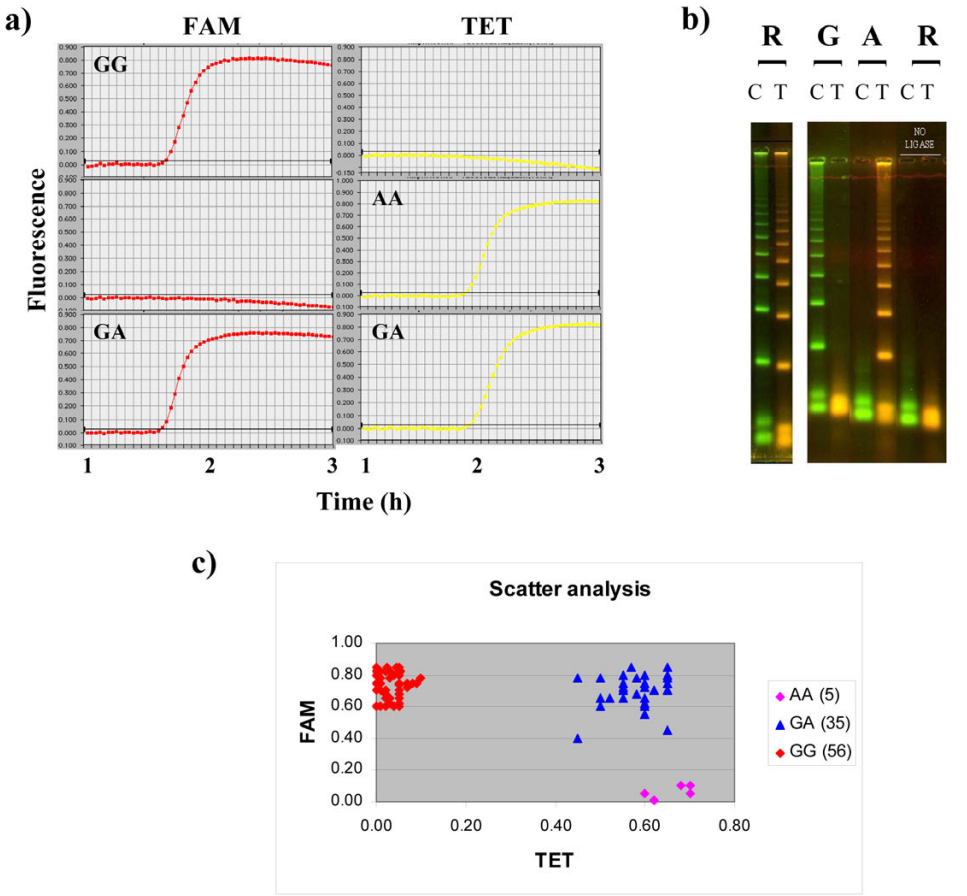


**Figure 5**
SNP identification on human genomic DNA. Alu I digested human genomic DNA (100 ng) was used as template for annealing and ligation of the OCP specific for the SNP G1822A. The ligation products were then amplified by ERCA. Three DNA samples representing the three genotypes were each analyzed in two reactions, each reaction containing an OCP/ P1 Amplifluor/ P2 combination corresponding to one of the two possible alleles. (a) Real-time profiles of the ERCA reactions indicate that an amplification signal is obtained only from reactions where the OCP/ P1 Amplifluor/ P2 combination matched the SNP nucleotide on the genomic DNA samples. When the combination OCPinC/ P1inFAM/ P2inC was used with DNA sample homozygous for the G allele, a fluorescent FAM signal was observed (top left panel). No signal was observed with the mismatched OCPocT and the corresponding primer pair P1ocTET/ P2ocT (top right panel). Similarly, when the combination OCPocT/ P1ocTET/ P2ocT was used with DNA sample homozygous for the A allele, only a fluorescent TET signal was observed (middle right panel). No signal was observed with the mismatched OCPinC and the corresponding primer pair P1inFAM/ P2inC (middle left panel). In the case of heterozygous DNA both the OCP/ primer combinations gave a fluorescent signal (bottom panels) (b) The endpoint products of these reactions were resolved on a 2% agarose gel. The characteristic double-stranded ERCA ladder was observed only from those reactions that gave a fluorescent signal in the real-time assay. (c) Scatter plot of the ERCA reaction endpoint fluorescence values obtained by screening a set of 96 different genomic DNA samples using the above assay.

## Discussion

We have demonstrated a solution-based, efficient, homogeneous and robust assay for genotyping SNPs directly from human genomic DNA utilizing ligation of open circle probes and rolling circle amplification. The design of the assay is fairly straightforward and the assay can be carried out in three hours (30 min for the denaturation / annealing / ligation reaction and 2 1/2 h for the amplification reaction) in a 96-well format. Specificity of the OCP to its target sequence is achieved by the complementarity of the two ends of the OCP to the target and the requirement of these ends to be adjacent for ligation. Single nucleotide discrimination is achieved at the ligation step by the use of the thermostable DNA ligase,

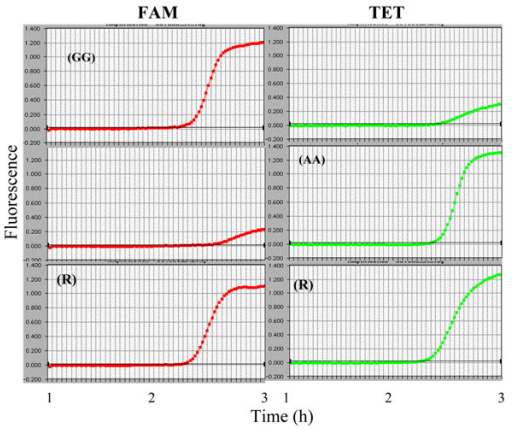


**Figure 6**
Genotyping of SNP G1822A on 1 ng of genomic DNA in a single-tube reaction. Ligation reactions were performed with 1 ng of Alu I digested genomic DNA in single tubes containing both of the allele-specific OCPs. ERCA reactions included both of the allele-specific P2 Amplifluor primers. Real-time amplification signal was observed only when the OCP/ P2 combination matched the SNP nucleotide on the genomic DNA. Therefore, FAM and TET P2s gave a signal with DNA homozygous for the A and G alleles respectively. DNA heterozygous for the SNP gave an amplification signal with both P2 Amplifluor primers.

Ampligase, which has a high affinity for a perfectly matched substrate at the 3' end of a DNA molecule [17]. We were able to enhance allele discrimination at the ligation step by designing the OCP such that the 5' complementary region is firmly hybridized to the target sequence whereas the 3' region is in equilibrium with its target at the ligation temperature [12]. This would result in increased specificity since the correctly matched OCP will have a greater chance to act as substrate for the ligase. It has been previously reported that 3'-T/G and 3' G/T mismatches are not good substrates for single nucleotide discrimination [17]. However, we found that the G1822A SNP, which would result in a 3'-T/G mismatch, was efficient for allele discrimination. The ability of this assay to genotype any SNP regardless of the base pair involved is an important advantage over assays based on primer extension such as PCR.

Allele discrimination achieved at the ligation step results in small circular DNA molecules topologically linked to the target DNA strand. These DNA circles are amplified in the powerful homogeneous ERCA reaction capable of $10^{12}$-fold signal amplification [6], similar to that achieved with PCR technology, the current gold standard in genetic analysis and quantitation. However, PCR involves exponential target amplification, thereby increasing the risk of amplicon cross-contamination. Even though this shortcoming can be overcome, it increases the cost and complexity of the assay, making it less attractive for high-throughput analysis. Since ERCA is a signal (circle) amplification method, it does not have the problems associated with PCR. In addition, ERCA is an isothermal reaction and the reaction endpoint can be used as the assay readout. Even though the present study was conducted on a real-time ABI 7700 Sequence Detector instrument the strategy can be easily adapted for a simple fluorescence plate reader coupled to an adjustable heating block. These properties make it an ideal system for high-throughput analysis.

The assay was tested for 10 SNPS on two sets of 96 different DNA samples (Table 2II). Results were compared to the known genotypes that we determined by RFLP or single nucleotide sequencing reactions. The assay had an average genotyping accuracy of 93% when samples were screened initially. When the mis-scored samples were repeated in triplicate, the genotyping accuracy jumped to over 99%. The majority of the mis-scoring involved a homozygous sample being called heterozygous, i.e., an amplification signal was observed with both sets of OCP/ primer combinations. A DNA sample homozygous for one allele was never genotyped as homozygous for the other allele. This implied that there is a low frequency of DNA synthesis artifacts resulting in a fluorescence signal.

Indeed when these reactions were analyzed on an agarose gel, the size of the characteristic ERCA DNA ladder was different from that obtained with amplification of the OCP (data not shown). We have used abasic residues at the 5' ends of P2 primers in order to reduce these artifacts. Other nucleotide modifications that have been reported to reduce primer-dimer formation will be tested to improve the accuracy of SNP genotyping. In addition, reagents that have been shown to reduce primer-dimer formation in PCR will be tested in the ERCA reaction.

Background signal is sometimes related to the amount of OCP used in the assay. For each SNP, the optimal OCP concentration needs to be determined before screening. Excess OCP concentrations result in an increase in non-specific fluorescence signal, therefore lowering the accuracy rate of genotyping. Potentially, un-ligated OCP can act as template or primer giving rise to a low level of non-specific DNA synthesis and subsequent fluorescent signal. We are currently developing a modified OCP design to overcome the need for concentration optimization. In this design, the 3' region of the OCP forms a stable hairpin-loop structure and opens up only to anneal to it target sequence (O. Alsmadi, unpublished). As with molecular beacons, this may also improve target specificity [19,20]. Any unused OCP is self primed and extend-

**Table II: Summary of 10 SNPs genotyped on genomic DNA.**

| | | 96 NIMH DNA samples | | 96 Coriell DNA samples (HD100CAU) | |
|---|---|---|---|---|---|
| SNP ID | SNP | # correct/ # screened | % accuracy | # correct/ # screened | % accuracy |
| MSI 1822 | G/A | 96/ 96 | 100 | 92/96 | 96 |
| MSI 251 | A/T | 90/ 96 | 94 | 90/96 | 94 |
| MSI 221 | T/C | 87/ 96 | 91 | 88/96 | 92 |
| GW 465 | G/A | 90/ 96 | 94 | 87/96 | 90 |
| GW 458 | G/A | 93/ 96 | 97 | 81/96 | 84 |
| AZ PDK2-2 | G/C | 89/ 96 | 93 | 82/96 | 85 |
| JT SNP21 | C/A | 90/ 96 | 94 | 90/96 | 94 |
| AZ VCAM | G/A | 93/96 | 97 | 93/96 | 97 |
| AZ IL8 | T/O | 90/96 | 94 | 86/96 | 90 |
| GW474 | G/C | 91/96 | 95 | 86/96 | 90 |

ed to form an inert double stranded molecule, thus eliminating OCP as a source for non-specific amplification. Initial experiments with this design have been encouraging.

## Conclusions

We have described a solution-based SNP genotyping assay that is simple, sensitive and robust and can be easily formatted to high-throughput analysis of single nucleotide polymorphisms and mutation detection directly from human genomic DNA. The SNP genotyping assay uses a simple, homogeneous, fluorescence readout and can be carried out on inexpensive instruments already available in many academic and industrial laboratories.

Amplifluor is a trademark of Intergen Company.

Ampligase is a trademark of Epicentre Technologies Corp.

## Materials and methods

### *Oligonucleotides*

Table 1I shows the sequences of the oligonucleotides used in the study. Gel-purified and phosphorylated OCP oligonucleotides were obtained from Integrated DNA Technologies, Inc. (Coralville, IA). Abasic and Amplifluor primers were synthesized in-house and purified by HPLC. The two Amplifluors are labeled with a differently colored fluorophore: fluorescein (FAM) and tetrachloro-6-carboxyfluorescein (TET) and contain an internal non-fluorescent quencher DABCYL.

### *Instrumentation*

DNA denaturation, annealing and ligation reactions were carried out in an Eppendorf Master Cycler (Eppendorf Scientific, Germany). ERCA reactions were performed in the Real-Time ABI 7700 Sequence Detector (Perkin Elmer).

### *Genomic DNA*

Genomic DNA samples were obtained from (a) National Institute of Mental Health (NIMH) Center for Genetic Studies, Rutgers University Cell and DNA Repository, Piscataway, NJ and (b) Coriell Cell Repository (HD 100 CAU). The DNA samples were digested with the restriction endonuclease Alu I before being used as template in the ligation reaction.

### *DNA annealing and ligation*

The reactions were set up in 96-well MicroAmp Optical plates (Perkin Elmer) in a 10 µl reaction volume containing 1 U Ampligase (Epicentre Technologies), 20 mM Tris-HCl (pH 8.3), 25 mM KCl, 10 mM $MgCl_2$, 0.5 mM NAD, and 0.01% Triton® X-100. Standard reactions contained 0.5 pM OCP and 100 ng of Alu I digested genomic DNA. DNA was denatured by heating the reactions at 95°C for 3 min followed by annealing and ligation at 60°C for 20 min.

### *ERCA reaction*

20 µl of ERCA mix was added to the 10 µl ligation reaction. The ERCA mix contained 5% tetramethyl ammonium oxalate [21], 400 µM dNTP mix, 1 uM each of the two primers, 8 u of Bst polymerase (New England Biolabs, MA), and 1X modified ThermoPol reaction buffer con-

taining 20 mM Tris-HCl (pH 8.8), 10 mM KCl, 10 mM $(NH_4)_2SO_4$ and 0.1% Triton X-100.

### *Kinetics of the ligation reaction*

All ligation reactions were in 1x Ampligase Reaction Buffer supplied by Epicentre Technologies: 20 mM Tris-HCl (pH8.3 at 25°C), 25 mM KCl, 10 mM $MgCl_2$, 0.5 mM $NAD^+$, and 0.01% Triton X-100. The OCP were synthesized with a 5' phosphate and labeled with $^{32}P$ by the T4 kinase exchange reaction (following Life Technologies suggested protocol for T4 PNK). 40 nM CFTR M1101K12 OCP (Table 1I) was annealed to 100 nM oligo target, at 95°C, 2 min, and slow cooled to room temperature. Reactions were started by adding Ampligase at 0.2 units/µl final concentration in a reaction volume of 0.1 ml. Reactions time points were taken at 15, 30, and 60 sec for matched OCP and 1, 2, 4, and 6 hours for mismatched OCP. 3.5 µl was resolved on an 8% PAGE sequencing gel.

### *OCP 3' arm length and ligation reaction temperature*

OCP were designed with 12, 15, 20, or 23 nt 3' arms (Table 1I). All components except Ampligase were prewarmed to the reaction temperature (40, 47, 55, or 63°C) for 20 min to allow OCP annealing to come to equilibrium. Reactions were started by adding Ampligase at 0.2 units/µl final concentration and a reaction volume of 0.1 ml. Reactions conditions were the same as for kinetics of the ligation reaction above except for the reaction temperatures indicated. Reactions were stopped at 1 min for matched OCP and 4 hours for mismatched OCP by digestion with 0.75 units/µl Exonuclease I and 1.25 units/µl T7 gene 6 exonuclease (Amersham Pharmacia Biotech, Inc, NJ) at 37°C for 1 h, except the zero time points. The digestion removes all but circular DNA. 3.5 µl was resolved on an 8% PAGE sequencing gel.

## References


1. Collins FS, Guyer MS, Charkravarti A: **Variations on a theme: cataloging human DNA sequence variation.** *Science* 1997, **278**:1580-1581
2. McCarthy JJ, Hilfiker R: **The use of single-nucleotide polymorphism maps in pharmacogenomics.** *Nat Biotechnol* 2000, **18**:505-508
3. Halushka MK, Fan JB, Bentley K, Hsie L, Shen N, Weder A, Cooper R, Lipshutz R, Chakravarti A: **Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis.** *Nat Genet* 1999, **22**:239-247
4. Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl MC, Nemesh J, Lane CR, Schaffner SF, Bolk S, Brewer C, Tuomi T, Gaudet D, Hudson TJ, Daly M, Groop L, Lander ES: **The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes.** *Nat Genet* 2000, **26**:76-80
5. Landegren U, Nilsson M, Kwok PY: **Reading bits of genetic information: methods for single-nucleotide polymorphism analysis.** *Genome Res* 1998, **8**:769-776
6. Lizardi PM, Huang X, Zhu Z, Bray-Ward P, Thomas DC, Ward DC: **Mutation detection and single-molecule counting using isothermal rolling-circle amplification.** *Nat Genet* 1998, **19**:225-232
7. Nilsson M, Krejci K, Koch J, Kwiatkowski M, Gustavsson P, Landegren U: **Padlock probes reveal single-nucleotide differences, parent of origin and in situ distribution of centromeric sequences in human chromosomes 13 and 21.** *Nat Genet* 1997, **16**:252-255
8. Nilsson M, Malmgren H, Samiotaki M, Kwiatkowski M, Chowdhary BP, Landegren U: **Padlock probes: circularizing oligonucleotides for localized DNA detection.** *Science* 1994, **265**:2085-2088
9. Baner J, Nilsson M, Mendel-Hartvig M, Landegren U: **Signal amplification of padlock probes by rolling circle replication.** *Nucleic Acids Res* 1998, **26**:5073-5078
10. Landegren U, Kaiser R, Sanders J, Hood L: **A ligase-mediated gene detection technique.** *Science* 1988, **241**:1077-1080
11. Zhang DY, Brandwein M, Hsuih TC, Li H: **Amplification of target-specific, ligation-dependent circular probe.** *Gene* 1998, **211**:277-285
12. Thomas DC, Nardone GA, Randall SK: **Amplification of padlock probes for DNA diagnostics by cascade rolling circle amplification or the polymerase chain reaction.** *Arch Pathol Lab Med* 1999, **123**:1170-1176
13. Nazarenko IA, Bhatnagar SK, Hohman RJ: **A closed tube format for amplification and detection of DNA based on energy transfer.** *Nucleic Acids Res* 1997, **25**:2516-2521
14. Nuovo GJ, Hohman RJ, Nardone GA, Nazarenko IA: **In situ amplification using universal energy transfer-labeled primers.** *J Histochem Cytochem* 1999, **47**:273-280
15. Uehara H, Nardone G, Nazarenko I, Hohman RJ: **Detection of telomerase activity utilizing energy transfer primers: comparison with gel- and ELISA-based detection.** *Biotechniques* 1999, **26**:552-558
16. Myakishev MV, Khripin Y, Hu S, Hamer DH: **High-throughput SNP genotyping by allele-specific PCR with universal energy-transfer-labeled primers.** *Genome Res* 2001, **11**:163-169
17. Luo J, Bergstrom DE, Barany F: **Improving the fidelity of Thermus thermophilus DNA ligase.** *Nucleic Acids Res* 1996, **24**:3071-3078
18. Hall JG, Eis PS, Law SM, Reynaldo LP, Prudent JR, Marshall DJ, Allawi HT, Mast AL, Dahlberg JE, Kwiatkowski RW, de Arruda M, Neri BP, Lyamichev VI: **From the cover: sensitive detection of DNA polymorphisms by the serial invasive signal amplification reaction.** *Proc Natl Acad Sci U S A* 2000, **97**:8272-8277
19. Tyagi S, Kramer FR: **Molecular beacons: probes that fluoresce upon hybridization.** *Nat Biotechnol* 1996, **14**:303-308
20. Bonnet G, Tyagi S, Libchaber A, Kramer FR: **Thermodynamic basis of the enhanced specificity of structured DNA probes.** *Proc Natl Acad Sci U S A* 1999, **96**:6171-6176
21. Kovarova M, Draber P: **New specificity and yield enhancer of polymerase chain reactions.** *Nucleic Acids Res* 2000, **28**:E70



Publish with **BioMed** Central and every scientist can read your work free of charge

*"BioMedcentral will be the most significant development for disseminating the results of biomedical research in our lifetime."*

Paul Nurse, Director-General, Imperial Cancer Research Fund

Publish with **BMC** and your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours - you keep the copyright

**BioMed**central.com

Submit your manuscript here:
http://www.biomedcentral.com/manuscript/

editorial@biomedcentral.com