



Exo-Proofreading, A Versatile SNP Scoring Technology

Patrick Cahill, Michele Bakis, James Hurley, et al.

Genome Res. 2003 13: 925-931

Access the most recent version at doi:[10.1101/gr.939903](https://doi.org/10.1101/gr.939903)

References This article cites 22 articles, 7 of which can be accessed free at:
<http://genome.cshlp.org/content/13/5/925.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

A horizontal banner advertisement with a teal background. On the left, the text reads "CRISPR and RNAi Genetic Screening. Your new superpower." in white. In the center, there is a white-bordered box containing the words "LEARN MORE" in black. On the right, there is a photograph of a woman wearing a red mask and a red cape, and the Cellecta logo, which consists of a green molecular structure and the word "CELLECTA" in white.

To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Cold Spring Harbor Laboratory Press

Methods

Exo-Proofreading, A Versatile SNP Scoring Technology

Patrick Cahill,¹ Michele Bakis, James Hurley, Veena Kamath, William Nielsen, Dina Weymouth, Josée Dupuis, Lynn Doucette-Stamm, and Douglas R. Smith

Genome Therapeutics Corporation, Waltham, Massachusetts 02453, USA

We report the validation of a new assay for typing single nucleotide polymorphisms (SNPs) that takes advantage of the 3'-to-5' exonuclease proofreading activity of many DNA polymerases. The assay uses one or more primers labeled on the 3' nucleotide base, and can be implemented in a variety of formats including a one-step PCR reaction that allows SNP typing directly from genomic DNA samples. The detection of genotypes can be accomplished by means of fluorescence detection on assays that have been purified to remove excess primer, or by means of fluorescence polarization without any additional cleanup. We also demonstrate that the Exo-Proofreading SNP assay can be used on pooled samples to obtain allele frequency data.

With the successes of the Human Genome Project came an increasing need for low-cost, high-throughput technologies for the accurate typing of single nucleotide polymorphisms (SNPs) on human DNA for use in genetic inheritance or disease association studies. Many such methods have been developed over the past decade, but there is no clear choice in terms of cost, efficiency, speed, accuracy, and versatility. The most widely used methods for high-throughput SNP typing rely on allele-specific hybridization (Saiki et al. 1986; Pease et al. 1994) or primer extension (Syvanen et al. 1990), but other popular methods include allele-specific amplification (Newton et al. 1989), allele-specific ligation (Nickerson et al. 1990), flap endonuclease cleavage (Marshall et al. 1997), restriction endonuclease cleavage, pyrosequencing (Nyren et al. 1997), and conventional sequencing. These core technologies can be applied in a variety of formats, such as single or multiplexed reactions in solution (typically carried out in multiwell plates), on the surfaces of microbeads (Uhlen 1989), or on the surfaces of microarrays (Chetverin and Kramer 1993; Lipshutz et al. 1995). Several signal generation and detection methods have been applied, with fluorescence readout being the most prevalent. Such methods typically use labeled primers or nucleotides in conjunction with specialized instrumentation using sensitive optical detection devices—either as stand-alone plate or chip readers or as integrated electrophoretic or flow separation devices. Recently, mass spectrometry (MS) has been gaining popularity thanks to improvements in the instrumentation and software, and to the development of new chemistries for sensitive detection of nucleic acids using matrix-assisted laser desorption ionization (Braun et al. 1997). The development of efficient automated sample preparation techniques has also influenced the practice of these technologies.

In this report, we describe a fundamentally different approach for SNP scoring that takes advantage of the 3'-5' exonuclease (proofreading) activity native to many DNA polymerases to discriminate whether the 3' nucleoside of a primer is hybridized correctly to a template, and to extend or remove

that nucleoside in the case of a match or mismatch, respectively. The assay described here is simple, inexpensive, and sufficiently versatile to be used in conjunction with any of the above-mentioned signal generation and detection methods.

RESULTS

The Exo-Proofreading assay is designed to detect single nucleotide mismatches by selective incorporation or removal of a labeled test base by 3'-5' exo-nuclease activity of a proofreading DNA polymerase as illustrated in Figure 1. The two essential components of the assay are a 3'-end-labeled primer and a polymerase with proofreading activity. In general, however, the assay is practiced using two labeled primers (corresponding to both allele variants of the polymorphic base) and a reverse primer to enable PCR amplification of the extension products. Each forward primer has its 3' end aligned with the polymorphic base (T and C in the example shown in Fig. 1) and is labeled on the nucleotide base with a detectable tag such as a fluorescent moiety. If the primer has a 3' mismatch with the template, the labeled nucleotide is removed by the proofreading polymerase and no tag is incorporated into the PCR product. On the other hand, when the 3' end of the primer is matched with the target, the labeled nucleotide is extended and incorporated into the PCR product. Fluorescent imaging or other suitable detection technology then visualizes the products. The assay can also be practiced using a four-primer configuration, in which each labeled primer is hybridized to a different strand and has its own reverse primer (Fig. 1B). Finally, the assay can be performed using linear amplification by removal of the reverse primer or primers.

In one of the standard assay formats, a primary PCR product is generated from genomic DNA surrounding the SNP of interest, typically between 400 and 600 bp in length, although larger fragments can be used. The PCR product is diluted 1000-fold and subjected to the Exo-Proofreading reaction after addition of the appropriate primers. After thermocycling, the excess primers are removed, and the fluorescence of the reaction products is measured using a standard fluorescence plate reader.

To test the utility of our SNP assay, we compared the results generated for six SNPs typed using the Exo-

¹Corresponding author.

E-MAIL pcahill@genomecorp.com; **FAX** (781) 398-2472.

Article and publication are at <http://www.genome.org/cgi/doi/10.1101/gr.939903>. Article published online before print in April 2003.

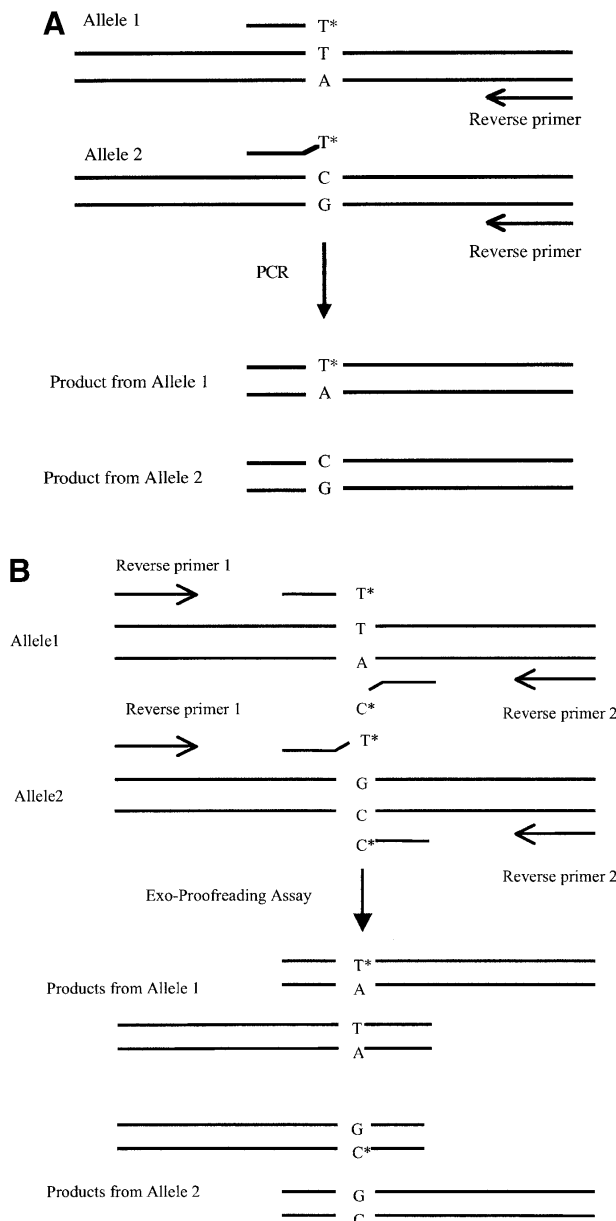


Figure 1 (A) Diagram of three-primer Exo-Proofreading assay. A primer is designed with its 3' terminus aligned at the polymorphic base of a DNA target (T to C in this example) and labeled with a detectable tag on the 3' nucleotide base. If the 3' end of the primer is matched with the target (allele 1), the labeled nucleotide is retained and incorporated into a PCR product generated with a reverse primer. If the primer has a mismatch with the template (allele 2), the labeled nucleotide is removed by the proofreading activity of the polymerase and no tag is incorporated into the PCR product. (B) Diagram of four-primer Exo-Proofreading assay. The four-primer assay is essentially the same as the three-primer assay except that the forward primers are interrogating the SNP from the plus and minus strands of DNA. An extra reverse primer is added, resulting in two PCR products.

Proofreading Assay with previous results obtained using allele-specific oligonucleotide (ASO) hybridization. This set of SNPs covers the full set of possible base variations, when both DNA strands are considered, using only the available 3'-amino-modified cytosine and thymidine primers. The

samples were typed for SNPs 216 S2, 216 S + 1, 216 V - 2, 216 M + 1, 216 ST + 4, 216 ST + 7 by the Exo-Proofreading assay and then compared with the same samples previously typed by ASO (Van Eerdewegh et al. 2002). The primers used in these assays are shown in Table 1. Note that SNPs 216 S2, 216 ST + 4, and 216 S + 1 used the four-primer assay; SNPs 216 V - 2 and 216 ST + 7 used the three-primer assay; and the linear amplification or two-primer assay was used to type SNP 216 M + 1.

Exo-Proofreading genotypes for each set of samples (192–432 individuals) were scored using an automated algorithm based on the raw fluorescence values measured at each of the two maximal emission wavelengths associated with the tags used in the assay. In addition to the genotype, the algorithm generated a confidence value for each call derived from the posterior probability of a correct genotype assignment. A typical scatter plot of the data is shown in Figure 2, and a summary of the data is presented in Table 2. The average call rate (where a genotype could be assigned) for these samples was 98.9%. All of the samples where genotypes were unable to be assigned were the result of failed primary PCR reactions as determined by agarose gel analysis of the PCR products. The percentage of calls with a confidence value of 90% or greater was 99%. The correlation between the Exo-Proofreading reaction and the ASO assays was >99%. However, for all but one of the SNPs, there were examples where the genotype obtained from the Exo-Proofreading assay did not match the results of the ASO assay. These samples were sequenced to attempt to resolve the discrepancies. In all of the cases, the genotype obtained from the sequencing assay corresponded to the Exo-Proofreading assay.

Because of the exponential nature of the PCR reaction, it was assumed that very little target DNA would be required to obtain a positive result. To test this, we performed the assay directly on genomic DNA using 23 ng of total genomic DNA. The results, shown in Figure 3, yielded a similar range of signal intensities (7.1–8.0 log[Tamra RFU]; 6.8–7.8 log[Fam RFU]; RFU = relative fluorescence units) as the two-stage assay. The grouping of samples into genotype clusters appeared to be slightly less distinct than was the case with the two-stage assay, but the genotypes were still clearly distinguishable and callable by the automated genotyping software. Similar results were obtained using 12 ng of genomic DNA (the smallest amount tested). The less distinct genotype clustering appeared to be more experiment-dependent than assay-dependent based on a small number of experiments (data not shown), but this was not studied extensively.

All of the results described so far were obtained using a purification step after the Exo-Proofreading reaction to remove any remaining label associated with the primers and excision products. However, because of the significant expected molecular size difference between the starting primers and the resulting extension and excision products, we were interested to see whether it would be possible to eliminate the need for the clean-up step by using fluorescence polarization detection (Chen et al. 1999). If successful, this would enable a single-step homogenous assay in which one could go from sample to result in one reaction vessel. Fluorescence polarization values, commonly expressed as millipolarization units (mP), are related to the rotational freedom and hence the molecular size of the fluorescently tagged molecules being measured. As the population of tagged molecules within the reaction shifts from oligonucleotide primers to extended PCR products, a shift in average millipolarization units of the com-

Table 1. Primer Design

SNP name	Type	Allele-specific primers	Label ^a	Reverse primers	Melting temperature	dbSNP ID (ss#)
216 S2	G/C	ACCAGGCCAGGCCGGCC TGCTGCCCTCTGCTCCCAGGC	Tam Fam	TGGCCCTGTGCAGGCTGAAA GGGAACCGCAGGAGTAGGCTCA	68	SS5112738
216 M+1	G/T	AAGGTCAGGGTGAGGGTTTCGT AAGGTCAGGGTGAGGGTTTCGC	Tam Fam	NA (Linear Assay)		SS5112737
216 ST+4	T/G	CAGCATTTGGGAACCTCAAGT GTCCACTCTCCTGCTTCTCC	Fam Rox	CAAATCACCTCTGTACACCC CCACCCAGGTACAGAGAAC	59	SS5112741
216 ST+7	T/C	TTTGACTGTCCCATCCCATCT TGACTGTCCCATCCCATCC	Fam Rox	GTCTGTGGGCAGAGGAGCAA	64	SS5112744
216 V-2	A/G	ACTGAGTGAGGATGGGCTCT CTGAGTGAGGATGGGCTCC	Tam Fam	TGCAGTTCAAGTTCTGGAG	59	SS5112757
216 S+1	T/A	GTGGCCTCCAGTCAAGCGT TGGGCAGGGATCCACCCCT	Tam Fam	GTCAAGGGTGAGGCAGCCAG CGAGCGCTGCTTCTGAGCC	60	SS5112740

^a(Tam) TAMRA; (Fam) 6FAM.

ponents can be detected. Figure 4 shows the results obtained using fluorescence polarization. In this experiment, samples were transferred from the thermocycling plate to the fluorescence read plate and fluorescence measurements were taken without any other manipulations. Although the range of millipolarization values in the resulting clusters appear larger than the corresponding range of fluorescence values obtained with the standard assay, the elongated shapes of the clusters are clearly defined and the separation is sufficient to allow an unambiguous genotype to be assigned in all cases. Other experiments (data not shown) gave similar results with fluorescent polarization (FP) values tending to have larger spreads than the corresponding fluorescent intensity (FI) values. In addition, fluorescent polarization detection tends to be less reproducible than Exo-Proofreading cleanup followed by straight fluorescent intensity measurements, and some assay-specific variation was observed.

Another potential use for the Exo-Proofreading assay is in allele frequency determination. To evaluate the utility for this application, we created an artificial set of samples with different proportions of wild-type and mutant alleles by mixing two homozygous samples. Figure 5, A and B, presents the data from those experiments. Figure 5A presents the raw fluorescent values for samples containing 100% C to 100% T and ratios of each of the alleles in between in 5%–10% increments. If the data are plotted as the ratio of the two fluorescence values versus the percent of one of the alleles, a linear relationship is observed (Fig. 5B). Comparison of the ratio of fluorescence values of a pooled set of samples for this SNP to the standard curve would, in principle, allow an estimate of the allele frequency to be determined. However, further work will be required to demonstrate the applicability of the method.

DISCUSSION

Most polymerases will not efficiently catalyze extension of a primer from a mismatched terminal base pair (although

some error-prone polymerases such as reverse transcriptase will do so; Preston et al. 1988). If the enzyme contains an integral 3'-to-5' exonuclease activity, the polymerase will instead rapidly excise the mismatched base pair and then produce a faithful extension product from a penultimate matched pair. This process is commonly referred to as proofreading, and is highly accurate. The error rate of proofreading polymerases (as measured by inappropriate extension of a mismatched nucleotide) typically falls in the range of 10^{-5} to 10^{-6} (Kunkel et al. 1984; Mattila et al. 1991; Cline et al. 1996). Mutant proofreading polymerases with lower error rates (such as the "antimutator" form of T4 polymerase), have been described (Drake et al. 1969), but have not yet been used in this assay. The Exo-Proofreading assay described here takes advantage of the inherent accuracy of proofreading polymerases to produce PCR products using 3'-end-labeled primers that accurately reflect the genotype of the starting sample at the position of the variable base. Several thermostable polymerases were tested in the assay, including Vent, Deep Vent, and Tgo, but the Pfu and Pwo polymerases, which are identical in

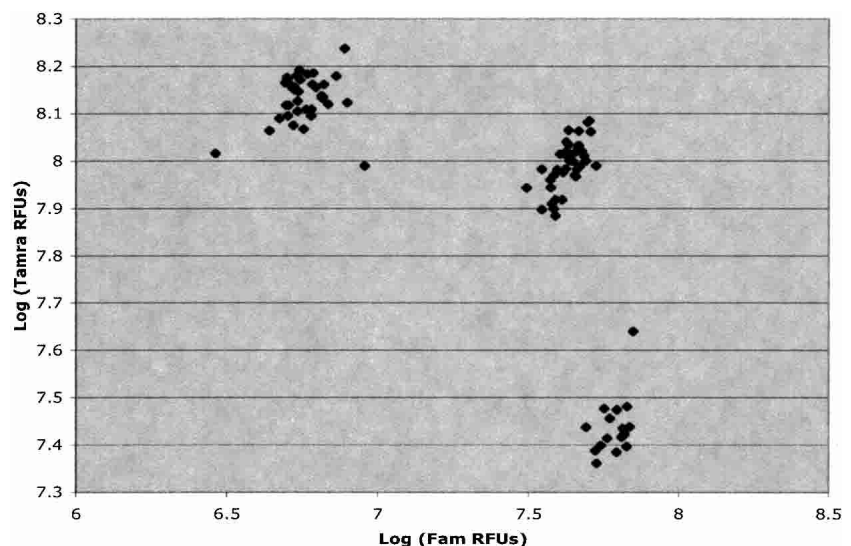


Figure 2 Scatter plot of Exo-Proofreading assay results 216 V – 2. Data from 96 samples assayed using the three-primer Exo-Proofreading assay. The data are plotted as the log of the fluorescence readings from the LJI instrument, in arbitrary relative fluorescence units (RFUs).

Table 2. Results of Exo-Proofreading ASO Comparison

SNP name	Number of samples	% called EXO	% called ASO	No. of discrepant samples	% correlation	% Exo accuracy
216 S2	432	100	99.77	1	99.8	100
216 S+1	432	100	92.82	5	98.8	100
216 V-2	384	99.74	99.22	1	99.7	100
216 M+1	432	100.00	100.00	1	99.8	100
216 ST+4	194	94.33	98.45	2	99.0	100
216 ST+7	194	99.48	98.97	0	100	100

sequence (Dabrowski and Kur 1998), appeared to work best in terms of the success rate for assay development and signal strength. Interestingly, we also found that consistent and accurate results can be produced using a mixture of Taq (lacking proofreading activity) and Pfu polymerases in ratios up to 16:1, respectively (results essentially indistinguishable from Fig. 2; data not shown).

The assay has proven to be an accurate and versatile assay that can be used to score SNPs in a cost-effective manner. We have used the assay to type moderate numbers of SNPs in genes associated with asthma (Van Eerdewegh et al. 2002), high bone density, drug metabolism, and numerous other conditions, demonstrating the utility of the assay for association studies and diagnostic applications. The preferred assay format is dictated by several factors, but primarily the number and availability of samples to be typed and the ease of assay design. For the present work, we elected to go with a primary PCR reaction followed by a separate Exo-Proofreading assay, rather than going directly from the genomic DNA. This format was chosen because of the limited supply of genomic DNA for the asthma samples. We also chose to go with a cleanup after the Exo-Proofreading assay rather than straight FP detection because of the higher success rate for assay design and higher reproducibility of fluorescence intensity measurements compared with FP measurements. In the comparison study all samples were processed in this manner. Examples of the other formats were presented to show the versatility and possibilities of the assay.

The use of PCR products as template introduces the chance of genotyping error caused by sample contamination (this is a problem for most SNP assays, the majority of which start with a PCR step). To reduce the chance of contamination, we physically segregate sample setup and postreaction workup in the laboratory. In addition, we use a significant amount of input template to reduce the impact of microdroplet cross-contamination. The data derived from over 2000 samples reported in Table 2 did not exhibit any signs of contamination, based on comparison with the ASO data. The potential for post-PCR contamination can be further reduced by using the linear assay format, or by reducing the number of thermal cycles to 20 or less.

The cost of the assay is determined largely by the price of the polymerase and the 3'-labeled primers. Although the cost of the 3'-labeled primers (\$40–\$60) is higher than the unlabeled primers used in some primer extension and allele-specific PCR-based assays, it is comparable to the cost of 5'-labeled or -biotinylated primers used in other primer extension assay formats, and less than the cost of double-labeled probes typically used in Taqman or molecular beacon assays. Of course, in most cases the initial cost of the primers for any given assay (labeled and unlabeled) is amortized over a large number of samples to be genotyped. The cost per genotype for two labeled and two unlabeled (PCR) primers is ~8 cents/sample for 2000 samples.

The labeled primers that we have used successfully all contain a linker arm to which the label is attached at the C5 position of a pyrimidine base. This configuration has proven successful in a number of other polymerase-compatible nucleotide labeling schemes (Langer et al. 1981; Foldes-Papp et al. 2001). The labeled primers could potentially be synthesized in a very cost-effective manner starting with controlled pore glass supports (CPGs) containing 5'-dimethoxytrityl-protected nucleosides with fluorescently labeled bases, because conventional primer synthesis proceeds in the 3'-to-5' direction (thus only the four standard phosphoramidites would be required during synthesis). However, only one of these labeled CPGs is available at the present time (3'-FAM C6 dT CPG; Glen Research). We have discussed the feasibility of synthesizing additional labeled CPGs with various manufacturers, and in the future, we will be exploring this option. However, for the present study, we used a less efficient alternative based on synthesis of

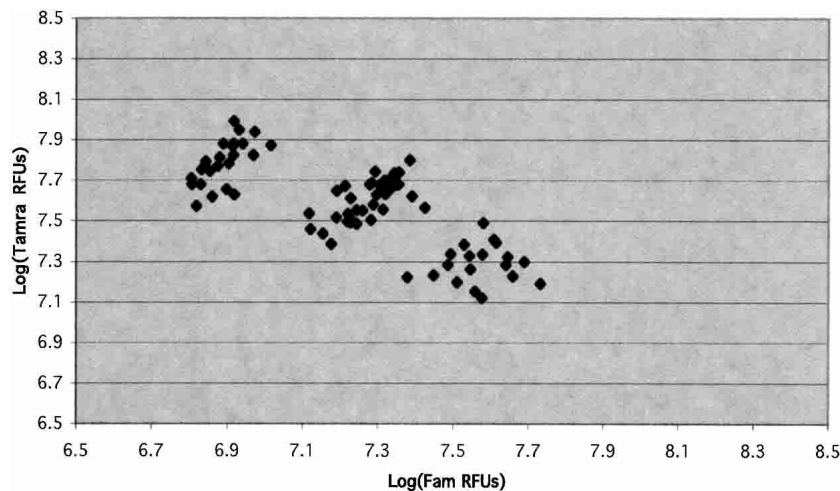


Figure 3 Exo-Proofreading assay performed on genomic DNA 216 S + 1. Data obtained when the Exo-Proofreading assay was performed directly on Genomic DNA. For the SNP assay, 22.5 ng of genomic DNA was used as a template; all other conditions remained the same.

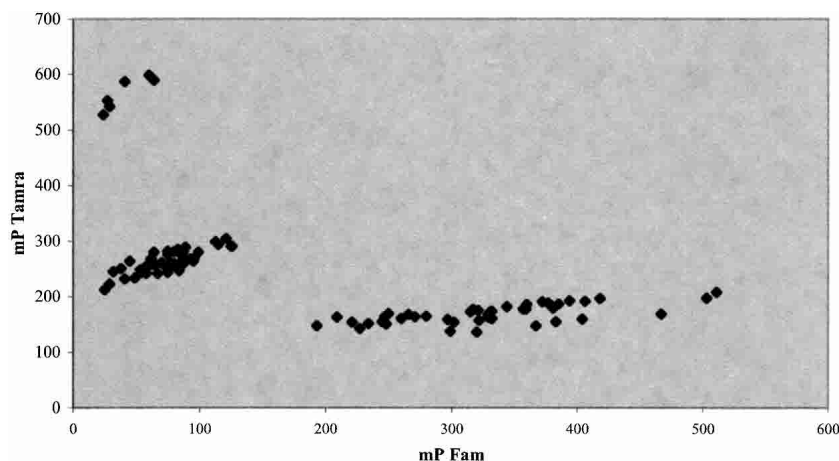


Figure 4 Detection by fluorescence polarization. Data obtained when using fluorescence polarization as a detection method with SNP 74D. FP detection was used on samples of unpurified Exo-proofreading reactions. This demonstrates the potential of the Exo-proofreading assay as a homogeneous method for SNP screening.

oligonucleotides starting with 3'-amino modifier CPGs (3'-Amino-Modifier C6 dC CPG and 3'-Amino-Modifier C6 dT CPG; Glen Research) followed by conjugation with the appropriate succinimidyl ester dye derivatives. These CPGs were used by our oligonucleotide suppliers to synthesize the desired oligonucleotides. One advantage to this two-step approach is the flexibility in the choice of fluorescent labels, because a large number of these are available commercially.

Unfortunately, there are no suitable 3'-amino- or fluorescent-modified purine CPGs (dA or dG) available at this time. However, the assay can still be used to type all possible base variations by typing both DNA strands. T/C and G/A polymorphisms are assayed using two 3'-pyrimidine-labeled primers that hybridize to the same strand. T/A, C/A, T/G, and C/G polymorphisms are assayed by using a four-primer configuration in which two 3'-pyrimidine-labeled primers overlap the position of the SNP on opposite strands and are flanked by two reverse primers on either side (see Fig. 1B). Although this approach limits the flexibility of primer design somewhat, it has proven to work effectively. The data from SNPs 216 S2, 216 ST + 4, and 216 S + 1 given in Table 2 were derived from four-primer assays of this type.

An inexpensive plate reader is the only instrumentation required for the assay. The potential to perform the assay directly on genomic DNA combined with the use of fluorescence polarization to detect the product provides one of the simplest SNP assay configurations yet described because the assay can be set up and performed and the products detected in a single reaction vessel in a truly homogeneous manner. However, further development work is required to address the apparent lack of robustness in fluorescence polarization detection. The assay has the potential to be multiplexed by using tagged primers in conjunction with universal tag arrays or microbeads with appropriate capture oligonucleotides.

METHODS

DNA Samples

The human DNA samples used in this study have been described by Van Eerdewegh et al. (2002). Samples were handled and the results analyzed using IRB-approved protocols.

Oligonucleotide Primers

The oligonucleotide primers used in this study are described in Table 1. Primers with 3'-amino-modified nucleotides were synthesized (Amitof Biotech) using 3'-amino modifiers having a linker attached to the 5-position of terminal pyrimidine moiety (3'-Amino-Modifier C6 dC CPG and 3'-Amino-Modifier C6 dT CPG; Glen Research). The resulting oligonucleotides were labeled using the following reaction conditions: 1 $\mu\text{g}/\mu\text{L}$ custom synthesized oligonucleotide, 2 $\mu\text{g}/\mu\text{L}$ succinimidyl ester dye derivative (FAM, ROX, or TAMRA; Molecular Probes), 75 mM sodium tetraborate (pH 8.5) at room temperature for 3 h. Labeled primers were purified by size exclusion chromatography using NAP columns (Amersham Biosciences) as per the manufacturer's instructions and concentrated using Nanosep microcentrators (Pall Filtron Corp.).

Primary PCR

All reactions described below were performed in 96-well polypropylene plates (Marsh BioMedical Products). Primary amplification reactions, 10 μL reaction volume, contained 22.5 ng of genomic DNA, 1 \times reaction buffer (60 mM Tris-HCl at pH 9.5, 15 mM $[\text{NH}_4]_2\text{SO}_4$, 2mM MgCl_2), 150 μM dNTPs (Roche), 500 nM forward and reverse primers (Invitrogen), 0.8 U of AmpliTaq (Applied Biosystems), and 1.6 U of TaqStart antibody (Clontech). Plates were sealed using Sealplate adhesive polyester film (Marsh), and reactions were performed in an MJ Research PTC-100 Thermocycler (MJ Research) using the following program: 94°C for 2 min followed by 35 cycles of 98°C for 10 sec, 68°C for 2 min, and a final extension step at 72°C for 3 min.

Exo-Proofreading Assay

The Exo-Proofreading assay was performed using either primary PCR amplified from genomic DNA or directly from genomic DNA. Primary PCR products (see above) were diluted 1:1000 with H_2O , and 1.3 μL of diluted sample was transferred to a 96-well plate (Marsh BioMedical); alternatively, 22.5 ng of genomic DNA was used. Sequences were amplified in a reaction containing 200 μM dNTPs, 1 \times polymerase buffer (60 mM Tris-HCl at pH 8.5, 15 mM $[\text{NH}_4]_2\text{SO}_4$, 1.5 mM MgCl_2), 300 nM each primer, and 0.625 U of Pwo DNA Polymerase (Roche). The reaction cycle conditions were as follows: 94°C for 1 min followed by 35 cycles of 94°C, 30 sec at the relevant annealing temperature (see T_A in Table 1), and 35 sec at 72°C, with a final extension step of 5 min at 72°C. Two-, three-, and four-primer assays were performed the same except that the linear or two-primer assay used 1:100 dilution of primary PCR product.

Allele Frequency

Samples of genomic DNA were quantitated using A260 measurements and the PicoGreen quantitation assay (Molecular Probes). Standard curves were prepared by combining homozygous samples of each allele together at different ratios (w/w) ranging from 100:0, 95:5, 90:10, down to 0:100. The standard Exo-Proofreading assay was performed on these mixed samples in groups of four. The ratio of the $\log(\text{rox})$ values to the $\log(\text{fam})$ values was fitted to the known frequency using linear regression to determine the standard curve. Then the

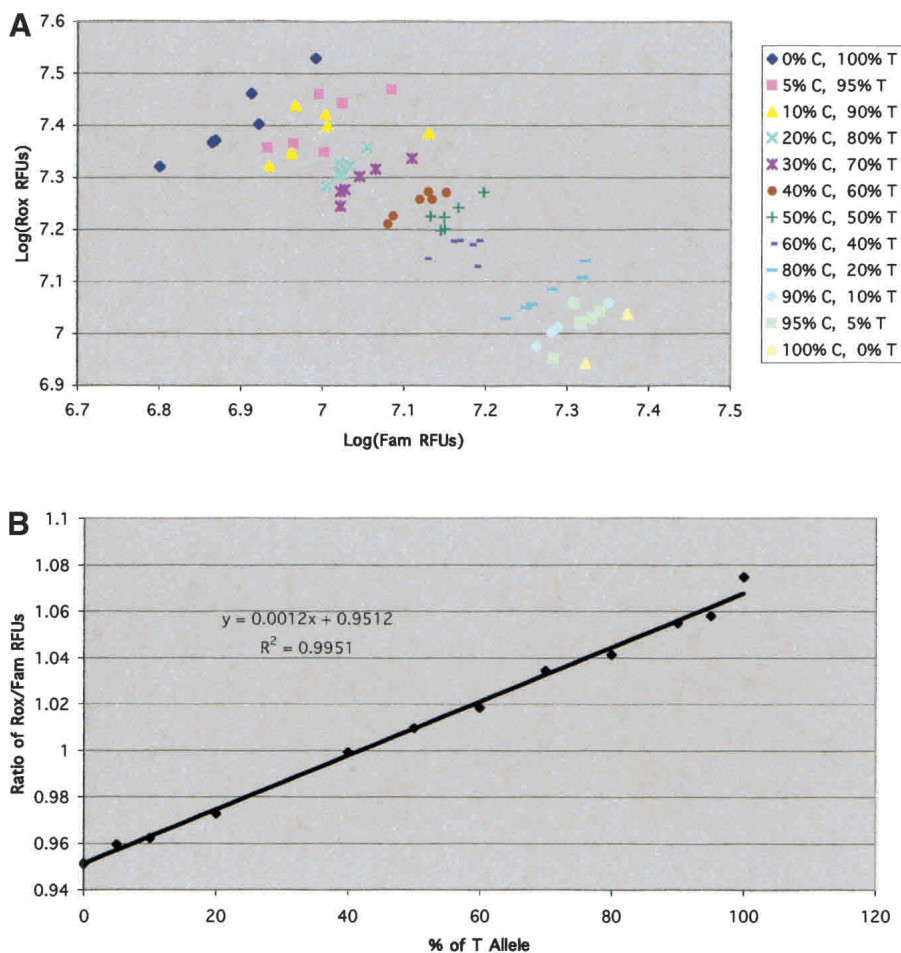


Figure 5 Standards for allele frequency determination. (A) Fluorescence intensity measurements obtained from samples containing known proportions of two alleles (in this case, C and T). Each sample group has $N = 6$ except 100% C, for which $N = 2$. (B) Plot of the ratios of Fam and Rox fluorescence intensity measurements against the percentage of T allele in the sample. Linear regression analysis reveals a linear relationship.

pool frequencies can be estimated with the following formula: $(\log[\text{rox}]/\log[\text{fam}] - \text{intercept estimate})/\text{slope estimate}$. Standard errors for the pool frequency estimates can be obtained using calibration methodology (Miller 1997, page 181.)

PCR Purification and Fluorescence Detection

After amplification, unincorporated primers and cleaved nucleotides were removed using the QIAquick-96 PCR purification kit (QIAGEN) according to the manufacturer's instructions. The elution volume was 65 μL , 25 μL of which was transferred into a black HE microplate (Molecular Devices). Incorporation of the two fluorescent markers into the PCR products was measured on an LJL Analyst HT multi-mode plate reader (Molecular Devices) using the following filters: ROX, excitation 580 nm, emission 610 nm; FAM, excitation 490 nm, emission 520 nm; TAMRA, excitation 550 nm, emission 580 nm.

Fluorescence Polarization

The Exo-Proofreading assay was performed exactly as described above. The mix was transferred to a 96-well black HE microplate and read with an LJL Analyst HT multi-mode plate reader in fluorescent polarization mode. The machine was set

with a G-factor of 0.85 for Fam and 1 for Tamra/Rox, with the light filters as previously described.

Allele-Specific Oligonucleotide Hybridization (ASO)

Individuals were typed using an allele-specific oligonucleotide method (Dietz et al. 1991). Amplicons were amplified by PCR, using the appropriate primers (see Table 1 for allele primers), and the PCR products were separated briefly by electrophoresis on agarose gels, blotted to nylon membranes, and then hybridized sequentially with radiolabeled oligonucleotide probes representing each allele.

Data Analysis

Genotypes were assigned based on the fluorescence values of both dyes simultaneously. A Bayesian algorithm based on a mixture of bivariate normal distributions was fitted. Prior means were determined using data points with known genotypes, when available. In addition, the algorithm computed a posterior probability of the correctness of each genotype assignment used as a confidence value. Details of this algorithm will be published separately.

Confirmation Sequencing

PCR products were diluted 1:20-fold, and 2.3 μL of the diluted sample was added to a Big Dye terminator sequencing reaction (Applied Biosystems). The sequencing reaction was performed according to the manufacturer's directions with the following exceptions. Cycle sequencing reactions were performed in one-half the standard volume and with one-half the standard amount of reaction mix. The reaction products were fractionated on the ABI 3700 instrument. Products were analyzed using ABI analysis software with visual inspection used to ascertain the nature of the genotype.

ACKNOWLEDGMENTS

We thank the GenomeVision Services group for their valuable assistance with the confirmation sequencing. Also we thank the Human Genetics group at GENE for their assistance, especially Ziyang Liu and Kathy Falls for their help with the samples and ASO data.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

- Braun, A., Little, D.P., Reuter, D., Muller-Mysok, B., and Koster, H. 1997. Improved analysis of microsatellites using mass spectrometry. *Genomics* **46**: 18–23.

- Chen, X., Levine, L., and Kwok, P.Y. 1999. Fluorescence polarization in homogeneous nucleic acid analysis. *Genome Res.* **9**: 492–498.
- Chetverin, A.B. and Kramer, F.R. 1993. Sequencing of pools of nucleic acids on oligonucleotide arrays. *Biosystems* **30**: 215–231.
- Cline, J., Braman, J.C., and Hogrefe, H.H. 1996. PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res.* **24**: 3546–3551.
- Dabrowski, S. and Kur, J. 1998. Cloning and expression in *Escherichia coli* of the recombinant his-tagged DNA polymerases from *Pyrococcus furiosus* and *Pyrococcus woesei*. *Protein Expr. Purif.* **14**: 131–138.
- Dietz, H.C., Pyeritz, R.E., Hall, B.D., Cadle, R.G., Hamosh, A., Schwartz, J., Meyers, D.A., Francomano, C.A. 1991. The Marfan syndrome locus: confirmation of assignments to chromosome 15 and identification of tightly linked markers at 15q15-q21.3. *Genomics* **9**: 355–361.
- Drake, J.W., Allen, E.F., Forsberg, S.A., Preparata, R.M., and Greening, E.O. 1969. Genetic control of mutation rates in bacteriophage T4. *Nature* **221**: 1128–1132.
- Foldes-Papp, Z., Angerer, B., Thyberg, P., Hinz, M., Wennmalm, S., Ankenbauer, W., Seliger, H., Holmgren, A., and Rigler, R. 2001. Fluorescently labeled model DNA sequences for exonucleolytic sequencing. *J. Biotechnol.* **86**: 203–224.
- Kunkel, T.A., Loeb, L.A., and Goodman, M.F. 1984. On the fidelity of DNA replication. The accuracy of T4 DNA polymerases in copying ϕ X174 DNA in vitro. *J. Biol. Chem.* **259**: 1539–1545.
- Langer, P.R., Waldrop, A.A., and Ward, D.C. 1981. Enzymatic synthesis of biotin-labeled polynucleotides: Novel nucleic acid affinity probes. *Proc. Natl. Acad. Sci.* **78**: 6633–6637.
- Lipshutz, R.J., Morris, D., Chee, M., Hubbell, E., Kozal, M.J., Shah, N., Shen, N., Yang, R., and Fodor, S.P. 1995. Using oligonucleotide probe arrays to access genetic diversity. *Biotechniques* **19**: 442–447.
- Marshall, D.J., Heisler, L.M., Lyamichev, V., Murvine, C., Olive, D.M., Ehrlich, G.D., Neri, B.P., and de Arruda, M. 1997. Determination of hepatitis C virus genotypes in the United States by cleavage fragment length polymorphism analysis. *J. Clin. Microbiol.* **35**: 3156–3162.
- Mattila, P., Korpela, J., Tenkanen, T., and Pitkanen, K. 1991. Fidelity of DNA synthesis by the *Thermococcus litoralis* DNA polymerase—An extremely heat stable enzyme with proofreading activity. *Nucleic Acids Res.* **19**: 4967–4973.
- Miller Jr., R.G. 1997. *Beyond ANOVA*. Chapman & Hall, London, UK.
- Newton, C.R., Graham, A., Heptinstall, L.E., Powell, S.J., Summers, C., Kalsheker, N., Smith, J.C., and Markham, A.F. 1989. Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res.* **17**: 2503–2516.
- Nickerson, D.A., Kaiser, R., Lappin, S., Stewart, J., Hood, L., and Landegren, U. 1990. Automated DNA diagnostics using an ELISA-based oligonucleotide ligation assay. *Proc. Natl. Acad. Sci.* **87**: 8923–8927.
- Nyren, P., Karamohamed, S., and Ronaghi, M. 1997. Detection of single-base changes using a bioluminometric primer extension assay. *Anal. Biochem.* **244**: 367–373.
- Pease, A.C., Solas, D., Sullivan, E.J., Cronin, M.T., Holmes, C.P., and Fodor, S.P. 1994. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc. Natl. Acad. Sci.* **91**: 5022–5026.
- Preston, B.D., Poiesz, B.J., and Loeb, L.A. 1988. Fidelity of HIV-1 reverse transcriptase. *Science* **242**: 1168–1171.
- Saiki, R.K., Bugawan, T.L., Horn, G.T., Mullis, K.B., and Erlich, H.A. 1986. Analysis of enzymatically amplified β -globin and HLA-DQ α DNA with allele-specific oligonucleotide probes. *Nature* **324**: 163–166.
- Syvanen, A.C., Aalto-Setälä, K., Harju, L., Kontula, K., and Soderlund, H. 1990. A primer-guided nucleotide incorporation assay in the genotyping of apolipoprotein E. *Genomics* **8**: 684–692.
- Uhlen, M. 1989. Magnetic separation of DNA. *Nature* **340**: 733–734.
- Van Eerdewegh, P., Little, R.D., Dupuis, J., Del Mastro, R.G., Falls, K., Simon, J., Torrey, D., Pandit, S., McKenny, J., Braunschweiger, K., et al. 2002. Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. *Nature* **418**: 426–430.

Received October 30, 2002; accepted in revised form February 18, 2003.