2003

Kinetic FP-TDI assay for SNP allele frequency determination

Ming Xiao
Washington University School of Medicine in St. Louis

Sherif Medhat Latif
Washington University School of Medicine in St. Louis

Pui-Yan Kwok
Washington University School of Medicine in St. Louis

DRUG DISCOVERY
AND GENOMIC TECHNOLOGIES

Research Report

Kinetic FP-TDI Assay for SNP Allele Frequency Determination

BioTechniques 34:190-197 (January 2003)

Ming Xiao, Sherif Medhat Latif, and Pui-Yan Kwok
Washington University School of Medicine, St. Louis, MO, USA

ABSTRACT

Strategies for identifying genetic risk factors in complex diseases by association studies require the comparison of allele frequencies of numerous SNPs between affected and control populations. Theoretically, hundreds of thousands of SNP markers across the genome will have to be genotyped in these studies. Genotyping SNPs one sample at a time is extremely costly and time consuming. To streamline whole genome association studies, some have proposed to screen SNPs by pooling the DNA samples initially for allele frequency determination and perform individual genotyping only when there is a significant discrepancy in allele frequencies between the affected and control populations. Here we describe a new method for determining the allele frequency of SNPs in pooled DNA samples using a two-color primer extension assay with real-time monitoring of fluorescence polarization (named kinetic FP-TDI assay). By comparing the ratio of the rate of incorporation of the two allele-specific dye-terminators, one can calculate the relative amounts of each allele in the pooled sample. The accuracy of allele frequency determination with pooled samples is within 3.3 ± 0.8% of that determined by genotyping individual samples that make up the pool.

INTRODUCTION

SNPs are the most abundant sequence variation among individuals. When two human genomes are compared, single base-pair variations are found at approximately 1200 nucleotide intervals. Because of their abundance and low mutation rate, SNPs are the markers of choice in association studies to identify the genetic risk factors in common diseases (10,13). In addition to the formidable task of acquiring the large number of SNP markers and choosing the right ones needed in these studies, a major challenge remains the cost and time needed to obtain the genotypes of numerous samples with hundreds of thousands of SNPs.

To reduce the time and cost associated with genotyping every individual in a study, it has been proposed that one works with pooled DNA samples constructed by mixing equal amounts of DNA from groups of individuals. Although the pooling strategy has some drawbacks, such as reduced insight into haplotype structure and lower precision when dealing with rare alleles, it is an appealing approach for initial screening of known SNPs in association studies. Several methods have been developed to estimate allele-frequencies in pooled samples (1,3,4,6,11,18–20). Because analysis of pooled samples requires only two reactions (for pools of cases and controls) per SNP, low assay development cost is of paramount importance. We report here an allele frequency determination method that is comparable in accuracy to the best methods currently in use but has the lowest assay development cost compared to all other approaches. The method is based on the template-directed dye-terminator incorporation (TDI) assay with fluorescence polarization (FP) detection (2,8). The only modification of the FP-TDI assay is the real-time detection strategy. At the end of PCR, a SNP probe is extended by one base when incubated with the two allele-specific dye-terminators in the presence of DNA polymerase. Because an incorporated dye-terminator is part of a much larger molecule, its fluorescence remains polarized (after excitation by plane-polarized light) for a longer time than that of a much smaller, free dye-terminator. If one monitors the FP of the two allele-specific dye-terminators throughout the primer extension reaction, then the rate of FP change can be followed. The rate of FP increase of a dye-terminator is proportional to the amount of DNA containing the corresponding allele present in the reaction. Therefore, one can calculate the relative abundance of each allele in a pooled sample by comparing the rate of FP change of the two dye-terminators.

The kinetic FP-TDI assay we developed is found to be very accurate, with the estimated allele frequencies differing only 3.32 ± 0.8% from the allele frequencies determined by genotyping all the individuals in the pool. Because only three unlabeled, "PCR-grade" primers are needed for each SNP, the development of kinetic FP-TDI assays is very simple and cost effective. Three reactions are needed for each marker (two population samples plus one reference sample). Therefore, one can determine the allele frequencies of 128 SNPs across two pooled samples in a 6-h experiment using a 384-well reaction for-

mat. It will take only 65 days to screen through 100 000 SNPs if one performs 12 plates of assays per day (one PCR machine with four heating blocks doing three runs/day even without the use of a real-time TDI-FP instrument) at a fraction of the cost of performing individual genotypes. Therefore, the kinetic FP-TDI screening approach can greatly reduce the cost and duration of an association study.

Principle of the Method

The FP-TDI assay (2) is a homogeneous SNP genotyping assay that is based on the fact that the degree of fluorescence polarization of a dye is a reflection of its molecular weight. The larger the molecular weight, the higher is the dye molecule's FP value. In the presence of DNA polymerase and the appropriate dye-terminators, the SNP primer (with its 3′ end adjacent to the polymorphic site) is extended specifically by one base as dictated by the target sequence. The primer extension products will have higher molecular weight and in turn higher FP value (12). When genotyping individual samples, one obtains the FP values of the two allelic dye-terminators after linear amplification during the primer extension step with thermal cycling. The allele(s) present in the DNA sample can be inferred by determining which of the two dye-terminators is incorporated in the reaction.

To measure the allele frequencies of a SNP in a pooled DNA sample, the incorporation rates of the two dye-terminators are compared kinetically, cycle by cycle. A heterozygous sample is used as the reference to determine the incorporation rates of the two alleles when they are present in a 50:50 ratio in the sample. These incorporation rates are then used in determining the relative amounts of each allele present in the pooled samples.

In a mixture of containing several fluorescent species bearing the same fluorophore, the total polarization FP of the mixture is given by the following expression (17):

$$(1/FP-1/3)= \Sigma f_i/(1/FP_i -1/3) \qquad [Eq. 1]$$

where f_i is the fraction of the ith species bearing the fluorophore being studied and FP_i is the FP value of the ith species.

When the above equation is expressed in anisotropy A (A=2FP/(3-FP)) (9), it gives a mathematically simpler equation:

$$A=\Sigma f_i A_i \qquad [Eq. 2]$$

where A_i is the anisotropy value of the ith species bearing the fluorophore under study.

In the FP-TDI assay, two species exist for each of the two allelic dye-terminators. For their anisotropy values at any particular cycle, we have the unincorporated dye-terminator ("free dye") at fraction f_f with a lower anisotropy value A_f and the incorporated dye-terminator ("bound dye") at fraction f_b with a higher anisotropy value A_b. The observed anisotropy, A, at any given time is therefore:

$$A=f_f A_f + f_b A_b \qquad [Eq. 3]$$

where $f_f + f_b = 1$.

Initially, all dye-terminators are unincorporated and the anisotropy reading is A_f, since $f_f = 1$ and $f_b = 0$. When the reaction is driven to completion, all dye-terminators are incorporated into primer and the anisotropy reading becomes A_b, with $f_f = 0$ and $f_b = 1$. The fraction of free and bound dye-terminators changes in opposite directions during the primer extension reaction, but the rate of this change depends on the amounts of the DNA template in the reaction.

In general, the rate of dye-terminator incorporation is determined by two factors: the amount of starting materials (PCR-amplified DNA fragments, SNP primers, and dye-terminators) and the incorporation efficiency of the dye-terminator by the DNA polymerase. If there are X copies of PCR-amplified fragments, then pX copies PCR fragments contain allele 1 and (1-p)X copies contain allele 2 (where p is the allele frequency of allele 1). In TDI reaction, the SNP primers and the dye-terminators are in vast excess over the PCR products, and the two dye-terminators added are in equal amounts (Y molecules each). Initially, dye-terminator incorporation is linear. As the dye-terminators and SNP primer are being consumed, dye-terminator incorporation becomes nonlinear and the anisotropy observed reaches a plateau when dye-terminators are used up.

During the first cycle of the primer extension reaction, each PCR-amplified fragment has one SNP primer hybridized to it. If every hybridized SNP primer gets extended, then pX and (1-p)X dye-terminators are incorporated onto all annealed SNP primers for each of the two dye-terminators, respectively. However, in reality, neither the hybridization nor the primer extension reaction is 100% efficient. Dissimilar incorporation efficiencies of the dye-terminators will result in uneven incorporation of the two dye-terminators (7). Denoting V_1 as the incorporation efficiency for the dye-terminator corresponding to allele 1 and V_2 for the dye-terminator corresponding to allele 2, pXV_1 is the number of allele 1-specific dye-terminators incorporated and $(1-p)XV_2$ is the number of allele 2-specific dye-terminators incorporated at the end of cycle one, since the incorporation of the two dye-terminators can be reasonably assumed to be independent. Accordingly, f_f can be represented as $(Y- pXV_1)/Y$ for allele 1 and $(Y-(1-p)XV_2)/Y$ for allele 2. Recalling that f_b equals to $1-f_f$, after n cycles of linear incorporation, f_f becomes $(Y-npXV_1)/Y$ for allele 1 and $(Y-n(1-p)XV_2)/Y$ for allele 2. Substituting f_f and f_b of the two alleles into Equation 3 separately and rearranging them, Equation 3 becomes Equation 4 for allele 1 and Equation 5 for allele 2.

$$(A_1-A_{1f})/(A_{1b}-A_{1f})=np(X/Y)V_1 \qquad [Eq. 4]$$

$$(A_2-A_{2f})/(A_{2b}-A_{2f})=n(1-p)(X/Y)V_2 \qquad [Eq. 5]$$

Equations 4 and 5 have four variables each: X copies of amplified PCR fragments, Y molecules of added dye-terminators, the incorporation efficiency V, and allele frequency p. X and Y vary from well to well and are hard to quantify, so it is impossible to calculate the allele frequency value, p, by solving Equations 4 and 5 separately. By taking the ratio of equations 4 and 5, two of the variables (X and Y) are canceled out. If one estimates the ratio of the dye-terminator incorporation efficiency ($K = V_1/V_2$) using a heterozygous individual (with "allele frequency" of each allele at 0.5 by definition) as reference, then the

Table 1. SNPs Used for the Evaluation of the Method

dbSNP markers	Allele	PCR primer 1	PCR primer 2	SNP primer 1	Allele detected by SNP primer 1	SNP Primer 2	Allele detected by SNP primer 2
D7s8	C/T	CTCAGGGACCCTGACCTTATTG	GCCTTCTCAGAGCTTATATTATTGTGTC	CCTGACCTTATTGCTCCCCTTTCC	C/T	TTATTGTGTCAGGACCAGCATTCC	G/A
rs1169	T/C	GCAGATCAGTCAGCCCACTT	CCCCATCCTACCAGTTGGT	AGAGGCTAGGCAGTGAACACATCA	C/T	CCAGTTGGTTATTTTCTCATTGCATAC	G/A
rs221	A/G	CCCCAGTATCTGGCACATC	TTTCTCAGGTTTCAGGGATTAGA	GCACATCTTTCCCTTTTCATCTCC	G/A	GATATTATTTGGCCAAACACACAAA	C/T
rs3157	G/A	CCGGCTGATTTTCTTCGACA	CCTAGGCCTCATCTGTGGGAA	AGGAGCTGTGGTGGGGGGCAGTAT	G/A	TCTGTGGGAAGGGAGTCCCTGGCT	C/T
rs3354	G/A	CATTCTGGTTTTGACATCAGCATT	GCAAAAGGTGCCATGGTGTT	ATCAGCATTAGTCACTTTGAAATGTAAC	G/A	AAACTTGGAATTGGTTGTAGTACCATT	C/T
rs1541612	G/A	TTTGGAGTCCTGCACTTTAC	TTATCTGGTGATTGAGGAAGA	CACGTTTTCCTGCTGCCT	G/A	N/A	
rs1951906	G/A	AAGGAGCAGCATTTTGACTA	CACGCTTCCTATTTCTCATC	TTTTTGTTAAATTAGTGCTGTCTACATTTT	G/A	N/A	
rs681	T/C	TGGGGCTTGACTTTCCAA	CAAATGCCTAATAGACACTATTGGTAGAT	N/A		TATTGGTAGATACCTTCTTGGCCC	G/A
rs688	T/C	TCCAAACTTCACTCCATCTCAA	GGGCCAGCCTCTTTTCAT	N/A		CAAGATGGTCTTCCGGTTGCCCCC	G/A
rs696	G/A	AAAAGAAGAAAAAATTTAAAGGGTGTACTTAT	AATGAGGGCTGATCCTACCAC	TTATATCCACACTGCACACTGCCT	G/A	N/A	
rs231	C/G	GAGGTTTCACTGGCTTGTGCT	CATGAGACATTTATCTAATGATTTTTTCTTATT	GCCTTGGGGCCCACTACTGCCATG	G/C	GAAGAAAGAGACAGCGATTGGCTAAC	G/C
rs241	A/C	TGCTGAATTTCCATCTCTGAGTTC	GCAGGATTCAGTGCCAGAAAG	AGATTTCAGAAAATATGATTAGAAAAAG	A/C	CAGAAAGTTATATTTTAAGATGTACCAC	G/T
rs245	del/T	TAAAGATGTTGGAGAACTGAAAAAGA	GACCACGGAAGGACACATACAT	AAAAAGAGAGCTTACATGCACCCC	T/A	GGAAATGTGTGGAGAGTTTTGCTATT	G/A

allele frequency, p, is now only related to the average anisotropy observed at the end of any particular cycle.

Although taking just one FP reading during the early phase of the primer extension reaction will yield a reasonable allele frequency estimate, the fact that this approach takes only one measurement during the presumed linear phase of the reaction makes it highly susceptible to random errors. Instead of taking the ratio of Equations 4 and 5 directly, however, one can monitor the FP at the end of every cycle and plot $(A-A_f)/(A_b-A_f)$ against cycle n for each of the two dye-terminators. Two straight lines will be obtained for the initial linear incorporation stage, and the slopes of the lines will be $p(X/Y)V_1$ for allele 1 and $(1-p)(X/Y)V_2$ for allele 2. Since X, Y, V_1, and V_2 remain constant for the same reaction, the ratio of the two slopes becomes Equation 6:

$$\frac{\text{Slope for allele 1 } (S_1)}{\text{Slope for allele 2 } (S_2)} = \frac{pV_1}{(1-p)V_2} \quad \text{[Eq. 6]}$$

with X and Y canceled out as before. Rearranging Equation 6 to solve for p, allele frequency can be obtained for the pooled DNA sample as Equation 7:

$$p = 1/(1+(S_2/S_1)K) \quad \text{[Eq. 7]}$$

S_1 and S_2 are determined in the reaction involving the pooled sample, while K is determined from the reaction using the heterozygous sample. Because the kinetic FP-TDI approach includes many measurements during the primer extension reaction, the variation is reduced significantly.

MATERIALS AND METHODS

DNA Samples and Pool Construction

DNA samples of 96 anonymous individuals were obtained from the Coriell Institute for Medical Research (Camden, NJ, USA). The DNA samples include 32 individuals each from the African-American, Asian-American, and European-American panels. We used both absorbance at 260 nm and a DNA specific fluorescence dye, PicoGreen® (Molecular Probes, Eugene, OR, USA), for the quantification of individual DNA samples. PicoGreen showed higher sensitivity and was used for fine-tuning the concentration. Equal amounts of DNA from individual samples were pooled together to construct three pools with final concentration of 3 ng/µL.

SNP Markers

Publicly available markers from the dbSNP database were used in this study. These SNPs were previously characterized by our group. Table 1 lists the dbSNP assay numbers, reference numbers, and primer sequences.

Assay Reagents

All reactions were run and read in 96-well black plate from MJ Research (Waltham, MA, USA). PCR and SNP primers were obtained from Invitrogen (Carlsbad, CA, USA). AmpliTaq Gold® DNA polymerase was from Applied Biosystems (Foster City, CA, USA). Exo-Sap® (Exonuclease I and Shrimp Alkaline phosphatase) was purchased from USB (Cleveland, OH, USA). SNP detection kits (AcycloPrime™) were generous gifts from Perkin Elmer (Boston, MA, USA) and include 10× reaction buffer, AcycloPol™ enzyme, and AcycloTerminators™ mixture consisting of equal amounts of R110 and Tamra acycloterminators™.

PCR Amplification, Degradation of Excess PCR Primers, and dNTPs

DNA (3 ng) was amplified in 5 µL reaction mixtures containing 2 mM

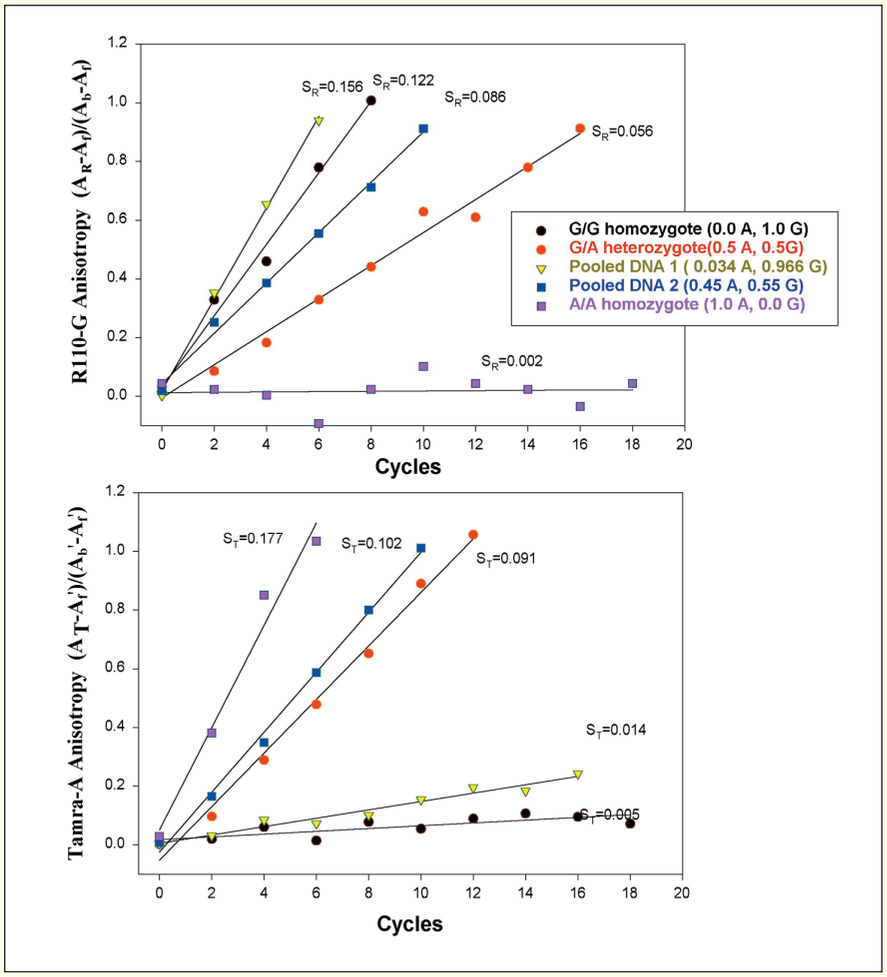

Figure 1. Plot of anisotropy versus cycles of primer extension reaction for rs688 (G/A alleles). The scatter plots are the anisotropy readings from microplate reader. The solid lines are the linear regression of the scatter plots. Error bars represent standard deviation. (A) The plot for R110-G terminator. (B) Plot for Tamra-A terminator.

$MgCl_2$ and 125 nM each primer. The reaction mixture was heated for 12 min at 95°C for enzyme activation followed by 35 amplification cycles. Each cycle consisted of 30 s of denaturation at 95°C, 40 s primer annealing at 58°C, and 40 s primer extension at 68°C. The reaction mixtures were then incubated at 72°C for 5 min for final primer extension. At the end of the reaction, the reaction mixtures were held at 4°C until further use. Two microliters of ExoSap were added to the PCR mixtures and incubated for 45 min at 37°C to degrade the excess PCR primers and dNTPs. The enzymes were heat-inactivated at 95°C for 15 min.

Primer Extension Reaction

SNP primers were designed using PrimerExpress™ (Applied Biosystems) to have a universal annealing temperature of 60°C. To the reaction mixtures from the previous step, 13 µL TDI cocktail were added according to the manufacturer's instructions [2 µL 10× TDI reaction buffer, 0.5 µL SNP primer (final concentration 384 nM), 0.05 µL Acyclo Enzyme, 1 µL two-dye Acyclo nucleotide terminators, and 8.95 µL water]. The reaction mixture was incubated at 93°C for 45 s, two cycles of single-base extension at 93°C for 10 s and 58°C for 30 s, and another 10 s of denaturation at 95°C before being taken out manually and reading on an LJL Analyst™ fluorescence reader (Molecular Devices, Sunnyvale, CA, USA). This procedure was repeated several times until the reaction goes well into the completion by inspecting the polarization readings. When individual samples were genotyped, the primer extension reaction was allowed to proceed for 20 cycles, and the fluorescence polarization readings were done just once at the end of the reaction.

Data Analysis

Data were analyzed with Microsoft® Excel® and SigmaPlot software. Linear regression was performed, and results were discarded with R^2 less than 0.95.

RESULTS

Figure 1 shows one typical real-time FP-TDI experiment of using SNP rs688 from the dbSNP database. The assays of five different samples were shown, including three individual samples (G/G, A/A homozygotes, and G/A

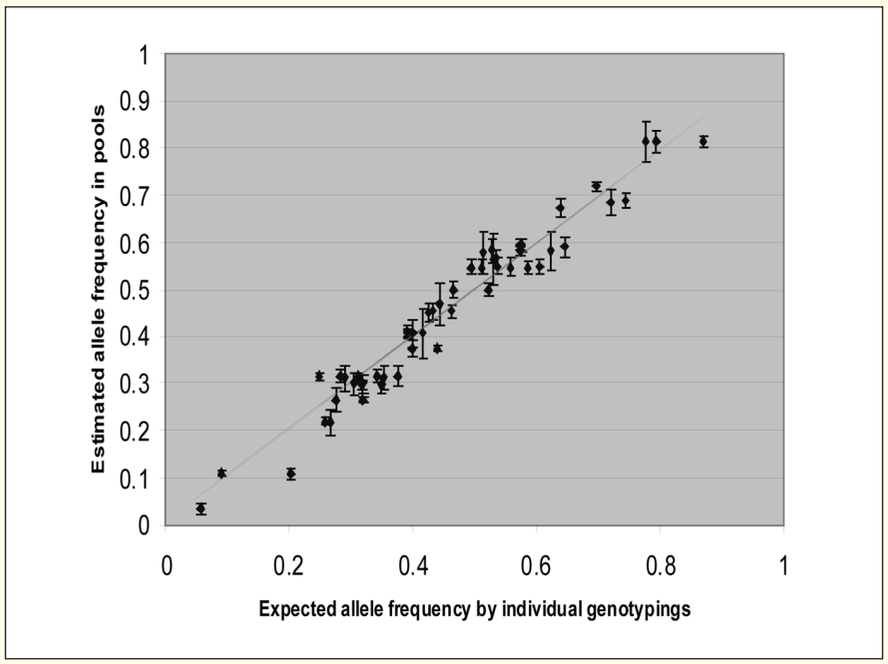

Figure 2. Allele frequency measurements in pools (y-axis) versus allele frequency obtained by individual genotypings (x-axis). A linear relation was observed $R^2 = 0.978$, slope = 1.015 ($P < 0.0001$).

heterozygote) and two pooled DNA samples of known composition based on individual genotyping of all the samples found in the pool by FP-TDI. Linear regression of the real-time anisotropy readings was performed. The higher the slope of the linear regression, the faster was the incorporation of a particular dye-terminator and the fewer the number cycles were needed to use up the dye-terminators. For the G/G homozygote, the slope of R110-G terminator has a high value of 0.122, while the slope of Tamra-A is very low (at 0.005). This is because the G/G sample incorporates R110-G terminators but not the Tamra-A terminators. The opposite was observed for the A/A homozygote, which has the high slope value for Tamra-A terminators of 0.177 and the low value of 0.002 for R110-G terminator. As for the G/A heterozygote, the slopes of both the R110-G and the Tamra-A lie in the middle. However, their values are not exact the same, 0.056 for R110-G versus 0.091 for Tamra-A. This indicates that the AcycloPol enzyme incorporates R110-G and Tamra-A with different efficiencies. As indicated before, this differential incorporation efficiency K (= $V_{Tamra\text{-}A}/V_{R110\text{-}G}$) can be calculated by taking the ratio of the two slopes for a reference heterozygous sample with known allele frequencies of 0.5 for both alleles. The K value for this particular marker is found to be 0.091/0.056 = 1.625. The incorporation efficiency of the Tamra-A is about 60% higher than that of R110-G. Pooled DNA sample 1 has the highest slope of R110-G with 0.156, which is even higher than that for the G/G homozygote. This seemingly contradictory result can be explained by the difference in the PCR-amplified fragments for the two reactions, with the PCR yield being higher for pooled sample 1 when compared to that for the G/G homozygote. The larger the amount of PCR products present, the faster the primer extension reaction will be. This variation in the yield of PCR is the very reason why one cannot accurately estimate the allele frequency by just looking at the incorporation of just one dye-terminator. Taking into account of differential incorporation efficiency of the R110-G and the Tamra-A (K = 1.625), the calculated allele frequency of allele A is 0.053 according to Equation 7. The difference between the true allele frequency obtained by genotyping individual samples in the pool 1 and allele frequency by real time FP-TDI is 1.9%. For the pooled DNA sample 2, the real-time FP-TDI assay found 0.421 allele frequency of allele A compared to 0.45 from individual genotyping, a 2.9% difference.

Allele frequencies of additional 13 SNP markers for three different pooled DNA samples were also obtained by the real-time FP-TDI assay using both sense and antisense SNP primers. The results are compared with the true allele frequencies by genotyping individual samples in the pool. The true allele frequencies of these 13 markers for three pooled samples cover a wide range of values, from the minor allele frequency 0.034 of marker rs688 in pooled sample 2 to 0.469 of marker rs681 in pooled DNA sample 3. The real-time FP-TDI assay works well in this wide range of allele frequencies and agrees well with the true allele frequencies across all three pooled DNA samples. Figure 2 shows the correlation between the estimated allele frequencies and the expected allele frequencies (obtained by genotyping individual samples) of these markers. The estimated allele frequency in pools correlated well with the allele frequency obtained by genotyping the individual (r = 0.978, slope = 1.015, $P < 0.0001$). One of the markers gave extremely poor estimated allele frequencies for all three pools for one SNP primer, with pool 1 yielding an estimate that was 17% higher than the observed allele frequency (rs241-C/A). When individual genotyping assays of rs241-C/A were performed, the polarization of R110-C was low and the assay did not produce four clear-cut clusters (C/C, A/A, C/A, and negative control clusters). Low FP values were an indication that the dynamic range of FP between free dye-terminator and incorporated terminators was small. Together with the unusually large disparity of the incorporation efficiency for Tamra-A and R110-C for this SNP (K = 2.49), the allele frequency estimate made is not reliable. To obtain good results of real-time FP-TDI assay, the underlying FP-TDI genotyping assay has to perform well. Using the antisense SNP primer for rs241, the allele frequency estimated by real-time FP-TDI assay agreed well with the known allele frequency across all three pooled DNA samples. Without considering the result of marker rs241-C/A, the overall average mean difference of allele frequency determined by real-time FP-TDI in pooled samples and that in individually genotyped samples is 3.32 ± 0.8% (5).

DISCUSSION

Effect of Differential Incorporation Efficiencies of Dye-Terminators

Depending on the local sequence context, the Acyclopol enzyme incorporates dye-terminators with different incorporation efficiency. For G/A (R110/Tamra) combination, AcycloPol enzyme prefers Tamra terminators for most of markers with K ranging from 1 to 1.89. Other dye-terminator combinations also give similar results. To calculate allele frequency accurately, one has to account for this differential incorporation rate using a heterozygous individual to estimate the K value. Finding a heterozygous individual for particular marker certainly adds additional cost and time and may even be difficult for some markers with low allele frequencies. Therefore, some individual genotyping assays must be done before the kinetic FP-TDI assay is performed on the pooled sample. To avoid this extra step of finding an individual heterozygous sample, one either finds an enzyme or designs new dye-terminators such that both dye-terminators are incorporated with equal efficiency. New enzymes and new dye-terminators are very difficult to come by, so performing some assays on individuals is the only practical solution at this time.

Sources of the Measurement Error and Potential Improvements

In the FP-TDI assay, we used the combination of R110 and Tamra terminators because of their lack of spectral overlap. However, extensive experience with these terminators showed

that the R110 fluorescence was quenched upon incorporation in many cases, with typical decreases of in fluorescence ranging from 30% to 90% (data not shown). The quenching by the SNP probe may involve complicated photophysical processes and complex conformation of dye-labeled oligonucleotides (14,15). This may be the reason for our observation that R110 terminators showed a wide range of polarization values, while the FP of Tamra terminators with very little quenching observed was rather stable. Low polarization value of R110 upon incorporation results in poor dynamic range, and in turn the linear regressions of cyclic data of R110 are not as good as those for Tamra. We are in the process of testing other dye-terminators with more stable polarization value upon their incorporation. Once the better dye-terminators are found, the accuracy of the allele frequency estimation by real-time FP-TDI assay will certainly be improved.

Improvements in instrumentation will also enhance the robustness of the assay. These proof-of-principle experiments are done by running the reaction in a thermal cycler for two cycles followed by FP reading in a microplate reader. This process is repeated every two cycles for 10–15 times until the TDI reaction goes well into completion. This is not only cumbersome (increasing the time it takes to perform the primer-extension reaction to 2 h/plate) but also will likely compromise the accuracy of the reaction. If a real-time FP plate reader is developed in the future, one can perform these experiments with little human handling, leading to enhancement of assay accuracy.

In summary, we have developed an allele frequency estimation strategy that is both accurate and cost-effective. Given the fact that FP-TDI assays have been designed by our group for 1.7 million SNPs found in the public databases (16), kinetic FP-TDI assays can be used for initial screening of SNPs in whole genome association studies, thereby reducing the time and cost associated with these large-scale studies. With improvements in instrumentation and dye-terminator incorporation rates, this approach will become even more robust and accurate.

REFERENCES

1. **Breen, G., D. Harold, S. Ralston, D. Shaw, and D. St. Clair.** 2000. Determining SNP allele frequencies in DNA pools. BioTechniques *28*:464-470.
2. **Chen, X., L. Levine, and P.Y. Kwok.** 1999. Fluorescence polarization in homogeneous nucleic acid analysis. Genome Res. *9*:492-498.
3. **Germer, S., M.J. Holland, and R. Higuchi.** 2000. High-throughput SNP allele-frequency determination in pooled DNA samples by kinetic PCR. Genome Res. *10*:258-266.
4. **Giordano, M., M. Mellai, B. Hoogendoorn, and P. Momigliano-Richiardi.** 2001. Determination of SNP allele frequencies in pooled DNAs by primer extension genotyping and denaturing high-performance liquid chromatography. J. Biochem. Biophys. Methods *47*:101-110.
5. **Glantz, S.A.** 1997. Primer of Biostatistics, 4th ed. McGraw-Hill, New York.
6. **Gruber, J.D., P.B. Colligan, and J.K. Wolford.** 2002. Estimation of single nucleotide polymorphism allele frequency in DNA pools by using Pyrosequencing. Hum. Genet. *110*:395-401.
7. **Haff, L.A. and I.P. Smirnov.** 1997. Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry. Genome Res. *7*:378-388.
8. **Hsu, T.M., X. Chen, S. Duan, R.D. Miller, and P.Y. Kwok.** 2001. Universal SNP genotyping assay with fluorescence polarization detection. BioTechniques *31*:560-568.
9. **Jaboński, A.** 1960. On the notion of emission anisotropy. Bull. Acad. Pol. Sci. Ser. A *8*:259-264.
10. **Kruglyak, L.** 1999. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat. Genet. *22*:139-144.
11. **Kwok, P.Y., C. Carlson, T.D. Yager, W. Ankener, and D.A. Nickerson.** 1994 Comparative analysis of human DNA variations by fluorescence-based sequencing of PCR products. Genomics *23*:138-144.
12. **Perrin, F.** 1926. Polarization de la lumiere de fluorescence. Vie moyene de molecules dans Iíetat excite. J. Phys. Radium *7*:390-401.
13. **Risch, N. and K. Merikangas.** 1996. The future of genetic studies of complex human diseases. Science *273*:1516-1517.
14. **Seidel, C.A.M., A. Schulz, and M.H.M. Sauer.** 1996. Nucleobase-specific quenching of fluorescent dyes.1. Nucleobase one-electron redox potentials and their correlation with static and dynamic quenching efficiencies. J. Phys. Chem.-US *100*:5541-5553.
15. **Torimura, M., S. Kurata, K. Yamada, T. Yokomaku, Y. Kamagata, T. Kanagawa, and R. Kurane.** 2001. Fluorescence-quenching phenomenon by photoinduced electron transfer between a fluorescent dye and a nucleotide base. Anal. Sci. *17*:155-160.
16. **Vieux, E.F., P.Y. Kwok, and R.D. Miller.** 2002. Primer design for PCR and sequencing in high-throughput analysis of SNPs. BioTechniques *32*:S28-32.
17. **Weber, G.** 1952. Polarization of the fluorescence of macromolecules. I. Theory and experimental method. Biochem. J. *51*:145-155.
18. **Werner, M., M. Sych, N. Herbon, T. Illig, I.R. Konig, and M. Wjst.** 2002. Large-scale determination of SNP allele frequencies in DNA pools using MALDI-TOF mass spectrometry. Hum. Mutat. *20*:57-64.
19. **Wolford, J.K., D. Blunt, C. Ballecer, and M. Prochazka.** 2000. High-throughput SNP detection by using DNA pooling and denaturing high performance liquid chromatography (DHPLC). Hum. Genet. *107*:483-487.
20. **Zhou, G., M. Kamahori, K. Okano, G. Chuan, K. Harada, and H. Kambara.** Quantitative detection of single nucleotide polymorphisms for a pooled sample by a bioluminometric assay coupled with modified primer extension reactions (BAMPER). Nucleic Acids Res. *29*:E93.

Received 26 August 2002; accepted 16 October 2002.

Address correspondence to:

Dr. Pui-Yan Kwok
Cardiovascular Research Institute
University of California, San Francisco
505 Parnassus Ave.
Long Hospital 1332A
Box 0130
San Francisco, CA 94143-0130, USA
e-mail: kwok@cvrimail.ucsf.edu