Common human diseases, such as cancer, have been associated with multiple types of variation in the genome, including sequence repeats and deletions and single-nucleotide polymorphisms (SNPs) (1). Of these, SNPs are the most abundant in the human genome (2, 3). As a result of the efforts of many groups (4), an estimated 5 million SNPs are now deposited in public databases (5), providing a resource for determining how genomic variation affects human biology. Building on this work, many groups have shown that drug response is also influenced by genomic variation (6). Multiple SNPs have been identified that have a major impact on response to chemotherapy (7–11); it is therefore necessary to have rapid and efficient SNP evaluation techniques to analyze genes that influence chemotherapy response.

Dihydropyrimidine dehydrogenase (DPD) is the initial and rate-limiting enzyme in the three-step degradation of uracil to β-alanine (12). This is the only endogenous pathway for production of the neurotransmitter β-alanine (13). DPD also degrades >80% of the anticancer agent 5-fluorouracil (5-FU), a pyrimidine analog used to treat colorectal cancer, and limits the oral absorption of the drug (13). DPD activity is found to be highest in liver and mononuclear cells, but it is also present in most other human tissues (14).

Variation in DPD activity can lead to nonbeneficial physiologic conditions. A complete absence of DPD has been associated with the hereditary metabolic disorder thymine-uraciluria (9). This is characterized by mental retardation and is sometimes accompanied by other neurologic disorders, such as microcephaly, motor retardation, and autism (9). In addition, decreased DPD activity leads to severe toxicity from 5-FU even when typical doses are administered. This is often accompanied by severe diarrhea, neutropenia, and sometimes, neurotoxicity (9).

The DPD gene located on chromosome 1p20 encodes DPD (14). At least 13 variant alleles of the DPD gene have been described (15). In particular, three genetic variants have been associated with severe toxicity after 5-FU therapy: DPD*2A, DPD*5, and DPD*6 (9, 12, 13). DPD*2A is the result of a G-to-A transition in a splice donor site that causes skipping of exon 14 (9). DPD*2B includes the exon 14 splice SNP and A1627G in exon 13 (I543V), respectively) at codons 543 and 732, respectively (9). DPD*9A is a point mutation in exon 2 (T to C) with variable effect on DPD activity, but it has a high frequency in the general population. DPD*5 and DPD*6 are point mutations (A-to-G and G-to-A polymorphisms, respectively) at codons 543 and 732, respectively (9).

Use of Pyrosequencing to Detect Clinically Relevant Polymorphisms in Dihydropyrimidine Dehydrogenase, Ranjeet Ahluwalia, Robert Freimuth, Howard L. McLeod, and Sharon Marsh* (Washington University School of Medicine, Division of Oncology, St. Louis, MO 63110; * address correspondence to this author at: Washington University School of Medicine, 660 South Euclid Ave., Campus Box 8069, St. Louis, MO 63110; fax 314-362-3764, e-mail smarsh@im.wustl.edu)
DPYD*2A, DPYD*5, DPYD*6, and DPYD*9A. PCR was performed on DNA from 95 healthy Caucasian volunteers and 95 healthy African volunteers. PCR primers were designed using Primer Express, Ver. 1.5 (ABI). Pyrosequencing primers were designed using Pyrosequencing Primer SNP Design, Ver. 1.01, software (16). PCR was carried out with 10–15 ng of genomic DNA, 0.6 nmol each of the forward and reverse oligonucleotide PCR primers (one of which is biotinylated; Integrated DNA Technologies), and 1× AmpliTaq Gold PCR Master Mix (Applied Biosystems), which contains 255 U (0.05 U/μL) of AmpliTaq Gold DNA polymerase, Gene Amp PCR Gold Buffer (30 mM Tris HCl, 100 mM KCl, pH 8.05), 400 μM deoxynucleotide triphosphates, and 5 mM MgCl₂. The primer details and PCR conditions are described in the Table in the Data Supplement that accompanies the online version of this Technical Brief at http://www.clinchem.org/content/vol49/issue10/. Pyrosequencing was carried out as described previously (16) with internal primer diluted in 1× Annealing Buffer (200 mM Tris-acetate, 50 mM magnesium acetate); 2× Binding Wash Buffer II, pH 7.6 (10 mM Tris-HCl, 2 M NaCl, 1 mM EDTA, 1 mL/L Tween 20); 100 μg of M280 Streptavidin Dynabeads (Dynal AS) per reaction well; 0.5 M NaOH; 1× Annealing Buffer, pH 7.6; and PSQ96 SNP reagents. Samples were analyzed on a PSQ96 instrument with pyrosequencing software (Pyrosequencing AB). A Tecan pipetting robot was used for all steps other than magnetic bead addition and transfer.

The genotyping data from Caucasian and African population sets are shown in Table 1, and pyrogram data from the DPYD*9A analysis are shown in Fig. 1. All results were in Hardy–Weinberg equilibrium (17). For the DPYD*2A SNP, no variant alleles were seen (q = 0) in either the Caucasian or the African populations. DPYD*6 had allele frequencies of 5% and 4% for the Caucasian and African populations, respectively. For the DPYD*5 allele, the observed frequencies in our two populations were 19% for Caucasians and 9% for the African population. In addition, for DPYD*9A, Caucasian samples had an allele frequency of 17%, whereas the African population had an allele frequency of 41%.

This study demonstrates that DPYD*5 occurs at lower frequency in an African population than a Caucasian population and that DPYD*9A has a higher frequency in an African population than a Caucasian population. In addition, in their study of 478 Egyptians, Hamdy et al. (18) reported a frequency of 11.5% for DPYD*5, which is between the frequencies for the Caucasian (19%) and sub-Saharan African populations (9%) found in this study. For DPYD*6, the reported frequency for the Egyptian population (9%) is higher than the frequency found in both the Caucasian (5%) and sub-Saharan African (4%) populations (18). In a study on 100 Japanese individuals, Wei et al. (12) reported frequencies of 35.2% for DPYD*5 and 4.4% for DPYD*6. These interethnic variations in allele frequencies for DPYD*5, DPYD*6, and DPYD*9A suggest the usefulness of documenting the allele frequencies of SNPs in genes that encode drug-metabolizing enzymes for different ethnic populations in the hope of finding therapeutic or diagnostic benefits in specific populations (6).

The DPYD*2A allele was previously found in only 1 of 135 alleles in Caucasians (14) and was not found in the 380 alleles evaluated for DPYD*2A in this study. The previously described frequency for DPYD*6 in Caucasians is similar to that observed in this study (Table 1). The previously reported frequency for DPYD*5 in Caucasians was 28% (14), which is slightly higher than the frequency of 19% found in this study, although within the 95% confidence interval (Table 1). Previously published allele frequencies for the DPYD gene were determined by use of restriction fragment length polymorphism (RFLP) analysis. Overall, pyrosequencing is a more efficient and flexible SNP detection method than RFLP. For RFLP analysis, the appropriate restriction enzyme is needed to cleave the desired DNA fragment at the site of the SNP. Often there are no restriction enzymes available for the sequence containing a SNP, limiting the use of RFLP analysis. Furthermore, the post-PCR steps in pyrosequencing take ~30 min for 96 samples compared with 1–2 h of digestion followed by gel electrophoresis for RFLP analysis. The robotic system (PSQHS 96A) recently introduced by Pyrosequencing AB is three times more sensitive than PSQ96. Reagent costs have been cut by one-third, and sample preparation time substantially reduced (16). In addition, the SNP profile for a single gene or multiple genes in a single patient can be completed in one pyrosequencing assay, making this system appealing as a rapid screening process in clinical laboratories. More than one SNP can be analyzed at the same time (multiplex) by pyrosequencing if the SNPs are within the same PCR

### Table 1. Genotype and allele frequencies for DPYD SNPs in Caucasian and African populations.

| SNP      | American Caucasian | | | | | | African | | | | |
|----------|-------------------|---|---|---|---|---|---|---|---|---|---|---|
|          | Wild type | Heterozygous | Variant | p* | q* | Wild type | Heterozygous | Variant | p* | q* |
| DPYD*2A  | 93         | 93         | 0       | 0       | 1       | 0       | 95         | 95         | 0       | 0       | 1       |
| DPYD*5   | 83         | 56         | 23      | 4       | 0.81 (0.71–0.88) | 0.19 (0.12–0.29) | 94         | 77         | 17      | 0       | 0.91 (0.83–0.95) | 0.09 (0.05–0.17) |
| DPYD*6   | 91         | 82         | 9       | 0       | 0.95 (0.88–0.98) | 0.05 (0.02–0.12) | 84         | 77         | 7       | 0       | 0.96 (0.89–0.99) | 0.04 (0.01–0.11) |
| DPYD*9A  | 95         | 65         | 28      | 2       | 0.83 (0.74–0.89) | 0.17 (0.11–0.26) | 85         | 31         | 38      | 16      | 0.59 (0.48–0.69) | 0.41 (0.31–0.52) |

a Number of passed samples of 95 samples tested.

b Values in parentheses are the 95% confidence intervals.
amplicon. This would allow, for example, DPYD polymorphisms in close proximity to be analyzed in one reaction (e.g., DPYD*4 and DPYD*5).

Haplotype analysis is an area of intense research for complex genetic phenotypes. The knowledge of haplotype for several SNPs in one gene is likely to provide more information about genotype-phenotype links than the underlying SNPs (19, 20). The ability to test for more than

---

Fig. 1. Pyrograms of three individuals from the Caucasian population in the DPYD*9A assay. (A), homozygous wild type (A/A); (B), heterozygous (A/G); (C), homozygous variant (G/G). E, enzyme; S, substrate.
one SNP at a time with the new software (PSQHS 96A) from Pyrosequencing AB is likely to speed up haplotype analysis. When pyrosequencing data for DPYD*5, DPYD*2A, DPYD*9A, and DPYD*6 were analyzed using the Polymorphism Haplotype Analysis Suite (PHASE), Ver. 0.9 (21), almost 90% of all Caucasian samples had one of the following three haplotypes (in the order DPYD*5, DPYD*2A, DPYD*9A, and DPYD*6): AGTG (63%), GGTG (16.5%), and AGCG (14.6%). However, the second most common Caucasian haplotype was seen in <5% of the African samples. This highlights the importance of comparing haplotype composition among different world populations before applying haplotype structure to clinical studies.

In conclusion, the technology from Pyrosequencing AB is a faster and more efficient SNP evaluation method than RFLP analysis. With the reduction in the amount of reagents used and the available software, cost is less of an obstacle for high-throughput SNP analysis in large population studies. This could aid in the rapid and efficient analysis of SNPs in many genes that encode drug-metabolizing enzymes in several patient and population sets in an effort to tailor drug therapy based on the genetic makeup of an individual.

This work was supported by Grant GM63340 from the NIH Pharmacogenetics Research Network (http://pharmacogenetics.wustl.edu).

References