Multiplex Assay for Comprehensive Genotyping of Genes Involved in Drug Metabolism, Excretion, and Transport

THOMAS M. DALY,1 CARMEN M. DUMAUAL,1 XIN MIAO,2 MARK W. FARMEN,1 REUBEN K. NJAU,1 DONG-JING FU,1† NANCY L. BAUER,1 SANDRA CLOSE,1 NANCY WATANABE,2‡ CARSTEN BRUCKNER,2 PAUL HARDENBOL,2 and RICHARD D. HOCKETT1*  

Background: Drug metabolism is a multistep process by which the body disposes of xenobiotic agents such as therapeutic drugs. Genetic variation in the enzymes involved in this process can lead to variability in a patient’s response to medication.

Methods: We used molecular-inversion probe technology to develop a multiplex genotyping assay that can simultaneously test for 1227 genetic variants in 169 genes involved in drug metabolism, excretion, and transport. Within this larger set of variants, we performed analytical validation of a clinically defined core set of 165 variants in 27 genes to assess accuracy, imprecision, and dynamic range.

Results: In a test set of 91 samples, genotyping accuracy for the core set probes was 99.8% for called genotypes, with a 1.2% no-call (NC) rate. The majority of the core set probes (133 of 165) had ≤1 genotyping failure in the test set; a subset of 12 probes was responsible for the majority of failures (mainly NC). Genotyping results were reproducible upon repeat testing with overall within- and between-run variation of 1.1% and 1.4%, respectively—again, primarily NCs in a subset of probes. The assay showed stable genotyping results over a 6-fold range of input DNA.

Conclusions: This assay generates a comprehensive assessment of a patient’s metabolic genotype and is a tool that can provide a more thorough understanding of patient-to-patient variability in pharmacokinetic responses to drugs.

© 2007 American Association for Clinical Chemistry

The pharmacokinetic profile of a drug depends on its absorption, distribution, metabolism, and excretion. These processes require interaction of the drug compound with multiple host proteins. Initial drug absorption and distribution occurs via a diverse family of transporters, depending largely on the molecular structure of the parent molecule (1, 2). After uptake of the drug, metabolism can occur in two distinct phases: enzymatic oxidation, reduction, and hydrolysis reactions that expose or add functional groups to produce polar molecules (phase I metabolism) (3), followed by addition of endogenous compounds to the molecules to further increase solubility (phase II metabolism) (4). The ultimate pharmacokinetics of a drug is largely determined by the relative function of these processes, which can be influenced by both genetic and environmental variables.

Significant interindividual variability exists in drug disposition and response (5). Much of the observed heterogeneity is thought to result from underlying genetic variation in the human population (6–9). Genetic variants such as single nucleotide polymorphisms (SNPs)3, insertions/deletions (indels), and gene duplications can all potentially affect either the expression level of an enzyme or the functional activity of the gene product, resulting in altered pharmacokinetics. These effects have been most clearly demonstrated for members of the cytochrome P450 (CYP) system, such as CYP2D6, CYP2C9, and

1 Eli Lilly and Company, Lilly Corporate Center, Indianapolis IN.
2 Affymetrix Inc., Santa Clara, CA.
† Current affiliation: Johnson & Johnson Pharmaceutical Research and Development, L.L.C., Raritan, NJ.
‡ Current affiliation: Genentech, Inc., San Francisco, CA.
* Address correspondence to this author at: Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN, 46285. Fax 317-276-5281; e-mail hockettrd@lilly.com.

Received January 25, 2007; accepted April 18, 2007. 
Previously published online at DOI: 10.1373/clinchem.2007.086348

3 Nonstandard abbreviations: SNP, single nucleotide polymorphism; indel, insertion/deletion; CYP, cytochrome P450; MIP, molecular inversion probe; MAF, minor allele frequency; NC, no call; PRA, possible rare allele; DMET, Drug Metabolizing Enzyme and Transporter.
CYP2C19, where patients can be readily classified into metabolic phenotypes based on genotype (10–13). These classifications have been demonstrated to have direct effects on drug treatments for various compounds (14–16). Similar information is known for some non-CYP enzymes such as TPMT (17) and UGT1A1 (18, 19), but the clinical relevance of genetic variations for many enzymes is less understood, as are the effects of multigenic variations on overall drug distribution. In part, this is a result of the lack of readily available genotyping assays for the majority of metabolic enzymes and the expense required to broadly genotype across multiple genes. A multiplex assay allowing comprehensive analysis of a patient’s metabolic genotype would be a useful tool for drug development and discovery.

Molecular inversion probe (MIP) technology is an oligonucleotide-based method that can be used to analyze several thousand SNPs in a single assay (20). This technique has been used extensively for the HapMap project (21) and offers several advantages for multiplex genotyping. First, PCR amplification occurs after mutation detection, at which time MIPs have been converted to standard-length oligonucleotides of similar sequence composition utilizing common PCR primers. Second, the MIPs are designed to be locus-specific rather than allele-specific and require only a single probe per marker, so any loss of performance of a probe will affect both alleles equally, avoiding unbalanced genotyping. Third, each probe has two recognition sequences as in PCR but retains unimolecular hybridization kinetics, enabling high multiplexing with low concentration requirements. Finally, the use of molecular tag sequences on the probes allows hybridization to the microarray to be independent of the gene sequence, increasing the specificity and sensitivity of detection and allowing for a high degree of flexibility in assay design (22).

We have applied MIP technology to develop an assay testing 1227 variants in 169 genes involved in drug metabolism. Within this set, we have chosen a core set of 169 variants in 27 genes for extensive analytical validation. This core set consists of genes for which substantial published evidence suggests functional significance of the mutations, as well as the “known valid biomarkers” listed in a recent US Food and Drug Administration guidance document (23). These core set assays were then characterized for accuracy, imprecision, and dynamic range.

Materials and Methods

Selection of Genetic Variants and Assay Design
We compiled an initial list of 169 genes involved in drug transport and phase I and II metabolism from literature sources and consultation with experts. We identified a set of 1227 variants within these genes for inclusion in the MIP pool using published literature and databases such as dbSNP (24), Ensembl (25), and the Karolinska CYP450 website (26). We chose this extended set based primarily on sequence data without considering allele frequency or clinical effects, and included exonic, intronic, and promoter variants (see Table 2 in the Data Supplement that accompanies the online version of this article at http://www.clinchem.org/content/vol53/issue7). Within this extended set, we selected 165 variants in 27 genes as a core set of variants for clinical validation. The core set includes genes and variants for which literature evidence suggests functional significance of the mutations, as well as all reported variants of CYP2D6, CYP2C9, and CYP2C19 (Table 1).

We designed MIP probes for all variants using sequences corresponding to build 35 of the human genome (UC Santa Cruz Golden Path assembly). We selected flanking sequences using proprietary software (Affymetrix SNP Probe Designer version 06.08.21.1) and assembled probe precursor oligonucleotides into final probes and probe pools. We used GeneChip Universal 3K Tag Arrays (Affymetrix) for hybridization. Each locus is measured in 3 well-separated features on the microarray.

Table 1. Variants included in the core set of genes. *A

<table>
<thead>
<tr>
<th>CYP450</th>
<th>No. variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP1A2</td>
<td>6</td>
</tr>
<tr>
<td>CYP2A6</td>
<td>7</td>
</tr>
<tr>
<td>CYP2B6</td>
<td>10</td>
</tr>
<tr>
<td>CYP2C19</td>
<td>12</td>
</tr>
<tr>
<td>CYP2C8</td>
<td>4</td>
</tr>
<tr>
<td>CYP2C9</td>
<td>10</td>
</tr>
<tr>
<td>CYP2D6</td>
<td>18</td>
</tr>
<tr>
<td>CYP2E1</td>
<td>1</td>
</tr>
<tr>
<td>CYP2J2</td>
<td>5</td>
</tr>
<tr>
<td>CYP3A4</td>
<td>2</td>
</tr>
<tr>
<td>CYP3A5</td>
<td>8</td>
</tr>
</tbody>
</table>

Non-CYP enzymes

| CDA                 | 2            |
| DPYD                | 6            |
| FM02                | 1            |
| FM03                | 2            |
| GSTP1               | 2            |
| NAT1                | 12           |
| NAT2                | 7            |
| TPMT                | 5            |
| UGT1A1              | 9            |

Transporters

| ABCB1               | 6            |
| ABCBC1              | 6            |
| SLC01B1             | 9            |
| SLC02B1             | 1            |
| SLC15A2             | 4            |
| SLC22A1             | 6            |
| SLC22A2             | 4            |

*Variants tested include SNPs, insertions, deletions, triallelic variants, and nucleotide repeats. Complete details of the variants tested, as well as a full list of probes and genes included in the extended set of probes, are available in Supplementary Table 1 of the online Data Supplement.
providing internal redundancy to reduce the chance of genotyping failure resulting from localized errors such as debris on the microarray. A final genotype call is determined by a majority-rules voting algorithm of the 3 features.

We initially synthesized 641 probe designs for the core set variants. This initial pool was screened on a subset of genomic DNA samples containing parent-child trios. We selected a single best probe design for each locus based on visual inspection of genotype cluster separation for each variant, as well as assay parameters such as call rate, repeatability, and trio accuracy. We then manufactured a final probe pool containing a single probe for each variant. (To fully interrogate the locus, indel variants require 2 probes, which are synthesized into a final genotype call. For this report, these are considered as a single probe per locus.)

**DETERMINATION OF CLUSTER BOUNDARIES AND CALL DEFINITIONS FOR FINAL ASSAY**

We ran the final probe pool against a training set of 529 ethnically diverse population-based DNA samples obtained from the index repository at Cogenics. Results were used to set initial genotyping cluster boundaries for each probe using predefined software algorithms (27). For the core set loci, any samples falling near software-defined boundaries were sequenced to confirm the appropriateness of boundary cuts, and some boundaries were manually adjusted before being fixed as the final cluster boundaries for validation experiments. A genotype had to be present in at least 4 samples in the training set to adequately define a cluster. Based on this cutoff, 52 core set probes had clusters defined for all 3 genotypes, whereas 41 probes had clusters defined for only the wild-type and heterozygous genotypes. Minor allele clusters were not defined for the remaining 72 probes, indicating that they had a minor allele frequency (MAF) of <0.5% in the training set samples.

The assay generates 2 primary classes of calls, genotype calls (for samples that fall within predefined cluster boundaries) or no call (NC) (for samples that fall outside cluster boundaries or fail other quality control tests). However, for probes that do not have 3 defined clusters owing to a low MAF in the training set, any samples that truly contained a rare allele would generate a NC result and be missed. To address this issue, we created a possible rare allele (PRA) call category defined as the subset of NC results occurring in such probes. We evaluated whether the PRA designation could serve as a screening test to identify samples containing previously undefined rare genotypes.

**ANALYTICAL PROCEDURE**

The Drug Metabolizing Enzyme and Transporter (DMET) assay used in this report is based on a modification of the previously described Affymetrix Targeted Genotyping assay (20). DNA samples were initially diluted to 150 ng/μL in Tris-EDTA buffer, and we used an initial PCR amplification step to amplify 32 loci that either had a pseudogene or did not generate sufficient signal using the routine Targeted Genotyping protocol. These preamplified products were then serially diluted and divided into aliquots, and an aliquot was combined with 13.4 μL genomic DNA. We then incubated this mixture with a multiplex PCR anneal cocktail containing the custom Targeted Human DMET assay 1.0 probe panel (Affymetrix). Samples were incubated on the 9700 Thermal Cycler for 4 min at 20 °C, 5 min at 95 °C, and then overnight at 58 °C. All remaining steps were carried out using Affymetrix Targeted Human DMET assays.

Microarrays were washed and scanned with 4-color detection using the Affymetrix GeneChip Scanner 3000 7G 4C. Raw signal values were background subtracted and normalized and genotypes were reported using the Affymetrix GeneChip Operating Software version 1.4 and Affymetrix Targeted Human DMET 1.0 software. Individual samples were assigned a status of “pass” based on routine chip performance metrics. We then performed single-sample genotyping by comparing each individual marker’s data to the specific, predefined cluster boundaries.

**ASSAY VALIDATION STUDIES**

**Accuracy.** We evaluated genotyping accuracy using a test set of 91 samples that had been sequenced at the 165 core set loci. This test set contained genomic DNA from 74 Epstein–Barr virus-transformed cell lines (Cogenics) as well as 17 samples of DNA isolated from normal donor peripheral blood mononuclear cells. We compared results from the DMET 1.0 assay with sequencing results for each of the core set loci. We sent samples with discrepant results to a 3rd party for tiebreaking via bidirectional sequencing using 2 primer sets different from those originally used. Fig. 1A outlines the process; final test set genotypes are included (see Table 2 in the online Data Supplement).

**Imprecision.** We used 5 genomic samples to evaluate the reproducibility of the assay for signal intensity, contrast, and genotype call. We tested 8 aliquots of each sample in a single run to evaluate within-day imprecision. Between-day imprecision was assessed over 10 individual runs using a single aliquot of each sample per run. We calculated variance on a signal level as percentage CV for intensity (y-axis) and SD for contrast (x-axis) for each of the core set probes. In addition, we calculated the reproducibility of genotype calls across the repeated runs for each of the 5 control samples.

**Dynamic range.** We performed a dilution series to determine the robustness of the assay to changes in input DNA. Four genomic samples were diluted from 0.25 to 8.0 μg, and each dilution was tested in duplicate in a single
run. We calculated genotype accuracy for each dilution by comparing the results at a given dilution to the results at 2 μg (the recommended starting amount).

CONSTRUCTION AND ANALYSIS OF PLASMIDS CONTAINING MINOR ALLELES
Because many of the variants included on the chip are very rare, not all possible genotypes are represented in the test set samples. To determine whether the assay was technically capable of differentiating these rare variants, we developed a series of plasmid controls containing major and minor alleles for the core set loci. Plasmids were constructed by Blue Heron Biotechnology, using genome synthesis techniques, and contained the complement of the probe homology regions with an additional 20 bases of genomic sequence 5’ and 3’ of the probe binding sequence. Sequence inserts were cloned into a pUC19 vector and transformed into Escherichia coli for propagation and amplification. We confirmed sequence accuracy of the cloned products by resequencing the final plasmid product. These plasmids were then genotyped using the DMET assay either individually or in 1:1 mixtures to represent homozygous major, heterozygous, and homozygous rare genotypes for all core set loci.

RESULTS
We set initial assay definitions using a 529-sample training set as described in Materials and Methods and then assessed genotyping accuracy on a separate test set of 91 samples using the protocol diagrammed in Fig. 1A. The genotyping success rate of the DMET 1.0 assay for the core set probes was 98.5%, with the majority of failures resulting from lack of a call (NC/PRA) rather than incorrect genotype (Fig. 1B). This resulted in 99.8% accuracy for called genotypes compared to sequencing. Accuracy rates were similar for wild-type and heterozygous samples, although the miscall rates were slightly higher in samples with rare homozygous variants and largely driven by a single probe (CYP1A2star1K-164CA) that accounted for 4 of the 7 miscalled rare homozygous samples.

Genotyping accuracy varied across the core set probes. The vast majority of probes were highly accurate, with 135 of the 165 core set loci showing no more than a single genotyping failure in the 91 test set samples (Fig. 2A). In contrast, 12 probes accounted for more than half of the total genotyping failures, the majority being NC/PRA. This was not unexpected, as certain probes have poor cluster separation resulting in a relatively high NC rate

---

Fig. 1. DMET chip genotyping accuracy. (A), genotypes generated from the DMET chip were compared to results from sequencing using this algorithm. “Miscalled” results are samples where DMET results did not agree with sequencing. “Inconsistent” results are samples where the 3rd-party tiebreaker sequencing produced results different from both the DMET and initial sequencing. The 32 samples in the hatched boxes are all counted as errors in the final accuracy calculations. (B), rates were calculated as the relevant number of events divided by the total number of genotyping opportunities. The majority of genotyping failures were NC/PRA results. Rates were comparable across the 3 genotypes. BB, homozygous major; Bb, heterozygous; bb, homozygous minor.

---

*Does not include triallelic loci
Fig. 2. Individual probe and sample genotyping performance.

(A), genotyping success rates were calculated on an individual probe basis for each of the core set variants. MM, miscal. The probe number on the x-axis is assigned solely for graphing purposes to group errors together and does not reflect arrangement of probes on microarray. (B), representative cluster plots for 2 probes. Vertical colored lines represent the predefined cluster boundaries used by the software to call genotypes. Upper panel, a probe with good cluster separation and high success rate. Lower panel, probe with the highest NC rate, caused by a significant overlap between 2 of the clusters. (C), genotyping accuracy was calculated on an individual sample basis for the 91 test set samples. The majority of samples showed >99% accuracy, with a subset of samples that had increased NC/PRA rates. The sample number on the x-axis is assigned solely for graphing purposes to group errors together and does not reflect processing order of samples.
(Fig. 2B). Accuracy on a per-sample basis was very high, with a median accuracy of 99.4% (range 89.8%–100%) for the test set samples (Fig. 2C). A subset of 12 samples accounted for more than half of the total genotyping failures. Interestingly, many of these samples showed concurrent increases in both NC and PRA calls, suggesting that there may be a matrix component in these samples contributing to a generalized increase in individual call failures. It is important to note, however, that even in samples with high NC/PRA rates, the remaining probes in those samples produced accurate genotypes.

The use of the PRA designation to identify samples containing extremely uncommon variants yielded mixed success. Fifteen of 18 rare alleles in the test set were correctly identified as PRAs by the assay, giving a sensitivity of 83.3% for the PRA designation. However, 76 alleles were incorrectly identified as PRAs, resulting in a positive predictive value of only 16.5%. This suggests that the PRA designation is best used as a screening test and requires confirmation by alternative methods.

**Imprecision**

We assessed the reproducibility of genotyping results by analyzing repeat runs of 5 genomic samples. We assessed variability for the core set probes by evaluating temporal changes in contrast ($x$-axis) and intensity ($y$-axis) values for each probe. Median intensity CVs ranged from 17% to 26% within-day and 25% to 34% between-day. Variability was somewhat correlated with signal intensity (Fig. 3A), with higher CVs seen for probes with lower intensities. Contrast variability was extremely low for the core set probes, with median SDs from 0.027 to 0.034 within-day and 0.031 to 0.040 between-day for the 5 controls (Fig. 3B). Ninety-five percent of probes showed <0.1 SD contrast variation between runs. The low contrast variance is critically important for genotyping purposes, as the cluster boundaries are set on the contrast axis, and is likely the result of the much greater correlation of the signal ratio of the 2 signals composing a call across the repeated samples than the absolute signal across the samples. This is reflected in the genotype calls for the 5 genomic controls (Fig. 3C), with reproducibility rates of 98.9% and 98.6% for within- and between-day runs, respectively.

**Dynamic Range**

Performance was relatively stable between 1 and 6 µg input DNA, with an increased error rate beyond this range (Fig. 4). The main type of error observed was an increase in NC/PRA calls, with only 2 instances of miscalled genotypes.

**Plasmid Analysis**

There was clear separation of genotypes for 137 of the 163 nontriallelic loci, with most probes falling near the idealized contrast values of −1, 0, and +1 for homozygous, heterozygous, and homozygous alleles, respectively (Fig. 5A). This separation included probes where not all clusters were defined in the training set (Fig. 5B), indicating that the probes for these loci are functional for the rare allele.

**Discussion**

Drug metabolism is a complex process involving the interaction of multiple genetic and environmental factors. We have developed a tool that will be useful to help understand the genetic factors that can influence this process. By simultaneously testing across multiple genes, this process can generate a more comprehensive description of a patient’s metabolic genotype than has previously been possible in a single assay. Applied to metabolism in this study, the approach could conceivably be applied to other complex genetic interactions as well. We believe a tool of this nature will be critical for understanding the complex multi-gene interactions underlying drug metabolism.

In addition to developing an assay suitable for research use, we also wanted to develop a tool that could generate validated results suitable for clinical trial support. To this end, we defined a core set of 165 genetic variants in 27 genes known to be relevant for drug metabolism and performed detailed validation on these probes. The analytical performance of this assay was robust, with 99.8% accuracy for called genotypes compared to the gold standard of sequencing. This may be a conservative estimate, as all discrepancies with sequencing were counted as errors, which does not account for possible sequencing errors. The most common cause of assay failure was the lack of a genotype call (NC/PRA), rather than a miscalled genotype. This is a critical distinction for clinical use, as an absent data point is less problematic than assigning an incorrect genotype to a patient. A relatively small number of probes were responsible for the majority of call failures. Further work to optimize these probes for future versions of the DMET assay is underway. Alternatively, as haplotype information on these enzymes becomes more widely available, it may be possible to eliminate some poorly performing probes that are adequately represented by neighboring variants in the gene. Imprecision of the assay was also very good for an assay of this complexity, as reflected in the high level of genotyping consistency across multiple runs.

A potential strength of a microarray-based approach is that rare variants can be readily included in the testing process with no appreciable increase in assay complexity or cost. The assay correctly identified 88 of 89 variants in the test set with a MAF of 1% to 5% and 27 of 31 variants with a MAF of <1%. The use of the PRA designation was helpful in this regard, but had a high false-positive rate and therefore is more useful as a screening test than as a definitive call. We are currently evaluating alternative software clustering algorithms that may improve the ability to call PRAs. In addition, ongoing effort to identify samples containing rare variants will allow further definition of rare clusters for many probes, improving the
Fig. 3. Imprecision analysis of core set probes. (A), between-day imprecision, intensity. (B), between-day imprecision, contrast. Homozygous samples (located at \(-1\) and \(+1\)) showed less variation than heterozygous samples (centered on 0). (C), genotype reproducibility. Columns and error bars represent mean and SD. Genotyping failures (y-axis) include both miscalled genotypes and NC/PRAs.
ability of the assay to detect rare alleles and reducing the overall NC/PRA rate.

Despite the wide range of genotyping done by this assay, in its current format allele quantification is not possible, so large-scale deletions or duplications [such as CYP2D6*5 (28) or *1XN (29)] cannot be readily detected, although homozygous deletions can be suspected by low signals across multiple probes for a gene. Also, enumerating small tandem repeats [such as TA repeats in UGT1A1*28 (30)] is difficult with the current format of the assay, limiting its ability to differentiate variants of this type. Further modifications of software algorithms and probe designs to enable detection of such variants will be required to make this a truly comprehensive assay. Whereas the analytical validation of this assay has focused on the core set of 165 variants in 27 key genes, the performance of the extended set of variants included on the DMET 1.0 assay has not been addressed. Although these probes were derived using the same algorithms as for the core set probes, the final accuracy and precision of the extended set has not been determined. However, the validation protocol shown here can be readily applied to additional genes on the chip as desired.

In conclusion, we have developed a single-microarray assay that allows for the comprehensive genetic analysis of genes involved in drug metabolism, transport, and excretion. The multiplex nature of this assay provides an advantage in that cost does not scale with increases in the number of variants tested, allowing large-scale genotyping at a reasonable cost. Such an approach may be useful to help understand complex multigenic interactions that impact pharmacokinetics beyond the more simple mono-

genic models currently described. Application of this tool early in drug development may help to identify patients at risk for undesirable drug reactions and provide a method to allow better tailoring of drug regimens for individual patients.

Grant/funding support: None declared.
Financial disclosures: Several authors (as noted on the title page) are employees of Affymetrix, the company that manufactures and sells the DMET assay. In addition, many authors are employees and shareholders of Eli Lilly and Company.

Acknowledgements: We thank David Flockhart, Kate Hillgren, Richard Kim, Jeff Miller, Richard Weinshilboum, and Steve Wrighton for their assistance in defining the initial gene list; Genaissance (now CogenicsTM, a division of Clin-
ical Data Inc.) for providing the training and test set samples; and Sharie Sipowicz for superior editorial help preparing the manuscript.

References