Characterizing the pharmacogenome using molecular inversion probes for targeted next-generation sequencing

Oscar Suzuki^{‡,1,2}, Olivia M Dong^{‡,1,2}, Rachel M Howard^{1,2} & Tim Wiltshire^{*,1,2,3}

¹Division of Pharmacotherapy & Experimental Therapeutics, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

²Center for Pharmacogenomics & Individualized Therapy, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

³Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

*Author for correspondence: Tel.: +1 919 843 5820; Fax: +1 919 966 5863; timw@unc.edu

‡ Authors contributed equally

Aim: This study assesses the technical performance and cost of a targeted next-generation sequencing (NGS) multigene pharmacogenetic (PGx) test. **Materials & methods:** A genetic test was developed for 21 PGx genes using molecular inversion probes to generate library fragments for NGS. Performance of this test was assessed using 53 unique reference control cell lines from the Genetic Testing Reference Materials Coordination Program (GeT-RM). **Results:** 93.7% of variants were successfully called and the repeatability rate was 99.9%. Reference calls were available for 78.4% of diplotype calls resulting from PGx testing, and concordance for the test was 85.7%. Cost per sample was \$32–\$56. **Conclusion:** A targeted NGS assay using molecular inversion probe technology is able to characterize the pharmacogenome efficiently.

First draft submitted: 16 April 2019; Accepted for publication: 24 July 2019; Published online: 27 September 2019

Keywords: molecular inversion probes • next-generation sequencing • pharmacogenetic • pharmacogenome • precision medicine

The pharmacogenome is comprised of pharmacogenetic (PGx) variants that are associated with an increased risk of adverse outcomes or decreased drug efficacy if standard drug dosing recommendations are followed [1]. A subset of these PGx variants is considered actionable because they have guidelines from the US FDA or the Clinical Pharmacogenetics Implementation Consortium (CPIC) that result in a change in drug dosing or medication selection from standard recommendations [2,3]. The FDA has approved of more than 1200 medications and approximately 7% are associated with actionable PGx germline variants in 17 genes [4]. Continued efforts that are aimed to help successful implementation of PGx testing are needed since only 7% of US hospitals offer PGx testing [5].

Clinical implementation of PGx testing has traditionally been reactive and limited to the immediate drug being prescribed, which typically involves genotype information for a single PGx gene [1]. This reactive method of implementing PGx testing has contributed to the low adoption of PGx testing since clinicians may not have the 0.3–16 days it requires to receive PGx results when making drug-prescribing decisions [6]. Clinical implementation of PGx information in the drug prescribing process may be more successful if the testing approach was pre-emptive and provided comprehensive information about the pharmacogenome, such that information was readily available at the time of drug prescribing. Pre-emptive, multigene PGx testing provides clinicians with comprehensive PGx information for drugs across multiple therapeutic areas and allows for immediate use of the PGx information when it becomes relevant in the drug prescribing process. Typically, multigene PGx testing has relied on genotyping technologies to provide data, but given the rapid decrease in sequencing cost in recent years, using next generation sequencing (NGS) is an affordable implementation strategy that can be used instead [7]. Sequencing provides a more detailed analysis of the gene being interrogated, and enables true precision therapy by comprehensively testing for all alleles of the gene, even less frequent ones. However, further research is required to fully understand the benefits of providing patients with pre-emptive, multigene PGx testing.

When implementing a PGx test for clinical or research use, there are important factors of a PGx test to consider. Assay characteristics, such as performance adequacy, genetic markers captured, flexibility in adding new genetic markers, sample processing time to receive genetic results and cost are important to evaluate. There are different approaches to capturing targeted genomic regions for sequencing. In this study, a multigene PGx test that uses molecular inversion probes (MIPs) for targeted NGS to interrogate the pharmacogenome will be explored. Once optimized, MIPs have been found to be an efficient method for large-scale SNP analysis (>1000 SNPs) with high call rate (95%), accuracy (>99%), concordance with the gold standard sequencing method of Sanger sequencing (>99.4%) and reproducibility (99.9%) [8]. This large-scale interrogation of the genome is also cost-efficient (reagent costs <\$1/gene/sample), making it a viable approach to characterize the dynamic and expanding pharmacogenome [9]. NGS enables the capture of all variant data for a locus and therefore can help ensure haplotype assignment with greater accuracy, as genotyping arrays rarely provide a comprehensive catalog of variants [10]. In addition, in comparison to genotyping arrays, this method allows the interrogated genomic regions to be revised with more flexibility as new PGx markers are discovered. Using NGS technologies can also be more time and cost-efficient than genotyping arrays because they allow for multiplexing and parallel sequencing of samples.

Previous studies using MIPs for targeted NGS focused on discovery of rare and novel variants given its ability to interrogate extensive genomics regions for large numbers of samples at an affordable cost [8,9,11]. One study used a combination of two NGS-based approaches, MIPs and a hybridization-based capture, to characterize the pharmacogenome with a focus on discovering rare and novel variants, and found NGS to be a suitable method to comprehensively profile the pharmacogenome [11]. However, the use of MIPs for targeted NGS has not been previously used to comprehensively characterize and focus on established genetic variants of the pharmacogenome, especially variants that are actionable per FDA or CPIC guidelines. Using MIPs for targeted NGS as the technological approach of a multigene PGx test has potential application for research and clinical use. To this end, the primary objective of this study is to demonstrate the technical performance and cost of a multigene PGx test, which we designate as DNA2Rx™, which uses MIPs for targeted NGS to characterize the pharmacogenome.

Methods

Selection of genes & genetic variants

CPIC assigns an evidence level (A–D) for PGx testing to each gene–drug pair based on evidence strength, and this classification is subject to change as new research findings become available. Only those gene–drug pairs that have had sufficient in-depth review of evidence to provide definitive CPIC level assignments are designated with guidelines. DNA2Rx includes genetic markers for 21 genes, and 15 genes have CPIC guidelines (Table 1): *ABCB1*, *BCHE*, *CYP2B6*, *CYP2C19*, *CYP2C9*, *CYP2D6*, *CYP3A4*, *CYP3A5*, *CYP4F2*, *DPYD*, *G6PD*, *IFNL3*, *ITPA*, *NAT2*, *NUDT15*, *RYR1*, *SLCO1B1*, *TPMT*, *TYMS*, *UGT1A1* and *VKORC1*. Table 2 lists the associated drugs that are prescribed with information for these PGx genes.

Performance assessment of the MIPs PGx Panel

The performance of DNA2Rx was assessed using 53 reference and quality control cell lines (Supplementary Table 1) from the Genetic Testing Reference Materials Coordination Program (GeT-RM) that have characterized PGx genes and determined consensus genotypes using genetic information results from different assays [13,14]. The PGx genes that overlap with DNA2Rx in these reference materials include *ABCB1*, *CYP2B6*, *CYP2C19*, *CYP2C9*, *CYP2D6*, *CYP3A4*, *CYP3A5*, *CYP4F2*, *DPYD*, *G6PD*, *IFNL3*, *ITPA*, *NAT2*, *SLCO1B1*, *TPMT*, *UGT1A1* and *VKORC1*. Genes on DNA2Rx that are not characterized by the GeT-RM include: *BCHE*, *NUDT15*, *RYR1* and *TYMS*. Fifteen cell lines that were used in the validation process have sequence data available from the 1000 Genomes Project [15], and we compared calls for the specific SNPs of interest for these four genes with available sequence data with 100% concordance of results. Cell lines were obtained from the National Institute of General Medical Sciences (NIGMS) Human Genetic Cell Repository at the Coriell Institute for Medical Research (NJ, USA). The 53 selected samples were chosen to obtain a diverse set of genotype call results and was based on being able to assess SNPs that had previously been called in the ‘gold-standard’ GET-RM dataset. The minimum number of cell lines was selected where heterozygous calls have been assessed for all of the genes and SNPs that we wanted to validate using a MIP technology. The primary consideration being able to call both alleles of a SNP, and this set of 53 cell lines met those requirements. 16 samples were sequenced twice to assess reproducibility. DNA was extracted from cell lines using the Blood and Cell Culture DNA Mini Kit (Qiagen, Hilden, Germany) per protocols from Qiagen.

Table 1. Genetic markers included on the pharmacogenetic assay.

Gene	Default haplotype [†]	Haplotype or single nucleotide polymorphism tested
1. <i>ABCB1</i>	Wild-type	rs1128503, rs2032582, rs1045642
2. <i>BCE</i>	Wild-type	rs1799807, rs28933389, rs28933390
3. <i>CYP2B6</i>	*1	*2, *4, *5, *6, *6A, *7, *11A, *16, *18, *19, *20, *22, *26, *27, *28, *33, *34, *35, *36
4. <i>CYP2C19</i>	*1	*2, *3, *4A, *4B, *5, *6, *7, *8, *9, *10, *11, *14, *15, *16, *17, *18, *19, *22, *24, *25, *26, *27, *29, *33, *34, *35
5. <i>CYP2C9</i>	*1	*2, *3, *4, *5, *6, *8, *9, *10, *11, *12, *13, *14, *15, *24, *25, *27, *31, *33, *35
6. <i>CYP2D6</i>	*1A	*2, *3, *4, *6, *7, *8, *9, *10, *11, *12, *14, *15, *17, *18, *19, *20, *21, *29, *30, *34, *38, *39, *40, *41, *42, *44, *45, *47, *49, *50, *51, *53, *56, *58, *64, *65, *69, *70, *72, *81, *82, *83, *91, *96, *100, *107, *109
7. <i>CYP3A4</i>	*1	*4, *6, *7, *8, *9, *10, *11, *12, *13, *15, *16, *17, *18, *22, *26
8. <i>CYP3A5</i>	*1	*2, *3, *4, *5, *6, *7, *8, *9
9. <i>CYP4F2</i>	*1	*3
10. <i>DPYD</i>	*1	*2A, *2B, *3, *5, *7, *9A, *9B, *12, *13
11. <i>G6PD</i>	B	A, A- 202A_376G, A- 968C_376G, Mediterranean haplotype, Asahi, Hechi, Sierra Leone, Ananindeua, Miaoli, Vangchan; Jammu, Kalyan-Kerala; Jamnaga; Rohini, Mira d'Aire, Bangkok Noi, Cosenza, Kaiping; Anant; Dhoni; Sapporo-like; Wosera
12. <i>IFNL3</i>	Wild-type	rs12979860
13. <i>ITPA</i>	Wild-type	rs1127354
14. <i>NAT2</i>	*4	*5, *6, *7, *11, *12, *13, *14, *17, *19, *22
15. <i>NUDT15</i>	*1	*2, *3, *4, *5, *6
16. <i>RYR1</i>	Wild-type	rs121918592, rs121918592, rs118192177, rs118192177, rs193922807, rs121918593, rs193922809, rs28933396, rs28933396, rs193922762, rs193922803
17. <i>SLCO1B1</i>	*1A	*1B, *2, *3, *4, *5, *6, *8, *9, *10, *11, *12, *13, *14, *15, *16, *17, *18, *19, *20, *21, *23, *29, *31, *34
18. <i>TPMT</i>	*1	*2, *3A, *3B, *3C, *4, *7, *8, *9, *10, *14, *16, *20, *22, *25, *29
19. <i>TYMS</i>	Wild-type	rs2853539
20. <i>UGT1A1</i>	*1	*6, *27, *28, *36, *37, *80, *80+*28, *80+*37
21. <i>VKORC1</i>	H1	H2, H3, H4, H6, H7, H9

[†] Designated haplotype when all other tested haplotypes are absent.

DNA concentration was quantified with the QuantiFluor[®] ONE dsDNA System (Promega, WI, USA) using the Infinite 200 plate reader (Tecan, Mannedorf, Switzerland).

Molecular inversion probe design

DNA2Rx uses a custom-capture method using MIPs to generate a short DNA fragment library of the pharmacogenome for NGS. MIPs were designed using the MIPgen program [16] to capture the regions of each gene where the selected variants are located. Each MIP was designed to comprise of 75 oligonucleotides: a five nucleotide molecular tag (randomized nucleotides added during the probe synthesis), a common 30-nucleotide linker sequence flanked by a 16- to 20-nucleotide extension arm and a 20- to 24-nucleotide ligation arm (extension and ligation arms vary in length and add up to 40 nucleotides) (Figure 1). The ligation and extension arms were designed to target and capture specific 112-bp regions of interest in a gap-fill and circularization reaction. MIP probes were synthesized by Integrated DNA Technologies (IDT, IA, USA) using the Ultramer[®] method.

Library generation for next-generation sequencing

Preparation of MIP libraries (Figure 2) followed a standard protocol previously published [9]. Briefly, pooled phosphorylated MIPs are hybridized to sample genomic DNA, for the capture reaction, which consists of a gap-filling and a ligation reaction that circularizes the MIPs and their targets. Remaining linear DNA is then digested in an exonuclease reaction. The circular MIPs, with captured target DNA, undergo a postcapture PCR amplification with primers that anneal to the common linker. This reaction appends sequencing adaptors and sample-specific eight-nucleotide index sequences (barcodes) to the captured DNA. Amplified libraries from multiple samples are then pooled together and purified. The sample-specific barcode is a unique identifier that allows libraries from different individuals to be included on the same sequencing run (multiplexed) and then have their sequences sorted according to the read barcode (demultiplexed). The MIP capture reaction was optimized for successful target

Table 2. DNA2Rx™ gene list and associated drugs with Clinical Pharmacogenetics Implementation Consortium or the US FDA Guidance.

Gene	CPIC guidelines (any rating level) [3]	CPIC rating level A or B (no CPIC guidelines), unless specified [3]	FDA Guidance [2] [†]
1. <i>ABCB1</i>	None	Antidepressants	None
2. <i>BCHE</i>	None	None	Succinylcholine
3. <i>CYP2B6</i>	None	Efavirenz, methadone, nevirapine	Efavirenz, prasugrel
4. <i>CYP2C19</i>	Amitriptyline, citalopram, clomipramine, clopidogrel, doxepin, escitalopram, imipramine, sertraline, trimipramine, voriconazole	Dexlansoprazole, esomeprazole, lansoprazole, omeprazole, pantoprazole, rabeprazole	Brivaracetam, carisoprodol, citalopram, clobazam, clopidogrel [§] , dexlansoprazole, diazepam, doxepin, drospirenone and ethinyl estradiol, escitalopram, esomeprazole, flibanserin, formoterol, lacosamide, lansoprazole, omeprazole, pantoprazole, phenytoin, prasugrel, rabeprazole, ticagrelor, voriconazole
5. <i>CYP2C9</i>	Phenytoin, warfarin	Acenocoumarol, celecoxib	Celecoxib, dronabinol, flibanserin, flurbiprofen, lesinurad, phenytoin, piroxicam, prasugrel, warfarin
6. <i>CYP2D6</i>	Amitriptyline, clomipramine, codeine, desipramine, doxepin, fluvoxamine, imipramine, nortriptyline, ondansetron, paroxetine, tamoxifen, trimipramine, tropisetron	Aripiprazole, atomoxetine, brexpiprazole, dextromethorphan, eliglustat, methylphenidate, mirtazapine, oxycodone, pimozone, protriptyline, quinidine, risperidone, sertraline, tramadol, venlafaxine, vortioxetine	Amitriptyline, arformoterol, aripiprazole, aripiprazole lauroxil, atomoxetine, brexpiprazole, cariprazine, carvedilol, cimetidine, citalopram, clomipramine, clozapine, codeine [§] , darifenacin, desipramine, desvenlafaxine, deutetrabenazine, dextromethorphan and quinidine, doxepin, duloxetine, eliglustat, escitalopram, fesoterodine, flibanserin, fluoxetine, fluvoxamine, formoterol, galantamine, gefitinib, iloperidone, imipramine, lofexidine, meclizine, metoprolol, mirabegron, modafinil, nebivolol, nefazodone, nortriptyline, ondansetron, palonosetron, paroxetine, perphenazine, pimozone, propafenone, propranolol, protriptyline, quinidine, quinine sulfate, risperidone, rucaparib, tetrabenazine, thioridazine, tolterodine, tramadol [§] , trimipramine, umecidinium, valbenazine, venlafaxine, vortioxetine
7. <i>CYP3A4</i>	None	Tacrolimus [‡]	None
8. <i>CYP3A5</i>	Tacrolimus	None	Prasugrel
9. <i>CYP4F2</i>	Warfarin	Acenocoumarol, phenprocoumon	None
10. <i>DPYD</i>	Capecitabine, fluorouracil, tegafur	None	Capecitabine, fluorouracil
11. <i>G6PD</i>	Rasburicase	Chloramphenicol, chloroquine, chlorpropamide, ciprofloxacin, dapsone, dimercaprol, erythromycin, glibenclamide, glipizide, hydroxychloroquine, levofloxacin, mafenide, mefloquine, mesalazine, methylene blue, moxifloxacin, nalidixic acid, nitrofurantoin, norfloxacin, pegloticase, phenazopyridine, primaquine, probenecid, quinine, sodium nitrite, sulfacetamide, sulfadiazine, sulfamethoxazole/trimethoprim, sulfasalazine, sulfisoxazole	Ascorbic acid/PEG-3350/ potassium chloride/sodium ascorbate/sodium chloride/sodium sulfate, chloroquine, chlorpropamide, dabrafenib, dapsone, erythromycin and sulfisoxazole, glimepiride, glipizide, glyburide, lidocaine and prilocaine, mafenide, methylene blue, metoclopramide, nalidixic acid, nitrofurantoin, pegloticase [§] , primaquine, quinine sulfate, rasburicase [§] , sodium nitrite, succimer, sulfamethoxazole and trimethoprim, sulfasalazine, trametinib
12. <i>IFNL3</i>	Peginterferon alfa-2a, peginterferon alfa-2b, ribavirin	None	Boceprevir, daclatasvir, dasabuvir/ombitasvir/paritaprevir/ritonavir, elbasvir and grazoprevir, ledipasvir and sofosbuvir, ombitasvir/paritaprevir/ritonavir, peginterferon alfa-2b, simeprevir, sofosbuvir, sofosbuvir and velpatasvir, sofosbuvir/velpatasvir/voxilaprevir, telaprevir
13. <i>ITPA</i>			
14. <i>NAT2</i>	None	Hydralazine [‡] , isosorbide dinitrate [‡] , isoniazid [‡]	None

FDA and CPIC guidelines were last accessed in January 2019.

[†]The US FDA Guidance for PGx does not necessarily indicate genetic testing is required or recommended [12].

[‡]CPIC Level C or D.

[§]Contains an FDA Boxed Warning that contains PGx information.

CPIC levels and clinical context:

CPIC level A: genetic information should be used to change prescribing of affected drug.

CPIC level B: genetic information could be used to change prescribing of the affected drug because alternative therapies/dosing are extremely likely to be as effective as safe as nongenetically based dosing.

CPIC level C: there are published studies at varying levels of evidence, some with mechanistic rationale, but no prescribing actions are recommended because (a) dosing based on genetics makes no convincing difference or (b) alternatives are unclear, possibly less effective, more toxic or otherwise impractical or (c) few published studies or mostly weak evidence and clinical actions are unclear. Most important for genes that are subject of other CPIC guidelines or genes that are commonly included in clinical or direct-to-consume (DTC) tests.

CPIC level D: There are few published studies, clinical actions are unclear, little mechanistic basis, mostly weak evidence or substantial conflicting data. If the genes are not widely tested for clinically, evaluations are not needed.

CPIC: Clinical Pharmacogenetics Implementation Consortium.

Table 2. DNA2Rx™ gene list and associated drugs with Clinical Pharmacogenetics Implementation Consortium or the US FDA Guidance (cont.).

Gene	CPIC guidelines (any rating level) [3]	CPIC rating level A or B (no CPIC guidelines), unless specified [3]	FDA Guidance [2] [†]
15. <i>NUDT15</i>	Azathioprine, mercaptopurine, thioguanine	None	Mercaptopurine, thioguanine
16. <i>RYR1</i>	Desflurane, enflurane, halothane, isoflurane, methoxyflurane, sevoflurane, succinylcholine	None	Sevoflurane
17. <i>SLCO1B1</i>	Simvastatin	Cerivastatin	Rosuvastatin
18. <i>TPMT</i>	Azathioprine, mercaptopurine, thioguanine	None	Azathioprine, cisplatin, mercaptopurine, thioguanine
19. <i>TYMS</i>	None	Capecitabine [‡] , fluorouracil [‡]	None
20. <i>UGT1A1</i>	Atazanavir	Belinostat, irinotecan	Arformoterol, belinostat, binimetinib, dolutegravir, indacaterol, irinotecan, nilotinib, pazopanib, raltegravir
21. <i>VKORC1</i>	Warfarin	None	Warfarin

FDA and CPIC guidelines were last accessed in January 2019.

[†]The US FDA Guidance for PGx does not necessarily indicate genetic testing is required or recommended [12].

[‡]CPIC Level C or D.

[§]Contains an FDA Boxed Warning that contains PGx information.

CPIC levels and clinical context:

CPIC level A: genetic information should be used to change prescribing of affected drug.

CPIC level B: genetic information could be used to change prescribing of the affected drug because alternative therapies/dosing are extremely likely to be as effective as safe as nongenetically based dosing.

CPIC level C: there are published studies at varying levels of evidence, some with mechanistic rationale, but no prescribing actions are recommended because (a) dosing based on genetics makes no convincing difference or (b) alternatives are unclear, possibly less effective, more toxic or otherwise impractical or (c) few published studies or mostly weak evidence and clinical actions are unclear. Most important for genes that are subject of other CPIC guidelines or genes that are commonly included in clinical or direct-to-consumer (DTC) tests.

CPIC level D: There are few published studies, clinical actions are unclear, little mechanistic basis, mostly weak evidence or substantial conflicting data. If the genes are not widely tested for clinically, evaluations are not needed.

CPIC: Clinical Pharmacogenetics Implementation Consortium.

Figure 1. Structure of a molecular inversion probe. A 75 oligonucleotides DNA strand designed to capture 112 bases of a target DNA region.

sequencing through the adjustment of probe concentrations (rebalancing). An initial NGS library was produced using MIPs at equimolar concentrations, MIPs were then ranked by their respective sequencing depth. Reaction concentrations of the lowest performing probes were increased by tenfold to 35-fold. A total of 453 MIPs were designed to target 20.8 kbp. Of the 453 probes, 222 were designed to sequence HLA regions (4.8 kbp); they were included in the library preparations but were not included in this analysis. After rebalancing experiments, the optimized concentration ratios for the target capture reaction included 309 probes at 0.789 pM (1×), 133 probes at 7.891 pM (10×) and 11 probes at 27.618 pM (35×).

A few key adaptations were made to the published protocol to generate the library. Hemo KlenTaq (New England BioLabs, MA, USA) was used for the gap-fill step of the MIP capture reaction. For each sample, three 25-μl PCR reactions (18 cycles each) were completed for each MIP capture reaction instead of one to minimize PCR duplicates. TruGrade® Ultramer DNA oligonucleotides (IDT) were used for barcoded primers to prevent cross-contamination

Figure 2. Library generation of the pharmacogenome for next-generation sequencing.
MIP: Molecular inversion probe.

that lead to demultiplexing errors. The pooled libraries were purified using AxyPrep Mag PCR clean-up kit (Axygen Scientific, Inc., CA, USA) at a ratio of 0.9 instead of 1.8 $\times$. The libraries underwent the purification process twice to improve isolation of the target capture. Libraries were resuspended in a final volume of 50 μ l of EB buffer (Qiagen, CA, USA) during the first purification process and then in 35 μ l of EB buffer (Qiagen) during the second purification process. Purified PCR fragments subsequently underwent paired-end sequencing on a single MiSeq lane with the 300-cycle MiSeq reagent kit v2 (Illumina, CA, USA). For the performance assessment of DNA2Rx, three library pools were prepared containing 22–24 samples each, and sequenced separately on a MiSeq.

Bioinformatics & genotype calling analysis

Paired-end sequences were merged using the Fast Length Adjustment of Short Reads (FLASH) tool (v1.2.11) [17] and PCR duplicates were removed using BBMap [18]. Deduplicated sequences were aligned to the human reference genome sequence (GRCh37 build) using the BWA-MEM algorithm from Burrows–Wheeler Aligner package (BWA, v0.7.17) [19]. Bases with <30 phred quality scores and sequences derived from the MIP targeting arms were masked using in-house programs. SAMtools (v1.3) [20] was used to convert and index sequence files for variant calling with Freebayes (v1.1.0) [21]. A minimum sequencing depth of 20 $\times$ was used to filter the called variants. If key SNPs (i.e., loss-of-function or gain-of-function variants) with a general US population frequency of >10% failed to be called, a diplotype call for the gene was not made. Diplotype calls were compared with the GeT-RM diplotype consensus calls for concordance.

Results

Selection of genes & genetic variants

A subset of genes on this panel is considered actionable, as there are associated CPIC guidelines with these genes, allowing for immediate clinical implementation of PGx information in the drug prescribing process. The presence of FDA guidance alone did not qualify associated PGx genes as actionable since directions for clinical implementation may not be available, and the presence of FDA guidance does not necessarily indicate genetic testing is required or recommended [12]. Therefore, based on the presence of CPIC guidelines, the 15 actionable PGx genes on the panel include: *CYP2B6*, *CYP2C19*, *CYP2C9*, *CYP2D6*, *CYP3A5*, *CYP4F2*, *DPYD*, *G6PD*, *IFNL3*, *NUDT15*, *RYR1*, *SLCO1B1*, *TPMT*, *UGT1A1* and *VKORC1*. These 15 actionable PGx genes provide guidance for 40 unique drugs. Table 3 summarizes the phenotype subgroups for each actionable gene per CPIC guidelines that would qualify for a PGx-based intervention when the associated drug is prescribed. Frequencies for phenotype subgroups are provided for major ethnic groups using data from CPIC supplements [22–38] for all genes except for *CYP2C9/VKORC1* where the ENGAGE AF-TIMI 48 trial was used [39]. Priority selection of genetic markers included on the panel was given to those that define key phenotype subgroups that require change in drug dosing or therapeutic agent selection per CPIC guidelines, which were mostly loss-of-function or gain-of-function variants [22–38]. Additional selected genetic markers aided in distinguishing among several haplotypes, as multiple genetic markers are sometimes used to define a haplotype. In highly polymorphic genes (i.e., *G6PD* and *CYP2D6*), frequency data were used to exclude very rare haplotypes.

The remaining six genes (*ABCB1*, *BCHE*, *CYP3A4*, *ITPA*, *NAT2* and *TYMS*) were included in the panel because of the potential of becoming part of the actionable pharmacogenome in the future. For instance, a few of these genes are associated with drugs that have a high CPIC rating of A or B. Key genetic markers for these six genes were determined using PharmGKB and PharmVar databases [40,41]. Table 1 summarizes the alleles and genetic markers included for each gene on DNA2Rx. Table 2 categorizes the associated drugs with CPIC guidelines, CPIC rating level A or B, and FDA guidance for each PGx gene that are captured.

In addition to these genes, *HLA-B*, which has CPIC guidelines, and *HLA-A* were also included in the panel, but were not analyzed here because of the complexities involved in targeting these regions for NGS and variant phasing. CPIC guidelines also exist for *CACNA1S*, however, they were published after the development of the assay, therefore this gene was not included in the panel.

Performance assessment of DNA2Rx

Across the sequenced samples, 71.3% (mean, SD = 5.1%) of deduplicated reads mapped to the desired locations. The sequencing depth for each SNP is presented in Supplementary Figure 1A, and the percentages of on-target reads from each MIP are presented in Supplementary Figure 1B and C. Approximately 50% of the MIPs target the genome with specificity with almost no off-target reads (Supplementary Figure 1C). Median sequencing depth ranged from 13x to 2837x, and the median sequencing depth of three SNPs fell below the 20x coverage cutoff for genotype calling.

Table 4 summarizes the performance assessment of DNA2Rx. DNA2Rx contains 21 genes and tests a total of 218 variants. A total of 53 unique samples were sequenced: 37 samples were sequenced once and 16 were sequenced twice to assess repeatability. Including replicates, there were a possible 14,995 variant calls that could have been made for the 53 samples 14,052 (93.7%) of them were successfully called. One of the probes, designed to capture the *CYP3A4**22 variant, was not sequenced in 47 samples. The 943 variant calls (6.3%) that were unsuccessful did not meet the minimum 20x sequencing depth threshold. Of the 14,052 successful calls, 3254 were in samples sequenced twice; discordant results were observed in three of these variant calls. Overall, the repeatability rate was 99.9%.

The GeT-RM reference genotypes are reported as diplotypes. For the 21 genes in 53 samples sequenced, there was a total of 1449 diplotypes that could have been called. 91.7% of possible diplotypes were successfully called (Figure 3). GeT-RM reference information was missing for 313 diplotypes so concordance with DNA2Rx could not be assessed for these diplotypes. From the 1136 diplotype calls that could be compared with the GeT-RM published results, 974 (85.7%) were concordant between the GeT-RM consensus genotype calls and calls made with DNA2Rx.

For the 162 discordant diplotype calls, major reasons for disagreement were due to differences in alleles tested, failed DNA2Rx variant calls, DNA2Rx not testing for *CYP2D6* copy number variants, GeT-RM samples with incorrect ploidy assignment which affects *G6PD* haplotypes/diplotypes calls, a combination of variants from

Table 3. Actionable pharmacogenome: phenotype subgroups per Clinical Pharmacogenetics Implementation Consortium guidelines and frequencies in major ethnic groups.

Gene	Subgroup	African American frequency	Caucasian frequency	East Asian frequency	Associated drug(s)
1. <i>CYP2C19</i>	Ultrarapid metabolizer	27.35%	31.53%	1.72%	Amitriptyline, citalopram, clomipramine, desipramine, doxepin, escitalopram, imipramine, trimipramine, voriconazole
	Intermediate metabolizer	32.35%	26.76%	47.04%	Clopidogrel
	Poor metabolizer	4.28%	2.53%	14.45%	Amitriptyline, citalopram, clomipramine, clopidogrel, desipramine, doxepin, escitalopram, imipramine, sertraline, trimipramine, voriconazole
2. <i>CYP2C9</i>	Intermediate metabolizer	23.06%	31.99%	6.63%	Phenytoin
	Poor metabolizer	1.77%	4.0%	0.12%	Phenytoin
3. <i>CYP2D6</i>	Ultrarapid metabolizer	3.44%	3.27%	1.15%	Amitriptyline, clomipramine, codeine, desipramine, doxepin, imipramine, nortriptyline, ondansetron, paroxetine, sertraline, trimipramine, tropisetron,
	Normal metabolizer or intermediate metabolizer	0.75%	1.14%	0.12%	Tamoxifen
	Intermediate metabolizer	13.21%	7.18%	8.66%	Amitriptyline, clomipramine, desipramine, doxepin, imipramine, nortriptyline, tamoxifen, trimipramine
	Poor metabolizer	3.05%	6.07%	0.90%	Amitriptyline, clomipramine, codeine, desipramine, doxepin, fluvoxamine, imipramine, nortriptyline, paroxetine, tamoxifen, trimipramine
4. <i>CYP3A5</i> [†]	Extensive metabolizers	20.52%	0.61%	6.60%	Tacrolimus
	Intermediate metabolizers	49.56%	14.38%	38.19%	Tacrolimus
5. <i>CYP4F2</i>	Decreased function	7.68%	29.96%	22.05%	Warfarin
6. <i>DPYD</i>	Intermediate metabolizer and poor metabolizer	NA	NA	NA	Capecitabine, fluorouracil, tegafur
7. <i>G6PD</i>	Deficient	7.5% [‡]	3.9% [§]	4.7% [¶]	Rasburicase
8. <i>IFNL3</i>	Unfavorable response genotype	85.0%	62.0%	23.0% [¶]	Peginterferon alfa-2a, peginterferon alfa-2b, ribavirin
9. <i>NUDT15</i>	Intermediate metabolizer	NA	0.40%	16.81%	Azathioprine, mercaptopurine, thioguanine
	Possible intermediate metabolizer	NA	0.00%	0.47%	Azathioprine, mercaptopurine, thioguanine
	Poor metabolizer	NA	0.00%	0.91%	Azathioprine, mercaptopurine, thioguanine
10. <i>RYR1</i>	Malignant hyperthermia susceptible	0.00077%	0.00068%	0.00022%	Desflurane, enflurane, halothane, isoflurane, methoxyflurane, sevoflurane, succinylcholine
11. <i>SLCO1B1</i> [†]	Intermediate function	1.98%	26.77%	25.67%	Simvastatin
	Low function	0.01%	2.53%	2.29%	Simvastatin
12. <i>TPMT</i>	Intermediate metabolizer	14.90%	11.72%	3.40%	Azathioprine, mercaptopurine, thioguanine
	Possible intermediate metabolizer	0.67%	0.03%	0.01%	Azathioprine, mercaptopurine, thioguanine
	Poor metabolizer	0.73%	0.44%	0.03%	Azathioprine, mercaptopurine, thioguanine
13. <i>UGT1A1</i>	Poor metabolizer	NA	NA	NA	Atazanavir
14. <i>CYP2C9/VKORC1</i>	Sensitive responders	NA	35.4%	NA	Warfarin
	Highly sensitive responders	NA	2.9%	NA	Warfarin

[†] Alleles were reported for *SLCO1B1* and *CYP3A5* in CPIC supplements, and the Hardy–Weinberg equation was used to approximate diplotype frequencies.

[‡] Africa.

[§] Europe.

[¶] Asia.

Frequencies from CPIC supplements [3] (accessed January 2019) for all genes except *CYP2C9/VKORC1* where the ENGAGE AF-TIMI 48 trial was used [39].

CPIC: Clinical Pharmacogenetics Implementation Consortium.

Table 4. Performance results for DNA2Rx™.

Parameter	N (%)
DNA2Rx assay	
Number of genes tested on DNA2Rx	21
Number of variants tested on DNA2Rx	218
GeT-RM reference samples tested with DNA2Rx	
Number of samples sequenced once	53
Number of samples sequenced twice (replicates)	16
Total number of samples sequenced	69
Breakdown by variants	
Successful variant calls [†]	93.7%
Repeatability [‡]	99.9%
Breakdown by diplotypes	
Successful diplotype calls [§]	91.7%
Concordance with reference calls [¶]	85.7%

[†] 14,052 variant calls were successfully made with DNA2Rx out of the possible 14,995 variant calls for 21 genes in the 69 samples sequenced.
[‡] Of the 14,052 successful variant calls made with DNA2Rx, 3254 variant calls were for the 16 replicate samples and discordant calls were observed in three variant calls.
[§] 1329 diplotype calls were successfully made with DNA2Rx out of the possible 1449 diplotype calls for 21 genes in the 69 samples sequenced. Cases where multiple diplotype were possible were considered successful calls.
[¶] 1136 of the 1449 diplotypes calls made with DNA2Rx have GeT-RM reference diplotypes. 974 diplotypes made with DNA2Rx were concordant with GeT-RM reference diplotype calls. 47 samples matched the GeT-RM result but had other possible diplotypes (unphased SNPs result in ambiguous haplotype combinations).
GeT-RM: Genetic Testing Reference Materials Coordination Program.

Figure 3. Flow diagram of concordance assessment.
[†]The categories are not mutually exclusive, some calls differ between GeT-RM and DNA2Rx™ for more than one reason.
CNV: Copy-number variant; GeT-RM: Genetic Testing Reference Materials Coordination Program.

Table 5. Cost structure of library preparation and next-generation sequencing[†].

Items	Scenario 1 [‡]	Scenario 2 [§]
Number of samples multiplexed	24	65
Reagents for MIP capture reaction with phosphorylated MIPs	\$39.26	\$106.33
Reagents for exonuclease reaction	\$12.05	\$32.64
Reagents for PCR reaction	\$87.41	\$236.74
Reagents for library purification and quantification	\$17.52	\$47.46
Reagents for next-generation sequencing (fixed cost)	\$1,080	\$1,521
Disposable supplies (pipette tips, microcentrifuge tubes, 96-well plates, 96-well plate seals and gloves)	\$84.28	\$114.19
Total cost:	\$1341	\$2078
– Total cost per sample	\$55.86	\$31.96
– Total cost per gene (21 genes)	\$2.66	\$1.52
– Total cost per actionable gene (14 genes)	\$3.99	\$2.28
Preparation time:		
– Complete preparation time up to sequence ready library preparation, including DNA preparation and quantitation	7 h	8 h

[†]The cost structure does not include costs associated with personnel time, isolating DNA samples for library preparation, equipment or data analysis. Costs are reflective of 2019 list prices.
[‡]Scenario 1 reflects sample preparation and next-generation sequencing using one lane on the Illumina MiSeq system (Illumina, CA, USA).
[§]Scenario 2 reflects sample preparation and next-generation sequencing using one lane on the Illumina HiSeq2500 Rapid Run system (Illumina).
MIP: Molecular inversion probe.

DNA2Rx that were incompatible with known haplotype data because of incomplete allele definitions, and DNA2Rx calls that could not be determined without phased genotypes.

Turnaround time & cost structure

It takes 4–6 days for DNA samples to undergo library generation (2–3 days), NGS (1–2 days) and data analysis (<1 day) to obtain genetic results. Table 5 summarizes the cost structure and technician time associated with library preparation and NGS using MiSeq and HiSeq instruments. Reagents for NGS are fixed costs regardless of the number of samples multiplexed. A HiSeq run generates more paired-end reads (600 million) than the MiSeq run (24–30 million) and is able to multiplex a higher number of samples. Subsequently, the cost for obtaining genetic sequencing per sample with a HiSeq run is lower with its higher capacity to multiplex more samples. The cost for library preparation and NGS per sample is \$31.96 with a HiSeq run at a multiplexing rate of 65 samples (\$1.52 per gene; \$2.28 per actionable gene) and \$55.86 with a MiSeq run at a multiplexing rate of 24 samples (\$2.66 per gene; \$3.99 per actionable gene). These are considered conservative multiplexing rates, as higher rates have been tested. This includes the cost of reagents and disposable laboratory supplies (i.e., pipette tips, microcentrifuge tubes, 96-well plates and covers, gloves), and excludes costs associated with personnel time, isolating DNA from samples, equipment and data analysis.

Discussion

This study assessed the ability for DNA2Rx, which uses MIPs for targeted NGS, to characterize the pharmacogenome. Overall, 1136 reference diplotype calls from the GeT-RM could be compared with DNA2Rx. Concordance was high (85.7%) despite some key constraints in achieving exactly the same diplotypes. A major reason for the 162 discordant calls was the different sets of haplotypes being tested on the assays, and not due to truly conflicting results where exactly the same data produces disparate results. This highlights the impact of disparities in variants interrogated by different genetic tests where different alleles may be targeted in the assay, or different alleles used to define a diplotype. Most observed differences between GeT-Rm and DNA2Rx involved haplotypes with normal or unknown function, and would not lead to discordant phenotypes or drug therapy recommendations. Genotyping for a comprehensive set of haplotypes is important to increase the accuracy of the results, especially if these genetic differences affect the way drugs are dosed or selected. Another source of discordant calls was due to failed variant calls, which reflects the ability of the MIP to capture the targeted genomic region.

The main reason for a lack of efficiency of a probe is competing sequences in the rest of the genome. Rebalancing mostly corrected the numbers of reads that we obtain to above 20x coverage level but this will not change probe specificity. The sequencing depth for *CYP2D6**4 (rs3892097), the most common loss-of-function *CYP2D6* allele,

did not reach the minimum sequencing depth of 20x for most samples, even after the corresponding MIP was redesigned and rebalanced. This highlights the limitations of using targeted NGS technologies, to capture certain genomic loci that are highly polymorphic or have high homology with other genomic regions.

We compared the cost of obtaining genetic results for one sample to a published study that has investigated the cost structure of using MIPs to undergo targeted NGS. The study multiplexed 192 samples with 2069 MIPs, which included information for 44 genes, on a single HiSeq run [9]. The cost per sample per gene was \$0.33, which included costs associated with molecular inversion probes, reagents for library generation and sequencing and plastic ware [9]. The cost to obtain information per gene for our assay on a HiSeq run was higher (\$1.52) because our assay targeted fewer genes and multiplexed fewer samples together on one HiSeq run. There is capacity to include more genes and genetic markers, thus decreasing the cost of obtaining genetic results per sample per gene.

As the pharmacogenome continues to expand with newly established gene–drug associations, DNA2Rx has the flexibility for new MIPs to be incorporated to capture additional relevant genomic regions. For instance, CPIC guidelines were recently updated to include the *RYR1* and *CACNA1S* genes [42]. DNA2Rx does not include all 48 *RYR1* variants and future revisions can include these, as well as the two variants in *CACNA1S* that are actionable per CPIC guidelines. One limitation of the pre-emptive PGx testing approach is that if newly established PGx genetic markers are discovered, samples may need to undergo genetic testing again to obtain information for these new genomic regions. Since this genotyping methodology requires little DNA input per sample library preparation (120 ng), samples can be re-run with an updated set of MIPs. The MIP probes capture 20.8 kb of genomic sequence and covers on average 30% of the coding region of each gene. DNA2Rx has the capacity to be used as a discovery assay given the extensive regions that were captured, but reporting on these new variants was outside the scope of this analysis. Only SNPs that would be considered to have functional significance, or was required to appropriately call the haplotype, as already defined from CPIC guidelines were reported.

Several limitations should be considered when interpreting the adequacy of DNA2Rx in characterizing the pharmacogenome. First, the technological approach of generating short-read sequencing read lengths (<300 bp) does not provide ideal sequence data for complex PGx genes, such as *CYP2D6*, *HLA-A* and *HLA-B*. Characterizing *CYP2D6* copy-number variants (CNV), which include full-gene deletion, duplication and rearrangements with *CYP2D7* [43], is critical in proper assignment of *CYP2D6* metabolizer status. Determination of *CYP2D6* CNV using short-read sequencing is possible but challenging given that *CYP2D6* is highly polymorphic (>100 alleles) and homology with neighboring pseudogenes (*CYP2D7* and *CYP2D8*) [43]. This limitation can be addressed by complementing DNA2Rx with a Taqman assay quantitative real-time PCR approach to detect *CYP2D6*-specific CNVs [44,45]. Software packages are being optimized for *CYP2D6* allele calls using NGS data, and future use can be incorporated for more accurate *CYP2D6* genotype designations [43,46]. Similarly, *HLA* typing with short-read sequencing is challenging since this region is the most highly polymorphic region of the genome, as *HLA-A* has more than 300 alleles and *HLA-B* has more than 500 alleles [47]. DNA2Rx does not provide *HLA* typing given the complexity of this region, but future inclusion may be possible as *HLA* typing software packages are optimized for short-read NGS data [48,49]. Additionally, for *NAT2*, where the actionable at-risk genotype is associated with the slow acetylator phenotype, the NAT2PRED algorithm can be used to predict the *NAT2* acetylator phenotype because of the difficulty in *NAT2* haplotype inference [50].

In addition to the limitations that are related to the type of genes that are optimal for inclusion on DNA2Rx, there are implementation limitations that need to be considered. A key quality that makes this assay affordable is its use of NGS, which allows for parallel sequencing and multiplexing of samples. A major limitation of this approach is that the affordability of this assay assumes multiple samples are being processed simultaneously, and is not cost-efficient method if only a few samples are being processed for genotype results. We propose that pre-emptive testing is the way to address some of these concerns, such that multiple samples are collected well in advance of the requirement for knowing the result and turn-around times become less relevant. The turnaround time for DNA2Rx matches the median turnaround time of 14 days (range: 1–249 days) that has been cited in the literature [6]. The turnaround time of 4–6 days to return genetic results may be prolonged if collection time is lengthy to reach an optimal sample size for sequencing.

Despite these limitations, MIP-based NGS is still a feasible targeted sequencing approach in characterizing the pharmacogenome. We have completed several major research and educational initiatives using this platform where cost-efficient PGx testing of large cohorts were needed. First, we have used the assay in a retrospective analysis of cardiac catheterization laboratory patients undergoing angiography to evaluate the benefits of providing patients

with multigene PGx testing instead of the traditional single-gene PGx testing [51]. Using MIP-based NGS in PGx research that involves genotyping large cohorts is an economical way of acquiring PGx information. Second, another major initiative completed with DNA2Rx was using it to offer PGx testing as part of an educational session on PGx at a scientific conference on precision medicine held at the UNC Eshelman School of Pharmacy in 2017 [52]. Attendees were practicing healthcare professionals and researchers, and offering PGx testing provided experiential learning of basic competencies in incorporating PGx information in the drug prescribing process [52]. Efforts to address the current lack of adequate instruction on PGx in training programs for students, residents, researchers and practicing pharmacists are needed, which have resulted in multiple educational initiatives across the USA. [53–58] In past years, we have used 23andMe® testing to provide PGx information for PharmD students at the UNC Eshelman School of Pharmacy as a way for them to directly experience the PGx testing process, although only a limited number of PGx genes and alleles are tested [59]. More recently, we have used DNA2Rx as the PGx test, and the ability of DNA2Rx to handle large sample sizes makes these educational initiatives financially economical. Additionally, DNA2Rx presents as a viable PGx assay for use clinically with proper Clinical Laboratory Improvement Amendments validation.

Targeting just the actionable pharmacogenome maximizes the relevance and application of the information provided and avoids unintended consequences and incidental findings associated with more comprehensive sequencing [60]. In recent years, whole exome and genome sequencing have become emerging approaches in obtaining comprehensive genetic data for numerous applications, including PGx testing, given the affordability of high-throughput genomic technologies [61,62]. However, whole exome sequencing is not specifically aimed to capture all PGx variants, as approximately 30% of PGx variants lie outside of coding regions and adequate coverage may vary from region to region [63]. Future research will likely demonstrate the benefits of whole exome and genome sequencing to obtain PGx information over targeted sequencing of PGx genes.

Conclusion

All genetic tests and the approach used to acquire genetic information are associated with some limitations. On one extreme, costs are still too high for comprehensive, high coverage whole genome sequencing while on the other extreme, costs are low for targeted genotyping but a limited number of alleles are captured. Using a multigene PGx test that uses MIPs to undergo targeted NGS, is able to characterize the pharmacogenome adequately. Future studies assessing its performance are warranted as new PGx genes and genetic markers are established and included on the assay.

Summary points

- Clinical implementation of pharmacogenetic information in the drug prescribing process may be more successful if the testing approach was pre-emptive and provided comprehensive information of the pharmacogenome.
- A multigene pharmacogenetic test, that uses molecular inversion probes for targeted next-generation sequencing to interrogate 21 pharmacogenetic genes was explored in this analysis.
- Key performance metrics of this test indicated that 93.7% of variants were successfully called with a repeatability rate of 99.9%. Reference calls were available for 78.4% of diplotype calls resulting from pharmacogenetic testing, and concordance between the panel and reference calls was 85.7%. Turnaround time to acquire genetic results from DNA samples was 4–6 days and the cost per sample was \$32–\$56.
- Molecular-inversion probe-based next-generation sequencing is a feasible targeted sequencing approach in characterizing the pharmacogenome although performance adequacy is limited in genomic areas that are highly polymorphic and have homology with other genomic regions.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: <https://www.futuremedicine.com/doi/suppl/10.2217/pgs-2019-0057>

Acknowledgments

The authors thank K Cheek for her assistance with the figures.

Financial & competing interests disclosure

The project described was supported by American Heart Association grant 18PRE33960079 to OM Dong and a UNC Eshelman Institute for Innovation grant R1020 RX03612214 to T Wiltshire. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

Institutional review board approval and informed consent was not necessary for this study.

References

Papers of special note have been highlighted as: ● of interest; ●● of considerable interest

- Schildcrout JS, Denny JC, Bowton E *et al.* Optimizing drug outcomes through pharmacogenetics: a case for preemptive genotyping. *Clin. Pharmacol. Ther.* 92(2), 235–42 (2012).
- **Characterizes the potential benefits of comprehensive pre-emptive pharmacogenetic testing in the drug prescribing process for patients.**
- US Food and Drug Administration. Table of pharmacogenomic biomarkers in drug labeling (2019). www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm
- CPIC: Clinical Pharmacogenetics Implementation Consortium (2019). <https://cpicpgx.org/>
- Relling MV, Evans WE. Pharmacogenomics in the clinic. *Nature* 526(7573), 343–350 (2015).
- Johnson JA, Weitzel KW. Advancing pharmacogenomics as a component of precision medicine: how, where, and who? *Clin. Pharmacol. Ther.* 99(2), 154–156 (2016).
- Luzum JA, Pakyz RE, Elsey AR *et al.* The pharmacogenomics research network translational pharmacogenetics program: outcomes and metrics of pharmacogenetic implementations across diverse healthcare systems. *Clin. Pharmacol. Ther.* 102(3), 502–510 (2017).
- **Real-world implementation of pharmacogenetic testing is discussed in this study.**
- Hayden EC. Technology: the \$1,000 genome. *Nature* 507(7492), 294–295 (2014).
- Hardenbol P, Baner J, Jain M *et al.* Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat. Biotechnol.* 21(6), 673–678 (2003).
- O’Roak BJ, Vives L, Fu W *et al.* Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 338(6114), 1619–1622 (2012).
- **The methods involved with using molecular inversion probes for targeted next-generation sequencing is discussed in considerable detail in this study.**
- Bielinski SJ, Olson JE, Pathak J *et al.* Preemptive genotyping for personalized medicine: design of the right drug, right dose, right time—using genomic data to individualize treatment protocol. *Mayo Clin. Proc.* 89(1), 25–33 (2014).
- Han SM, Park J, Lee JH *et al.* Targeted next-generation sequencing for comprehensive genetic profiling of pharmacogenes. *Clin. Pharmacol. Ther.* 101(3), 396–405 (2017).
- Vivot A, Boutron I, Ravaud P, Porcher R. Guidance for pharmacogenomic biomarker testing in labels of FDA-approved drugs. *Genet. Med.* 17(9), 733–738 (2015).
- Pratt VM, Everts RE, Aggarwal P *et al.* Characterization of 137 genomic dna reference materials for 28 pharmacogenetic genes: a GeT-RM collaborative project. *J. Mol. Diagn.* 18(1), 109–123 (2016).
- **A control set of DNA samples was used to establish consensus calls among pharmacogenetic genes and this process is detailed in this study.**
- Pratt VM, Zehnbauser B, Wilson JA *et al.* Characterization of 107 genomic DNA reference materials for CYP2D6, CYP2C19, CYP2C9, VKORC1, and UGT1A1: a GeT-RM and association for molecular pathology collaborative project. *J. Mol. Diagn.* 12(6), 835–846 (2010).
- **A control set of DNA samples was used to establish consensus calls among pharmacogenetic genes and this process is detailed in this study.**
- 1000 Genomes Project Consortium, Auton A, Brooks LD *et al.* A global reference for human genetic variation. *Nature* 526(7571), 68–74 (2015).
- Boyle EA, O’Roak BJ, Martin BK, Kumar A, Shendure J. MIPgen: optimized modeling and design of molecular inversion probes for targeted resequencing. *Bioinformatics* 30(18), 2670–2672 (2014).
- Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27(21), 2957–2963 (2011).
- Bushnell B. BMap short read aligner, and other bioinformatic tools (2019). <https://sourceforge.net/projects/bbmap/>

19. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25(14), 1754–1760 (2009).
20. Li H, Handsaker B, Wysoker A *et al.* The sequence alignment/Map format and SAMtools. *Bioinformatics* 25(16), 2078–2079 (2009).
21. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv* arXiv:1207.3907 [q-bio.GN] (2012). <https://ui.adsabs.harvard.edu/abs/2012arXiv1207.3907G/exportcitation>
22. Hicks JK, Swen JJ, Thorn CF *et al.* Clinical Pharmacogenetics Implementation Consortium guideline for *CYP2D6* and *CYP2C19* genotypes and dosing of tricyclic antidepressants. *Clin. Pharmacol. Ther.* 93(5), 402–408 (2013).
23. Gammal RS, Court MH, Haidar CE *et al.* Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for *UGT1A1* and atazanavir prescribing. *Clin. Pharmacol. Ther.* 99(4), 363–369 (2016).
24. Relling MV, Schwab M, Whirl-Carrillo M *et al.* Clinical Pharmacogenetics Implementation Consortium guideline for thiopurine dosing based on *TPMT* and *NUDT15* genotypes: 2018 update. *Clin. Pharmacol. Ther.* 105(5), 1095–1105 (2019).
25. Caudle KE, Thorn CF, Klein TE *et al.* Clinical Pharmacogenetics Implementation Consortium guidelines for dihydropyrimidine dehydrogenase genotype and fluoropyrimidine dosing. *Clin. Pharmacol. Ther.* 94(6), 640–645 (2013).
26. Hicks JK, Bishop JR, Sangkuhl K *et al.* Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for *CYP2D6* and *CYP2C19* genotypes and dosing of selective serotonin reuptake inhibitors. *Clin. Pharmacol. Ther.* 98(2), 127–134 (2015).
27. Scott SA, Sangkuhl K, Stein CM *et al.* Clinical Pharmacogenetics Implementation Consortium guidelines for *CYP2C19* genotype and clopidogrel therapy: 2013 update. *Clin. Pharmacol. Ther.* 94(3), 317–323 (2013).
28. Crews KR, Gaedigk A, Dunnenberger HM *et al.* Clinical Pharmacogenetics Implementation Consortium guidelines for cytochrome P450 2D6 genotype and codeine therapy: 2014 update. *Clin. Pharmacol. Ther.* 95(4), 376–382 (2014).
29. Gonsalves SG, Dirksen RT, Sangkuhl K *et al.* Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for the use of potent volatile anesthetic agents and succinylcholine in the context of *RYR1* or *CACNA1S* genotypes. *Clin. Pharmacol. Ther.* 105(6), 1338–1344 (2019).
30. Bell GC, Caudle KE, Whirl-Carrillo M *et al.* Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for *CYP2D6* genotype and use of ondansetron and tropisetron. *Clin. Pharmacol. Ther.* 102(2), 213–218 (2017).
31. Muir AJ, Gong L, Johnson SG *et al.* Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines for *IFNL3* (IL28B) genotype and PEG interferon-alpha-based regimens. *Clin. Pharmacol. Ther.* 95(2), 141–146 (2014).
32. Caudle KE, Rettie AE, Whirl-Carrillo M *et al.* Clinical pharmacogenetics implementation consortium guidelines for *CYP2C9* and *HLA-B* genotypes and phenytoin dosing. *Clin. Pharmacol. Ther.* 96(5), 542–548 (2014).
33. Relling MV, McDonagh EM, Chang T *et al.* Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines for rasburicase therapy in the context of G6PD deficiency genotype. *Clin. Pharmacol. Ther.* 96(2), 169–174 (2014).
34. Wilke RA, Ramsey LB, Johnson SG *et al.* The clinical pharmacogenomics implementation consortium: CPIC guideline for *SLCO1B1* and simvastatin-induced myopathy. *Clin. Pharmacol. Ther.* 92(1), 112–117 (2012).
35. Birdwell KA, Decker B, Barbarino JM *et al.* Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines for *CYP3A5* genotype and tacrolimus dosing. *Clin. Pharmacol. Ther.* 98(1), 19–24 (2015).
36. Goetz MP, Sangkuhl K, Guchelaar H-J *et al.* Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for *CYP2D6* and tamoxifen therapy. *Clin. Pharmacol. Ther.* 103(5), 770–777 (2018).
37. Moriyama B, Obeng AO, Barbarino J *et al.* Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines for *CYP2C19* and voriconazole therapy. *Clin. Pharmacol. Ther.* 102(1), 45–51 (2017).
38. Johnson JA, Caudle KE, Gong L *et al.* Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for pharmacogenetics-guided warfarin dosing: 2017 update. *Clin. Pharmacol. Ther.* 102(3), 397–404 (2017).
39. Mega JL, Walker JR, Ruff CT *et al.* Genetics and the clinical response to warfarin and edoxaban: findings from the randomised, double-blind ENGAGE AF-TIMI 48 trial. *Lancet* 385(9984), 2280–2287 (2015).
40. Gaedigk A, Ingelman-Sundberg M, Miller NA *et al.* The Pharmacogene Variation (PharmVar) Consortium: incorporation of the human cytochrome P450 (CYP) allele nomenclature database. *Clin. Pharmacol. Ther.* 103(3), 399–401 (2018).
41. Thorn CF, Klein TE, Altman RB. PharmGKB: the pharmacogenomics knowledge base. *Methods Mol. Biol.* 1015, 311–320 (2013).
42. Gonsalves SG, Dirksen RT, Sangkuhl K *et al.* Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for the use of potent volatile anesthetic agents and succinylcholine in the context of *RYR1* or *CACNA1S* genotypes. *Clin. Pharmacol. Ther.* 105(6), 1338–1344 (2019).
43. Yang Y, Botton MR, Scott ER, Scott SA. Sequencing the *CYP2D6* gene: from variant allele discovery to clinical pharmacogenetic testing. *Pharmacogenomics* 18(7), 673–685 (2017).
44. Nguyen DL, Staeker J, Laika B, Steimer W. TaqMan real-time PCR quantification strategy of *CYP2D6* gene copy number for the LightCycler 2.0. *Clin. Chim. Acta* 403(1–2), 207–211 (2009).
45. Leung EKY, Agolini E, Pei X *et al.* Validation of an extensive *CYP2D6* assay panel based on invader and TaqMan copy number assays. *J. Appl. Lab Med.* 1(5), 471–482 (2017).

46. Lee S-B, Wheeler MM, Patterson K *et al.* Stargazer: a software tool for calling star alleles from next-generation sequencing data using CYP2D6 as a model. *Genet. Med.* 21(2), 361–372 (2019).
47. Choo SY. The HLA system: genetics, immunology, clinical testing, and clinical implications. *Yonsei Med. J.* 48(1), 11–23 (2007).
48. Ka S, Lee S, Hong J *et al.* HLAScan: genotyping of the HLA region using next-generation sequencing data. *BMC Bioinformatics* 18(1), 258 (2017).
49. Xie C, Yeo ZX, Wong M *et al.* Fast and accurate HLA typing from short-read next-generation sequence data with xHLA. *Proc. Natl Acad. Sci. USA* 114(30), 8059–8064 (2017).
50. Kuznetsov IB, McDuffie M, Moslehi R. A web server for inferring the human N-acetyltransferase-2 (NAT2) enzymatic phenotype from NAT2 genotype. *Bioinformatics* 25(9), 1185–1186 (2009).
51. Dong OM, Li A, Suzuki O *et al.* Projected impact of a multigene pharmacogenetic test to optimize medication prescribing in cardiovascular patients. *Pharmacogenomics* 19(9), 771–782 (2018).
52. Dong OM, Howard RM, Church R *et al.* Challenges and solutions for future pharmacy practice in the era of precision medicine. *Am. J. Pharm. Educ.* 82(6), 620–626 (2018).
53. Weitzel KW, Aquilante CL, Johnson S, Kisor DF, Empey PE. Educational strategies to enable expansion of pharmacogenomics-based care. *Am. J. Health Syst. Pharm.* 73(23), 1986–1998 (2016).
54. Adams SM, Anderson KB, Coons JC *et al.* Advancing pharmacogenomics education in the core PharmD curriculum through student personal genomic testing. *Am. J. Pharm. Educ.* 80(1), 3 (2016).
55. Springer JA, Iannotti NV, Kane MD, Haynes K, Sprague JE. Pharmacogenomics training using an instructional software system. *Am. J. Pharm. Educ.* 75(2), 32 (2011).
56. Remsberg CM, Bray BS, Wright SK *et al.* Design, implementation, and assessment approaches within a pharmacogenomics course. *Am. J. Pharm. Educ.* 81(1), 11 (2017).
57. Owusu-Obeng A, Weitzel KW, Hatton RC *et al.* Emerging roles for pharmacists in clinical implementation of pharmacogenomics. *Pharmacotherapy* 34(10), 1102–12 (2014).
58. American Society of Health-System Pharmacists: Pharmacogenomics Certificate Program (2018). <http://elearning.ashp.org/products/6673/pharmacogenomics-certificate-program>
59. Frick A, Benton C, Suzuki O *et al.* Implementing clinical pharmacogenomics in the classroom: student pharmacist impressions of an educational intervention including personal genotyping. *Pharmacy (Basel)* 6(4), 1–15 (2018).
60. Christensen KD, Vassy JL, Phillips KA *et al.* Short-term costs of integrating whole-genome sequencing into primary care and cardiology settings: a pilot randomized trial. *Genet. Med.* 20(12), 1544–1553 (2018).
61. van El CG, Cornel MC, Borry P *et al.* Whole-genome sequencing in health care: recommendations of the European Society of Human Genetics. *Eur. J. Hum. Genet.* 21(6), 580–584 (2013).
- **The challenges associated with incorporating whole-genome sequencing into patient care are discussed in this paper.**
62. Iglesias A, Anyane-Yeboah K, Wynn J *et al.* The usefulness of whole-exome sequencing in routine clinical practice. *Genet. Med.* 16(12), 922–931 (2014).
63. Londin ER, Clark P, Sponziello M, Kricka LJ, Fortina P, Park JY. Performance of exome sequencing for pharmacogenomics. *Per. Med.* 12(2), 109–115 (2014).