A complete workflow for pharmacogenomics using the QuantStudio 12K Flex Real-Time PCR System

Introduction

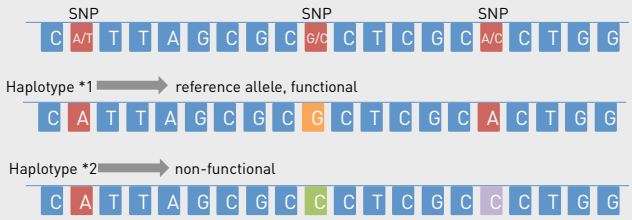
Pharmacogenomics (PGx) is the study of genetic variation as it relates to drug response. PGx studies involve testing individuals for multiple variants in drug metabolism enzyme (DME) and transporter genes. These studies are increasing in significance as personalized medicine becomes a reality in standard practice. As PGx studies continue to grow, there is a need for better solutions for sample preparation, flexibility in target selection, high sample throughput, data analysis, and lower costs.

This application note describes a complete sample-to-results PGx workflow solution using the Applied Biosystems™ QuantStudio™ 12K Flex Real-Time PCR System. We also demonstrate that TaqMan™ OpenArray™ plates run on the QuantStudio 12K Flex system provide flexibility in content, high sample throughput, quick sample-to-results turnaround time, and low cost. Finally, we describe a simple and effective data analysis method using TaqMan™ Genotyper™ Software to analyze genotyping assay results, CopyCaller™ Software to analyze copy number assay results, and the newly developed AlleleTyper™ Software for translation of these data to star allele genotypes.

ThermoFisher SCIENTIFIC

What is a star allele?

Star (*) alleles are gene-level haplotypes associated with DME phenotypes. Genetic variants within a haplotype can include single-nucleotide polymorphisms (SNPs), insertions or deletions (indels), and copy number variations (CNVs). For phenotype interpretation purposes, genotyping results must be translated to star allele nomenclature.



Materials and methods

Figure 1 illustrates the workflow described in this application note. Briefly, polyester buccal swabs were used for sample collection from 30 unrelated individuals, and DNA purified from these samples was subjected to both DME SNP genotyping and copy number variation experiments. Samples were run on the fixed-format TaqMan™ OpenArray™ PGx Panel and in 384-well plates for copy number analysis with TaqMan™ Copy Number Assays. Data were analyzed using TaqMan Genotyper Software and CopyCaller Software, and star allele results were generated with AlleleTyper Software.

Sample collection: Buccal swab samples from 30 unrelated individuals were collected using polyester swabs (Pur-Wraps™ Sterile Polyester Tipped Applicator by Puritan Hardwood). In our experience, polyester swabs are superior to cotton swabs for DNA analysis, as cotton swabs appear to contain PCR inhibitors (data not shown). Individuals were instructed to swab for 30 seconds on each side of their mouth (60 seconds total). Buccal swabs were placed swab down in 96-well plates and stored at –20°C before DNA was extracted. Buccal swabs were found to be stable at room temperature for 1–2 weeks before processing.

Sample processing: DNA was extracted from buccal swabs using the MagMAX™-96 DNA Multi-Sample Kit (we now recommend using the MagMAX™ DNA Multi-Sample Ultra Kit, Cat. No. A25597) and the MagMAX™ Express-96 Magnetic Particle Processor (we now recommend the KingFisher™ Flex Magnetic Particle Processor, Cat. No. 5400630), using an optimized protocol that was based on the MagMAX blood sample purification protocol [1]. This protocol employs Proteinase K digestion followed by treatment with a guanidinium thiocyanate–based solution. The detailed modified protocol is described in the Appendix.

DNA samples were quantified using the TaqMan™ RNase P Detection Reagents Kit (Cat. No. 4316831) and TaqMan™ DNA Template Reagents (Cat. No. 401970) to create a standard curve. Sample concentrations ranged from 10.9 to 116.2 ng/µL. The average sample concentration was

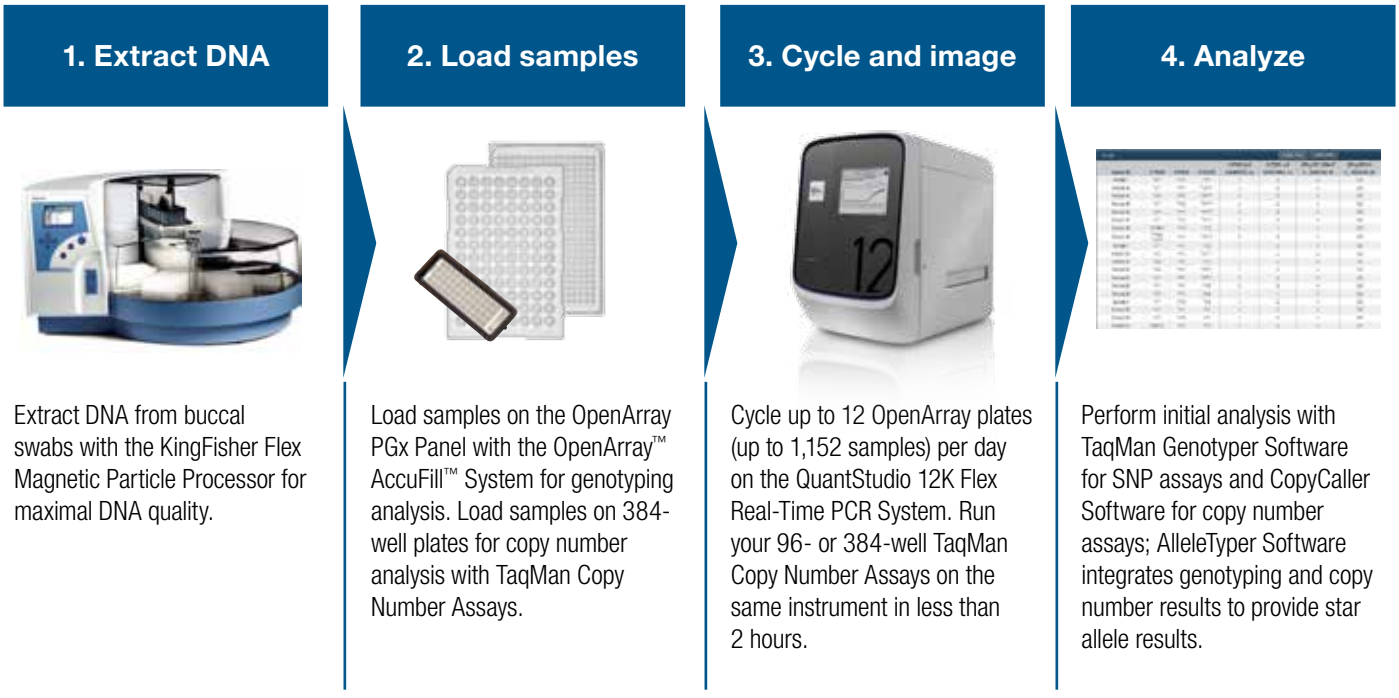


Figure 1. Experimental workflow for PGx studies, from sample preparation to data analysis.

43.0 ng/µL and the median concentration 36.8 ng/µL. (Note: DNA yields of 70–200 ng/µL are now routinely obtained using 4N6FLOQSwabs™ (Cat. No. 4473979)). The recommended concentration for OpenArray experiments is 50 ng/µL, so the samples were not diluted for use in OpenArray experiments. Aliquots of the samples were adjusted to 5 ng/µL for use in copy number variation experiments.

SNP genotyping experiments: Genotyping experiments were performed using the fixed-format TaqMan OpenArray PGx Panel (Cat. No. 4475395), which contains 158 TaqMan™ drug metabolism assays derived from the PharmaADME core marker panel (pharmaadme.org). Up to 15 samples plus one no-template control (NTC) were run per OpenArray plate; 2 plates were used to run 30 samples. Reactions were prepared and run on OpenArray plates according to the protocol in the OpenArray Experiments User Guide [2].

Copy number variation experiments: Three DME genes that exhibit copy number variation were interrogated with TaqMan Copy Number Assays: Hs00010001_cn and Hs04083572_cn (targeting exon 9 and intron 2, respectively, of *CYP2D6*), Hs07545275_cn (targeting *CYP2A6* intron 7), and Hs02575461_cn (targeting *GSTM1* exon 1). Copy number assays were run in the same well with the TaqMan™ Copy Number Reference Assay (RNase P) in PCR reactions containing 10 ng of purified DNA (30 test samples plus 4 control Coriell DNA samples) and TaqMan™ Genotyping Master Mix. Reactions were prepared and run on 384-well plates on the QuantStudio 12K Flex Real-Time PCR System according to the TaqMan Copy Number Assays protocol [3].

Data analysis: An overview of the data analysis workflow is shown in Figure 2. TaqMan SNP Genotyping Assay data (experiment .eds files) were first analyzed using the QuantStudio™ 12K Flex Software using the "Real-Time Rn - Median (Rna to Rnb)" analysis setting. Analyzed .eds files were then imported into TaqMan Genotyper Software, and data were analyzed using the "Real-Time Experiment" type and "Autocalling" method settings. Allele discrimination plots were reviewed, calls were edited as needed (e.g., occasionally a sample is called as undetermined, yet it is closely associated with a genotype cluster), and then results were exported in a .txt file using the "Advanced" settings to export the genotype calls.

TaqMan Copy Number Assay results were first analyzed using the QuantStudio 12K Flex Software with a manual C_t threshold level of 0.2 and auto-baseline settings to determine C_t values for the FAM™ dye–labeled test assays and the VIC™ dye–labeled RNase P reference assay. Exported results (.txt files) were then imported into CopyCaller Software for copy number analysis by the $\Delta\Delta C_t$ method. The median sample ΔC_t value was used as the calibrator, with a copy number value of 2 for the *CYP2D6* and *CYP2A6* assays and value of 1 for the *GSTM1* assay. Copy number results were exported in .txt files.

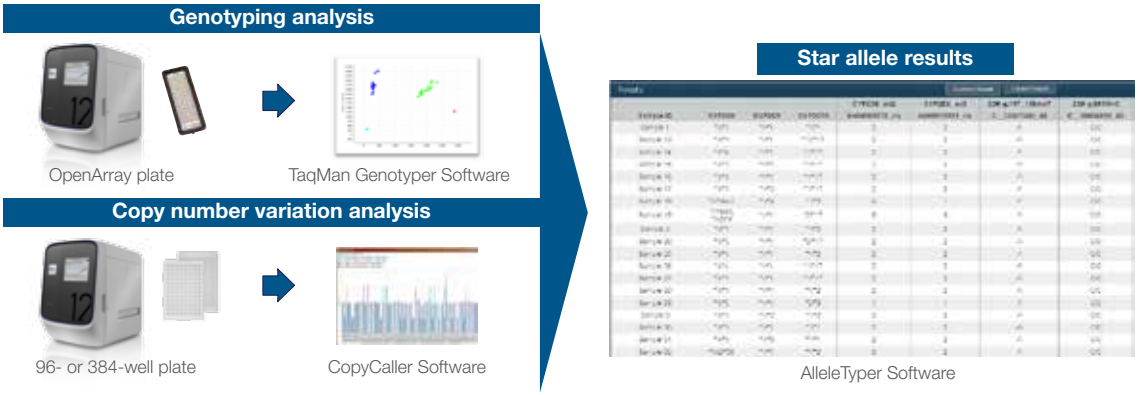


Figure 2. Data analysis workflow for PGx studies.

AlleleTyper Software

AlleleTyper Software facilitates the translation of results to star allele or other nomenclature, for individuals genotyped using TaqMan Drug Metabolism Genotyping Assays and TaqMan Copy Number Assays. User-defined monoallelic translation tables containing haplotype information for the targeted gene variants in a study are automatically converted by the software to biallelic translators containing diplotype genetic patterns. AlleleTyper Software matches the sample genotypes in results files from TaqMan Genotyper Software and CopyCaller Software to the patterns in the biallelic translator, and reports the star allele genotypes determined for each individual.

AlleleTyper Software was used to determine the sample star allele genotypes for the *CYP2D6*, *CYP2C9*, and *CYP2C19* gene variants in the OpenArray PGx panel and copy number variation experiments. A monoallelic translation table was first prepared that contained the expected genotypes for each star allele, as noted on the Cytochrome P450 Allele Nomenclature site (cypalleles.ki.se), that could be determined using the DME assays in the panel. AlleleTyper Software was used to convert the monoallelic translation table to a biallelic translator containing all possible star allele diplotype genetic patterns. TaqMan Genotyper and CopyCaller results files were then imported into AlleTyper Software, which automatically translated the sample genotype patterns to star allele genotype calls.

Results and discussion

Figure 3 shows examples of the allele discrimination cluster plot data for *CYP* gene variants that were present in individuals in this study. In general, the assays performed extremely well with these sample preparations. Of the 4,710 data points examined, a total of 4,704 data points had unambiguous genotypes, for a call rate of 99.9%.

One important consideration in selecting a platform for PGx studies is to ensure that there is enough flexibility in assay content to support evolving target selection requirements. In addition, studies are often more affordable when content is easily customizable and users are not locked into fixed content. The TaqMan OpenArray PGx Panel used in this study covers many important targets, but it may cover more genes and variants than many users will need for their studies. Additional targets may also be desired for a given panel (e.g., some key CYP2D6 targets are missing). For more specific studies, the TaqMan™ OpenArray™ PGx Express Panel (Cat. No. 4488847) covers 60 of the most common targets for PGx research applications. Users can also customize and order their own OpenArray DME and SNP genotyping assay panels. Thermo Fisher Scientific offers 2,700 TaqMan Drug Metabolism Assays designed to known and putative causal SNP, MNP, and indel variants, as well as 7 million predesigned SNP assays and custom SNP assays for other targets of interest.

Previous studies have demonstrated that several key DME genes exhibit copy number variation (CNV) [4]. To determine the genotype of an individual for variants in such genes, both SNP genotyping and CNV analysis must be done. We tested the 30 samples purified with the MagMAX system with four TaqMan Copy Number Assays to analyze three DME genes in the PharmaADME panel: *CYP2D6*, *CYP2A6* and *GSTM1*. All samples performed very well with all assays: copy number results were generated with confidence values of >99.9% for all data points. As shown in Figure 4, the copy number variation noted with this sample set was very diverse; all but two samples showed deletion or duplication of the target gene with at least one assay.

Knowing the combination of variants within a given haplotype and the diplotype content in an individual are of key importance for studying drug metabolism, drug response, and adverse drug reactions. AlleleTyper Software matches genetic information from the biallelic translator (Figure 5) and reports star allele genotypes for each individual (Figure 6). Example results are summarized below; information on the expected phenotype for the deduced genotypes is from the Pharmacogenomics Knowledgebase website (www.pharmgkb.org).

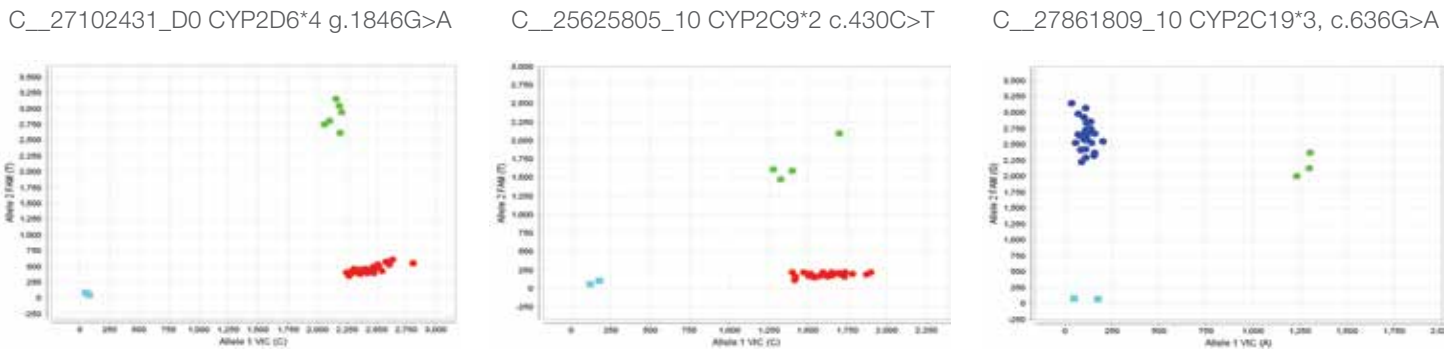


Figure 3. Example data for the 30 buccal cell DNA samples run on the TaqMan OpenArray PGx Panel and analyzed with TaqMan Genotyper Software.

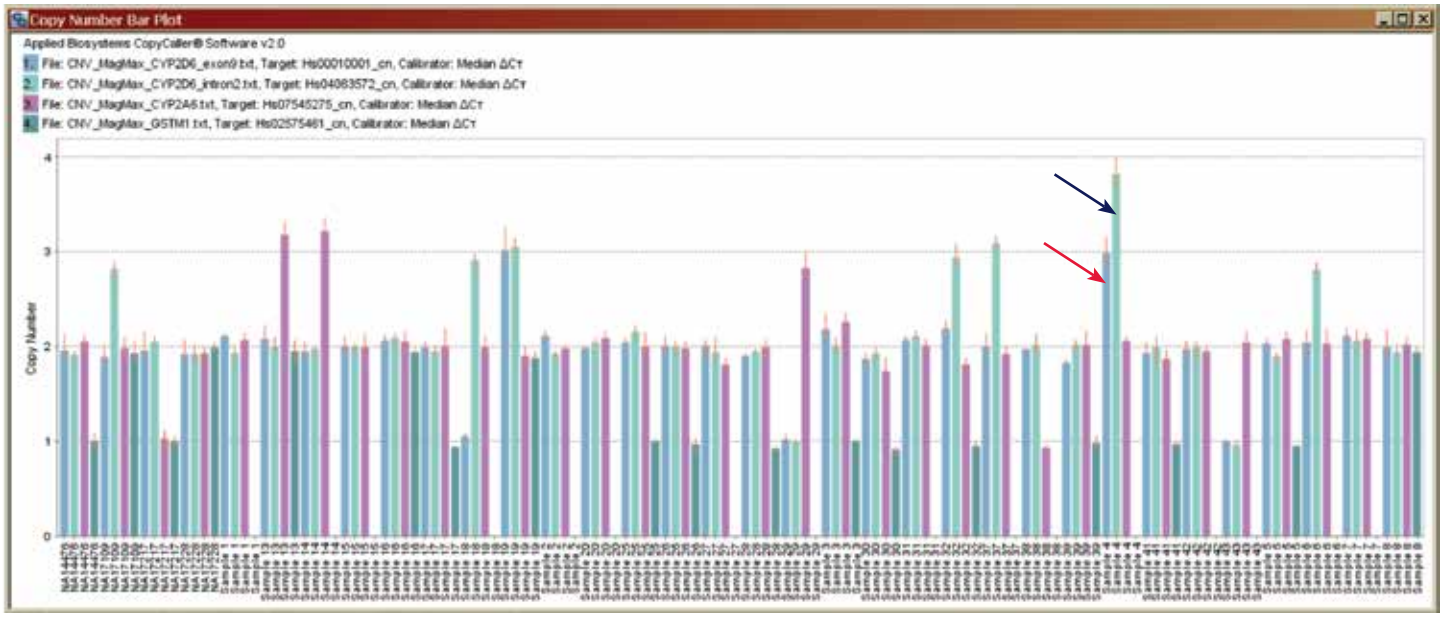


Figure 4. Copy number variation analysis of DME genes. Results are shown for all 30 samples and 4 Coriell gDNA controls in CopyCaller Software. Sample 4 carries 4 copies of *CYP2D6*, as detected by the *CYP2D6* intron 2 assay (blue arrow). The *CYP2D6* exon 9 assay (red arrow) detects 3 full-length *CYP2D6* alleles; the fourth copy is a *CYP2D6*36* allele with an exon 9 gene conversion to *CYP2D7*.

Example data

- Sample 20 is heterozygous for the rare, nonfunctional *CYP2D6*3* (2549delA) frameshift mutation and would be typed as an extensive metabolizer of some drugs. This sample also carries the loss-of-function *CYP2C19*2* (c.681G>A) splicing mutation and the gain-of-function *CYP2C19*17* (g.-806C>T) promoter mutation and has an intermediate metabolizer phenotype.
- Sample 29 carries a *CYP2D*5* gene deletion allele and a wild type allele and thus is an extensive metabolizer of some drugs. This sample also carries two *CYP2C19* loss-of-function alleles—the *CYP2C19*2* (g.-806C>T) promoter mutation and the *CYP2C19*3* (c.636G>A; W212X) nonsense mutation—and has a poor metabolizer phenotype.
- Neither Sample 20 nor 29 carries deleterious *CYP2C9* alleles; these alleles were *CYP2C9*2* (430C>T, found in four other samples) and *CYP2C9*3* (1075A>C, found in three other samples).
- Five samples carry one or more *CYP2D6*36* alleles, which are classified as having no *CYP2D6* function and which are often found in tandem with other *CYP2D6* genes [5]. The *CYP2D6*36* allele has a gene conversion to *CYP2D7* pseudogene sequences in exon 9; the copy number assay targeting intron 2 will amplify this allele whereas the one targeting exon 9 will not. Because Sample 4 carries three copies of the *CYP2D6*1* wild type allele in addition to a *CYP2D6*36* allele, it is classified as an ultra-rapid metabolizer.

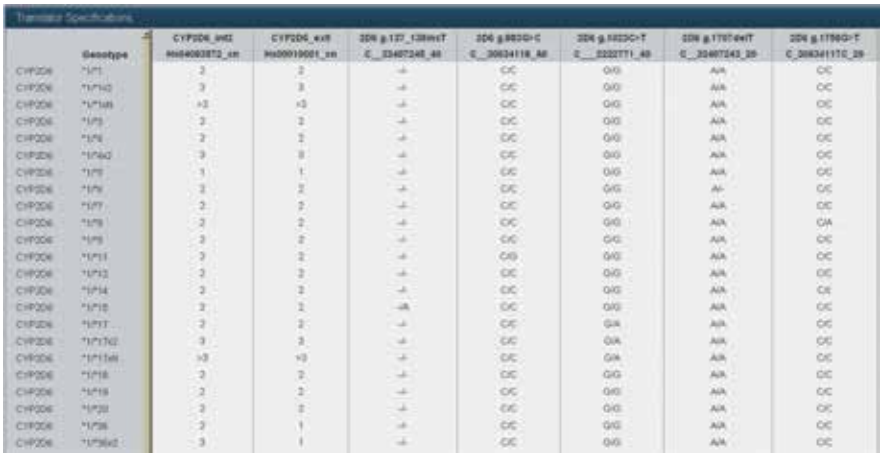


Figure 5. AlleleTyper Software showing a portion of a biallelic translation table for *CYP2D6* variants.

Results

Summary Report | Detailed Report

Sample ID	CYP2D6	CYP2C9	CYP2C19	CYP2D6_int2 Hs04083572_cn	CYP2D6_ex9 Hs00010001_cn	2D6 g.137_138insT C__32407245_40	2D6 g.883G>C C__30634118_A0
Sample 1	*1/*1	*1/*1	*1/*1	2	2	-/-	C/C
Sample 13	*1/*1	*1/*1	*17/*17	2	2	-/-	C/C
Sample 14	*1/*4	*1/*1	*1/*17	2	2	-/-	C/C
Sample 15	*1/*1	*1/*2	*1/*17	2	2	-/-	C/C
Sample 16	*1/*4	*1/*1	*1/*17	2	2	-/-	C/C
Sample 17	*1/*1	*1/*2	*1/*17	2	2	-/-	C/C
Sample 18	*1/*36x2	*1/*3	*1/*2	3	1	-/-	C/C
Sample 19	*1/*4x2, *1x2/*4	*1/*1	*2/*17	3	3	-/-	C/C
Sample 2	*1/*1	*1/*1	*1/*2	2	2	-/-	C/C
Sample 20	*1/*3	*1/*1	*2/*17	2	2	-/-	C/C
Sample 25	*1/*4	*1/*1	*1/*2	2	2	-/-	C/C
Sample 26	*1/*4	*1/*1	*1/*17	2	2	-/-	C/C
Sample 27	*1/*1	*1/*1	*1/*17	2	2	-/-	C/C
Sample 28	*1/*1	*1/*1	*1/*2	2	2	-/-	C/C
Sample 29	*1/*5	*1/*1	*2/*3	1	1	-/-	C/C
Sample 3	*1/*1	*1/*2	*1/*2	2	2	-/-	C/C
Sample 30	*1/*1	*1/*1	*1/*1	2	2	-/-	C/C
Sample 31	*1/*1	*1/*3	*1/*1	2	2	-/-	C/C
Sample 32	*1x2/*36	*1/*1	*1/*2	3	2	-/-	C/C
Sample 37	*1x2/*36	*1/*1	*1/*3	3	2	-/-	C/C
Sample 38	*1/*1	*1/*1	*2/*3	2	2	-/-	C/C
Sample 39	*1/*1	*1/*3	*1/*1	2	2	-/-	C/C
Sample 4	*1xN/*36	*1/*1	*2/*2	4	3	-/-	C/C
Sample 41	*1/*1	*1/*1	*1/*1	2	2	-/-	C/C
Sample 42	*1/*1	*1/*1	*1/*2	2	2	-/-	C/C
Sample 43	*1/*5	*1/*1	*1/*2	1	1	-/-	C/C
Sample 5	*1/*1	*1/*1	*1/*2	2	2	-/-	C/C
Sample 6	*1x2/*36	*1/*1	*1/*1	3	2	-/-	C/C
Sample 7	*1/*1	*1/*1	*1/*2	2	2	-/-	C/C
Sample 8	*1/*4	*1/*2	*1/*1	2	2	-/-	C/C

Figure 6. AlleleTyper Software showing a portion of a biallelic translation table for *CYP2D6* variants.

Conclusions

As pharmacogenomics becomes more widespread, there is a need for simple, streamlined workflow solutions to accomplish studies on varying numbers of samples and target genes. This application note addresses some of the common challenges in preparing samples for PGx studies and describes an optimized, complete workflow solution for performing both genotyping and copy number analysis using TaqMan Assays in OpenArray and 384-well plate formats on the QuantStudio 12K Flex system. We show that the MagMAX DNA Multi-Sample Kit, which can be used to isolate DNA from a number of different sample types, can provide excellent quality and yield using an optimized protocol to easily process up to 96 samples at one time. MagMAX technology offers a simple and rapid solution for the preparation of high-quality genomic DNA for OpenArray genotyping and copy number analysis workflows.

PGx studies typically use genotyping and copy number analysis to understand the impact of genetic variation on drug metabolism, drug efficacy, and adverse drug effects. Flexible genotyping assay content, short time-to-results, and an affordable platform are vital for these studies. This application note demonstrates the flexibility and high throughput of the QuantStudio 12K Flex system when combined with OpenArray plates and TaqMan Assays. The fixed-content OpenArray PGx or PGx Express panels can be used for analysis with a standard set of markers, but customized panels are also available from a broad selection of 2,700 validated TaqMan DME Assays, 7 million predesigned SNP genotyping assays, and custom SNP assays. Finally, we demonstrate time-to-results with the new automated AlleleTyper Software, where data from SNP genotyping and copy number assay experiments can be easily interpreted to finish the entire workflow from sample preparation to data analysis in a single day.

References

1. MagMAX-96 DNA Multi-Sample Kit, MagMAX Express-96 Magnetic Particle Processor protocol. P/N 4428202, Rev. B.
2. Applied Biosystems QuantStudio 12K Flex Real-Time PCR System: OpenArray Experiments User Guide. P/N 4470935, Rev. B.
3. TaqMan Copy Number Assays Quick Reference Card. P/N 4397425, Rev. D.
4. Glasel JA (1995) Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios. *Biotechniques* 18:62–63.
5. Gaedigk A, Bradford LD, Alander SW, Leeder JS (2006) CYP2D6*36 gene arrangements within the cyp2d6 locus: association of CYP2D6*36 with poor metabolizer status. *Drug Metab Dispos* 34:563–569.
6. TaqMan PreAmp Master Mix Kit protocol. P/N 4384557, Rev. B.

Appendix: Detailed protocol and protocol notes

A complete, detailed protocol can be found at **thermofisher.com/pgx**

Sample processing

DNA was extracted from buccal swabs using the MagMAX-96 DNA Multi-Sample Kit (Cat. No. 4413021) and the MagMAX Express-96 Magnetic Particle Processor (Cat. No. 4400077) using an optimized protocol based on the MagMAX blood sample purification protocol [1]. The following modifications to this protocol were made:

To extract the DNA, swab ends were placed in the wells of a 96-well plate. Proteinase K digestion was performed using a mix of 184 µL of Proteinase K buffer and 16 µL of Proteinase K solution per swab. Plates were sealed with adhesive seal and shaken for 3 min at speed 7, and then incubated for 45 min at 65°C. To each swab-containing well, 200 µL of Multi-Sample DNA Lysis Buffer was added, and then plates were sealed and shaken for 15 min at speed 7. Swabs were removed from the wells, leaving behind as much liquid as possible.

To purify the DNA, 40 µL of magnetic DNA-binding beads (32 µL DNA Binding Beads at 10 mg/mL, plus 8 µL water) was added to each well, and then plates were sealed and shaken for 10 min at speed 7; 240 µL of 100% isopropanol was added to each sample, and then plates were sealed and shaken for 5 min at speed 7. The standard blood sample protocol was followed for the remainder of the steps: plates were placed on the magnetic stand for 5 min, the supernatant was removed, and then samples were washed once with 150 µL of Wash Solution 1 and twice with 150 µL of Wash Solution 2. DNA was eluted using 50 µL each of DNA Elution Buffers 1 and 2. Samples were stored at –20°C.

Note: An updated MagMax DNA Multi-Sample Ultra Kit protocol (P/N MAN0010293) is available for high-throughput isolation of PCR-ready DNA from buccal swabs.

Assessment of sample quality

Buccal swab DNA samples collected and prepared by the methods described above were of sufficiently high purity and concentration to use directly in SNP genotyping and copy number analysis experiments. However, in some cases it may be difficult to control the sample collection (e.g., the swab type used and the amount of swabbing done), and it is possible that even the optimized MagMAX preparation method described here may not provide samples of sufficient quantity or quality for genotyping analysis on the OpenArray platform. A general guideline is to use A_{260} and A_{280} readings to assess both the concentration and the quality of the sample; conventionally, we recommend using 50 ng/µL DNA stock solutions (such that 250 ng will be loaded on the array) and A_{260}/A_{280} ratios that fall within the recommended range of 1.8–2.0. In practice, less DNA may work well (as in this study, where down to 10.9 ng/µL DNA was used). However, samples with concentrations of less than approximately 10 ng/µL may not be genotyped correctly if the concentration of amplifiable DNA is too low to properly amplify both target alleles. In addition, samples may not amplify well if PCR inhibitors are present in the preparation.

Recommendations for low-quality samples

If samples are of low quantity or contain PCR inhibitors, samples may fail to amplify, may fail to cluster properly, and in the worst-case scenario, may provide incorrect genotypes (e.g., apparent loss of heterozygosity if only one chromosomal copy is predominantly amplified and a heterozygous sample clusters within a homozygote cluster). If sample preparations perform poorly on the OpenArray platform with a number of assays, we recommend preamplifying samples (see below), which can serve both to generate higher concentrations of target amplicons and to dilute PCR inhibitors. In our experience, low-quality sample preparations can frequently be rescued by this method. In one pilot genotyping study using low-concentration and low-quality DNA preparations, the call rate was only 86.3%, but after preamplification of samples, the call rate increased to 99.7%.

Note: For copy number variation experiments, 10 ng of high-quality purified DNA is required for each of 4 replicate PCR reactions. Preamplification of DNA may be possible but is likely to compromise the data quality.

Sample preamplification

DNA preamplification reactions were performed similarly to that described in the TaqMan™ PreAmp Master Mix Kit protocol [6]. Reactions were prepared in a 96-well plate using 2.5 µL of 2X TaqMan PreAmp Master Mix (Cat. No. 4391128), 1.25 µL of preamplification assay pool (0.2X of each TaqMan DME assay used in genotyping experiments), and 1.25 µL of 0.4–4.0 ng/µL DNA. The PCR cycling conditions used were: enzyme activation at 95°C for 10 min, followed by 12 PCR cycles of 95°C for 15 sec and 60°C for 4 min, followed by enzyme inactivation at 99.9°C for 10 min. Preamplified samples were stored at 4°C. Prior to use in genotyping experiments, samples were diluted 1:20 by adding 95 µL of 1X TE solution to each sample. Please see the user bulletin TaqMan OpenArray Genotyping Sample Preamplification Guide, P/N MAN0011116.

Note: Each custom TaqMan PreAmp Pool contains a unique set of primers. We recommend performing a concordance experiment using samples of known genotype to verify performance. For assistance in designing and ordering a pool of preamplification primers, please contact your local sales representative.

Ordering information

Product	Cat. No.
MagMAX DNA Multi-Sample Ultra Kit, 500 preps	A25597
KingFisher Flex Magnetic Particle Processor	5400630
TaqMan RNase P Detection Reagents Kit	4316831
TaqMan DNA Template Reagents	401970
TaqMan OpenArray PGx Panel	4475395
TaqMan OpenArray PGx Express Panel	4488847
TaqMan PreAmp Master Mix	4391128
TaqMan Copy Number Assays	Go to **thermofisher.com/cnv**
QuantStudio 12K Flex Real-Time PCR System	Go to **thermofisher.com/quantstudio**

Find out more at **thermofisher.com/pgx**

ThermoFisher SCIENTIFIC