Research Article Open Access

# A High-throughput Real-time PCR Approach to Pharmacogenomics Studies

Toinette Hartshorne<sup>1\*</sup>, Ferrier Le<sup>1</sup>, Jordan Lang<sup>1</sup>, Harrison Leong<sup>1</sup>, Kathleen Hayashibara<sup>1</sup>, Dominique Dewolf<sup>2</sup> and Elliot Shelton<sup>1</sup>

<sup>1</sup>Thermo Fisher Scientific, Genomic Applications, South San Francisco, CA, USA

\*Corresponding author: Toinette Hartshorne, Thermo Fisher Scientific, Genomic Applications, 180 Oyster Point Blvd., South San Francisco, CA 94061, USA, Tel: 1-650-246-4081; E-mail: toinette.hartshorne@thermofisher.com

Received date: April 09, 2014, Accepted date: May 20, 2014, Published date: May 28, 2014

Copyright: © 2014 Hartshorne T, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

#### **Abstract**

Advances in personalized medicine have led to an increase in pharmacogenomics studies that involve testing individuals for drug metabolism enzyme and transporter gene polymorphisms implicated in drug response. As a consequence, there is a growing demand for affordable, easy to use technologies with fast sample-to-results workflows that can accommodate testing customizable sets of target gene variants and a changeable number of samples. Additionally, data analysis tools are needed to facilitate translation of an individual's genetic information to their diploid content of gene-level star allele haplotypes, which can be correlated with drug metabolism enzyme phenotypes. Here we describe the development of a comprehensive pharmacogenomics experiments workflow solution to meet this need. High quality data was generated from purified buccal swab DNAs run with TaqMan® SNP genotyping and copy number assays in OpenArray® and 384-well plate formats, respectively, on the QuantStudio™ 12K Flex system. Data analysis was accomplished using TaqMan® Genotyper™ Software to examine SNP genotyping assay results and CopyCaller® Software to examine copy number assay results, followed by translation of this genetic data for individual samples to star allele genotypes using the recently developed AlleleTyper™ Software. The specific TaqMan® SNP Genotyping and Copy Number Assays to gene variants used can be tailored to suit the needs of a given pharmacogenomics study. This low cost, high throughput pharmacogenomics workflow can be completed in a single day, from sample preparation to data analysis.

**Keywords:** Pharmacogenomics; DME; SNP; CNV; OpenArray; Taqman; PCR

# Introduction

Pharmacogenomics (PGx) is the study of genetic variation as it relates to drug response. PGx studies, consisting of testing individuals for multiple variants in drug metabolism enzyme (DME) and transporter genes, are increasing in significance as personalized medicine becomes a reality in standard practice. Technologies that are fast, accurate, cost-effective, and have broad marker coverage are required to support the growing number of PGx studies being conducted. Real-time PCR platforms such as the Applied Biosystems\* Quantstudio™ 12k Flex Real-time PCR System (Life Technologies, Thermo Fisher Scientific), can provide comprehensive yet flexible content and enable PGx studies to proceed quickly, accurately, and at lower costs than previously possible.

For phenotype interpretation of DME gene variants, genotyping results must be translated to star (\*) allele nomenclature [1,2]. Star alleles are haplotype patterns that have been defined at the gene level and, in many cases, associated with protein activity levels. Genetic variants within a haplotype can include single nucleotide polymorphisms (SNPs), insertions or deletions (InDels), and copy number variations (CNVs). Knowing the combination of variants within a given haplotype, and the diploid content in an individual, is of key importance for studying drug metabolism, drug response and adverse drug reactions. Software tools that facilitate translation analysis could benefit PGx research.

The objective of the study described herein was to create a comprehensive, robust workflow for high throughput PGx genotyping experiments from sample preparation to data analysis. We demonstrate a high success rate of genotyping calls for DME gene variants through our newly developed workflow, starting with DNA preparation from buccal swabs, proceeding through genotyping using the Quantstudio™12k Flex Real-time PCR System with TaqMan® SNP Genotyping Assays on OpenArray® Plates and TaqMan® Copy Number Assays on 384-well plates, and concluding with data analysis. 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.

## Materials and Methods

### Sample collection

Buccal swab samples from 30 unrelated individuals were collected using polyester swabs (Pur-Wraps\* Sterile Polyester Tipped Applicator by Puritan Hardwood). 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.

<sup>&</sup>lt;sup>2</sup>Thermo Fisher Scientific, Life Sciences Solution, Saint Aubin, France

#### Sample processing

DNA was extracted from buccal swabs using the MagMAX™-96 DNA Multi-Sample Kit (Cat. No. 4413021) plus additional components (Cat. No. 4489110, 4489111 and 4489112) and the MagMAX™ Express-96 Magnetic Particle Processor (Cat. No. 4400077), using an optimized protocol that was based on the MagMAX™ blood sample purification protocol versus the MagMAX™ buccal swab purification protocol [3]. This protocol employs an enzymatic Proteinase K digestion followed by treatment with a guanidinium thiocyanate-based solution. Optimization steps included doubling the quantities of the Proteinase K enzyme solution and the magnetic DNA-binding beads. The detailed modified protocol is described in the Pharmacogenomics Experiments User Guide (Pub. No. MAN0009612).

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 43.0 ng/µL and the median concentration was 36.8 ng/μL. The recommended concentration for OpenArray® experiments is 50 ng/ $\mu L$ , so the samples were not diluted for use in OpenArray experiments. Aliquots of the samples had their concentrations adjusted to 5 ng/µL for use in copy number variation experiments.

### SNP genotyping experiments

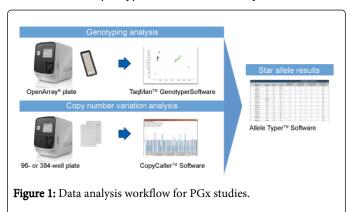
Genotyping experiments were performed using the fixed-format TagMan OpenArray PGx Panel (Cat. No. 4475395), which contains 158 TaqMan drug metabolism assays to SNP and InDel variants derived from the PharmaADME core marker panel (see www.pharmaadme.org). Up to 15 samples plus 1 no-template control (NTC) 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 Applied Biosystems OpenArray Experiments User Guide [4].

# 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 [5].

# Data analysis

An overview of the data analysis workflow is shown in Figure 1. 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<sub>t</sub> threshold level of 0.2 and auto baseline settings to determine C<sub>t</sub> 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  $\Delta$ Ct 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.



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 haplotype information for each star allele, as noted on the Cytochrome P450 Allele Nomenclature site (www.cypalleles.ki.se) [2], that could be determined by 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<sup>®</sup> Genotyper<sup>™</sup> and CopyCaller<sup>®</sup> results files were then imported into AlleleTyper™ Software, which automatically translated the sample genotype patterns to star allele genotype calls.

#### **Results and Discussion**

The study described in this paper was undertaken to develop a straightforward, comprehensive PGx experiments workflow. Briefly, polyester buccal swabs were used for sample collection from 30 unrelated individuals, and DNAs purified from these samples were subjected to both DME gene variant 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 and CopyCaller Software, and star allele results were generated with AlleleTyper™ Software.

Easy-to-collect buccal swab samples are often used in PGx studies, yet obtaining high-quality, inhibitor-free DNA preparations in sufficient quantity from swabs can be difficult. To address this problem, we developed an optimized DNA purification protocol using the MagMAX™ DNA isolation system. DNA samples collected and prepared by the described methods were of sufficiently high purity and concentration to use in both SNP genotyping and copy number analysis. In our experience, polyester swabs are superior to cotton swabs for DNA analysis, as cotton swabs appear to contain PCR inhibitors (data not shown). Since this study was done, we have significantly improved DNA yields per polyester swab from 11-116 ng/ µL to 70–200 ng/µL by using 4N6FLOQSwabs™.



Figure 2: Example data for the 30 buccal swab samples run on the TaqMan OpenArray PGx Panel and analyzed with TaqMan Genotyper Software.

The DME SNP genotyping assays on the TaqMan OpenArray PGx Panel were derived from the PharmaADME core marker list (see www.PharmaADME.org). Many of the targeted DME variants are rare, thus only 36% of the assays detected the minor allele in the 30 individuals tested in this study; wild type homozygous and heterozygous clusters were noted with 23 assays and all 3 genotypes were present with 34 assays. Figure 2 shows examples of the allele discrimination cluster plot data for CYP gene variants that were present in these individuals. 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%.

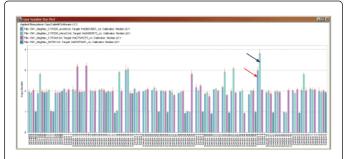


Figure 3: Copy Number Variation analysis of DME genes. Results are shown for all 30 samples and 4 Coriell gDNA controls in CopyCaller Software. Sample 4 carries 4 copies of CYP2D6 as detected by the CYP2D6 intron 2 assays (blue arrow). The CYP2D6 exon 9 assay (red arrow) detects 3 full-length CYP2D6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

Several key DME genes are known to exhibit copy number variation [6]. To determine the genotype of an individual for variants in such genes, both SNP genotyping and CNV analysis must be done. Figure 3 shows the copy number analysis results for the 30 samples tested with four TaqMan Copy Number Assays to probe 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. The copy number variation noted with this sample set was diverse, with all but two samples showed CNV of at least one gene. In most individuals, one or both copies of the GSTM1 gene was deleted (deletion allele frequency of 70%). For the CYP2A6 and CYP2D6 genes, individuals having single gene deletion or duplication alleles were noted, as were 6 individuals having nonfunctional CYP2D6/CYP2D7 hybrid alleles carrying an exon 9 conversion to CYP2D7 sequences (e.g. CYP2D6\*36 is not detected by the exon 9 assay but is detected by the intron 2 assay) [7,8].



Figure 4: AlleleTyper Software showing a portion of a biallelic translation table for CYP2D6 variants.

Star allele nomenclature is used to describe gene-level haplotypes that have been associated with DME phenotypes (e.g., functional or nonfunctional variants) [1,2]. The combination of SNPs, InDels and CNVs (if applicable) within a given gene must be taken into account to determine the haplotypes and their diploid content within an individual. In this study, the newly developed AlleleTyper™ Software was used to facilitate this analysis. The software aided the making of a biallelic translator from a monoallelic table that contained the genetic pattern information for the CYP gene variants tested. The software then matched the genetic information in SNP genotyping and copy number assay results files to the patterns in the biallelic translator (Figure 4) and reported the star allele genotypes for each individual (Figure 5). Example results are summarized below; information on the expected phenotype for the deduced genotypes is from the Pharmacogenomics Knowledgebase website (www.pharmgkb.org) [9].

# 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 consequently has an intermediate CYP2C19 metabolizer phenotype.
- Sample 29 carries a CYP2D6\*5 gene deletion allele and a wild type allele and so 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 therefore has a poor metabolizer phenotype for drugs metabolized by CYP2C19.
- Neither Sample 20 nor 29 carry the deleterious CYP2C9 alleles, CYP2C9\*2 (430C>T) and CYP2C9\*3 (1075A>C), which were found in four and three other samples, respectively.
- Five samples carry one or more CYP2D6\*36 alleles, which are classified as nonfunctional having no CYP2D6 function and which are often found in tandem with other CYP2D6 genes [7]. 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 ultrarapid metabolizer.

|           |                     |        |         | Summery Report Detailed Report |               |                   |              |  |
|-----------|---------------------|--------|---------|--------------------------------|---------------|-------------------|--------------|--|
|           |                     |        |         | CYP2D6_int2                    | CYP2D6_ex9    | 2D6 g.137_138insT | 2D6 g.883G>C |  |
| Sample ID | CYP2D6              | CYP2C9 | CYP2C19 | Hs04083572_cn                  | Hs00010001_cn | C_32407245_40     | C30634118_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,<br>*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 5:** Allele Typer Software showing the translated star allele diplotypes for CYP2D6, CYP2C9 and CYP2C19 gene variants for the 30 buccal swab samples. A portion of the genotype data for each sample is shown in this screenshot.

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). To this end, users can customize and order their own OpenArray\* DME and SNP genotyping assay panels. Life Technologies offers 2,700 TaqMan\* Drug Metabolism Assays designed to known and putative causal SNP, and InDel variants, as well as 4.5 million predesigned SNP assays and custom SNP assays for other targets of interest.

Pharmacogenomics studies typically use genotyping and copy number analysis to understand the impact of genetic variation on drug

metabolism, drug efficacy, and adverse drug effects. 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. We demonstrate a flexible and efficient workflow that generates a high success rate of genotyping calls and copy number results, and can be completed in a single day, from sample preparation to data analysis.

We also demonstrate that 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.

# Acknowledgements

We thank the members of the Life Technologies pharmacogenomics project team for their contributions and helpful discussions.

### References

- Robarge JD, Li L, Desta Z, Nguyen A, Flockhart DA (2007) The star-allele nomenclature: retooling for translational genomics. Clin Pharmacol Ther 82: 244-248
- Sim SC, Ingelman-Sundberg M (2010) The Human Cytochrome P450 (CYP) Allele Nomenclature website: a peer-reviewed database of CYP variants and their associated effects. Hum Genomics 4: 278–281.
- 3. Pharmacogenomics Experiments User Guide. P/N MAN0009612, Rev. A.
- Applied Biosystems\* QuantStudio™ 12K Flex Real-Time PCR System: OpenArray\* Experiments User Guide. P/N 4470935, Rev B.
- 5. TaqMan\* Copy Number Protocol. P/N 4397425, Rev D.
- 6. He Y, Hoskins JM, McLeod HL (2011) Copy number variants in pharmacogenetic genes. Trends Mol Med 17: 244-251.
- 7. 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.
- Ramamoorthy A, Flockhart DA, Hosono N, Kubo M, Nakamura Y, et al. (2010) Differential quantification of CYP2D6 gene copy number by four different quantitative real-time PCR assays. Pharmacogenet Genomics 20: 451-454.
- Michelle Whirl-Carrillo T, Hebert JM, Gong L, Owen R, Gong M, et al. (2008) The pharmacogenetics and pharmacogenomics knowledge base: accentuating the knowledge. Nucleic Acids Res 36: D913–D918.