Applied Biosystems

TaqMan® Drug Metabolism Genotyping Assays for Triallelic SNPs

Introduction

Applied Biosystems currently offers over 2,600 TaqMan® Drug Metabolism Genotyping Assays to detect biologically important polymorphisms in coding regions and regulatory elements of human drug metabolism enzyme (DME) and transporter genes. Single nucleotide polymorphisms (SNPs) or other mutations in these genes can affect drug efficacy, thus knowledge of their genotypes can be critical during pharmaceutical development and clinical trials. TaqMan® DME and SNP Genotyping Assays are designed to detect typical biallelic SNPs, yet an increasing number of SNPs of pharmaceutical and clinical value are known to be triallelic (i.e., three different nucleotide bases occur at the same genomic location in human populations). Here we describe the use of two paired, conventional TaqMan SNP Assays to detect all three alleles of a triallelic SNP and generate accurate sample genotypes.

Overview

Pairs of TaqMan DME Assays are now available to interrogate functionally important triallelic SNPs in the ABCB1 (MDR1) transporter[1,2] and in several Cytochrome P450 (CYP) genes[3] (Table 1). Each assay is a standard TaqMan SNP Assay containing allele-specific FAM™ and VIC® dye-labeled probes for the detection of just two SNP alleles. In both assays for a triallelic SNP, identical target-specific primer pairs are used to generate amplicons by polymerase chain reaction (PCR). Each assay contains one probe for the major SNP allele, which is labeled with the same reporter dye in both assays (e.g. VIC dye), and one probe for one of the minor alleles, which is labeled with the second reporter dye (e.g. FAM dye). To generate accurate sample genotypes, the two assays must be run independently on the same panel of samples, and the resulting allelic discrimination plots must be analyzed in concert, comparing the expected cluster positions from both assays to a map of the true sample genotypes (Table 2).

Here, assay validation data generated with the ABCB1 rs2032582 assays, C_11711720C_30 and C_11711720D_40, is used for genotype analysis of a triallelic SNP with paired SNP assays. Each allele of the rs2032582 SNP is found with relatively high frequency in different human populations (Table 3). It is critical that both assays for this SNP are run to generate correct sample genotypes. In contrast, the CYP gene SNP assays listed in Table 1 are designed to rare alleles (i.e. mutations) that are infrequently detected in general populations. There is a low risk that sample genotypes could be miscalled using a single assay due to the presence of the unreported minor allele. To accurately determine sample genotypes, it is highly recommended that paired assays are used to genotype the CYP gene triallelic SNPs.

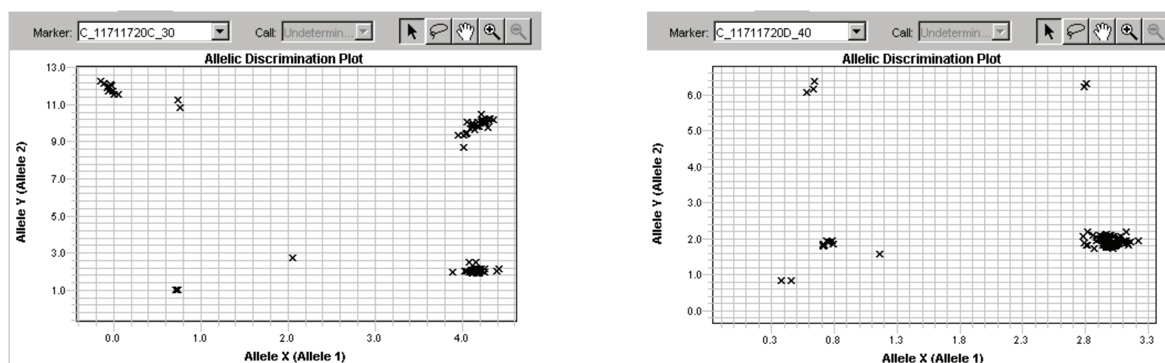


Figure 1. The allelic discrimination plots for the C_11711720C_30 (C/A) and C_11711720D_40 (C/T) assays are shown in SDS Software v2.3. Automatic allele calling has been turned off so that no genotypes are called.

Experimental Workflow

To interrogate a triallelic SNP, paired assays are run independently on the same genomic DNA samples, following the basic genotyping protocol for TaqMan DME Assays. After thermal cycling, plates are read on an Applied Biosystems Real-Time PCR System. Allelic discrimination plots are examined using either the Sequence Detection System (SDS) or AutoCaller™ Software (for the 7900HT Fast Real-Time PCR System only).

In the experiments shown in Figures 1–4, 3 ng DNA each for 45 African American and 45 Caucasian Coriell samples were dried down in 384-well plates. The 20X TaqMan Assays were individually mixed with 2X TaqMan® Universal PCR Master Mix (No AmpErase® UNG) and water to achieve a 1X concentration of all reagents, and 5 µL of this reaction mix was added to the wells. Plates were run on an Applied Biosystems GeneAmp® PCR System 9700 Thermal Cycler and an endpoint read was done on the 7900HT Fast Real-Time PCR System.

Assay Results

Preliminary genotype calls will usually need to be made manually to ensure that all samples are called appropriately and to facilitate the second stage of the data analysis (see below). Call all possible genotypes, but avoid calling samples that do not amplify well. For a given assay, the samples that fall into or outside of clusters can have the following properties:

- **Heterozygote clusters**—samples observed to fall within a heterozygote cluster are heterozygous for the two alleles reported by probes in the given assay.
- **Homozygote clusters**—samples found in or near a FAM or VIC homozygote cluster can be either a true homozygote for the reported allele, or can be a heterozygote for the FAM or VIC reported allele and the unreported allele (the allele without a probe in the assay). Samples carrying one or two copies of the reported allele can either cluster together or can run slightly differently

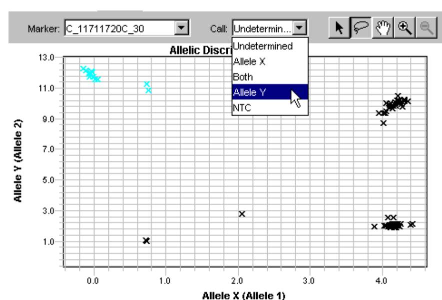
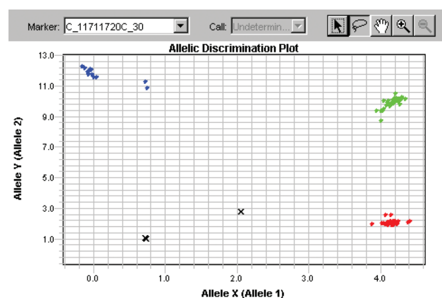


Figure 2. Genotypes of samples that amplify well are manually called and samples that do not amplify, or that amplify poorly, are left uncalled as 'undetermined'.



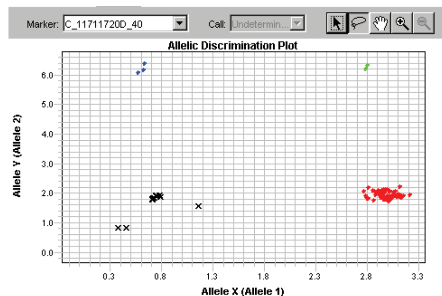
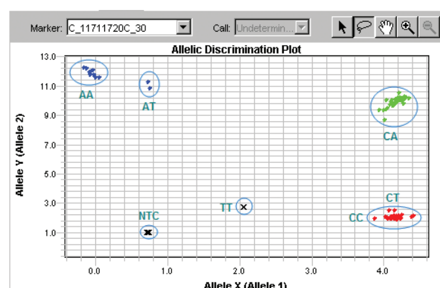


Figure 3. These panels show the completed manual assignment of samples to genotype clusters for the C_11711720C_30 (C/A) and C_11711720D_40 (C/T) assays.

Assay Name	Alleles Tested	rs ID#	Gene Symbol	Gene Name	Allele Nomenclature
C_11711720C_30	C/A	rs2032582	ABCB1	ATP-binding cassette, sub-family B (MDR/TAP), member 1	ABCB1, g.2677G>T, c.3095, A893S
C_11711720D_40	C/T	rs2032582	ABCB1	ATP-binding cassette, sub-family B (MDR/TAP), member 1	ABCB1, g.2677G>A, c.3095, A893T
C_30634117C_20	C/A	rs5030865	CYP2D6	cytochrome P450, family 2, subfamily D, polypeptide 6	CYP2D6*8, c.595G>T, g.1758G>T, G169X
C_30634117D_30	C/T	rs5030865	CYP2D6	cytochrome P450, family 2, subfamily D, polypeptide 6	CYP2D6*14, c.595G>A, g.1758G>A, G169R
C_72650009C_10	G/A	none	CYP2C8	cytochrome P450, family 2, subfamily C, polypeptide 8	CYP2C8*7, c.556C>T, g.4517C>T, R186X
C_72650009D_20	G/C	none	CYP2C8	cytochrome P450, family 2, subfamily C, polypeptide 8	CYP2C8*8, c.556C>G, g.4517 C>G, R186G
C__25625804_10	G/A	rs7900194	CYP2C9	cytochrome P450, family 2, subfamily C, polypeptide 9	CYP2C9*8,c.449G>A, g.3627G>A, R150H
C_25625804D_20	G/T	rs7900194	CYP2C9	cytochrome P450, family 2, subfamily C, polypeptide 9	CYP2C9*27,c.449G>T, g.3627G>T, R150L

Table 1. TaqMan® DME Assay pairs to detect triallelic coding SNPs in four DME genes. Note that the bases in the Alleles Tested and the Allele Nomenclature columns are the reverse complement of one another when the allele nomenclature references a cDNA or genomic location that maps to the reference genome (-) strand. TaqMan® SNP Assay alleles and context sequences are mapped to the (+) genome strand of the reference genome assembly (NCBI B36).



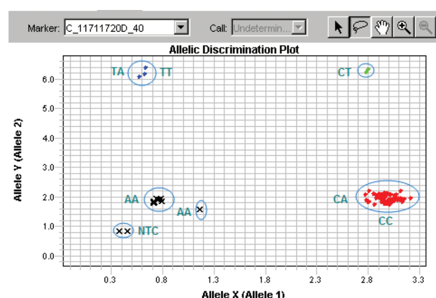


Figure 4. Shown are the final sample genotype assignments and their locations in the C_11711720C_30 (C/A) and C_11711720D_40 (C/T) assay allelic discrimination plots.

from one another and form split clusters (Figures 1–4). Manual comparison of the results of the two assays is required to determine the true sample genotype calls.

- **Undetermined**—weak or no amplification. Samples that are homozygous for the unreported allele in an assay can exhibit weak amplification, relative to samples carrying reported alleles, due to nonspecific activity of the assay probes for the unreported allele (Figures 1–4).

Stage 1 Data Analysis: Preliminary Sample Genotype Calls Per Assay

In SDS Software: call genotypes manually. From the Menu Bar 'Analysis' pull down menu, select 'Analysis Settings' and disable 'auto caller' by unchecking the 'auto caller enabled (SDS v2.3)' or 'automatic allele calling (SDS v1.4)' box.

In AutoCaller Software: when using the 7900HT Fast Real-Time PCR System and AutoCaller Software, import the .sds file and edit genotyping calls manually to call all amplified samples.

Use the software's manual calling feature to select and assign genotypes to samples that cluster in homozygote and heterozygote patterns (Figures 1–3). Assign the same genotype to all samples that fall within any clusters having a 'split' appearance (Figure 2). Do not call samples that do not amplify or that amplify relatively weakly.

Stage 2 Data Analysis: True Sample Genotype Calls by Comparison of Paired Assay Results

Use AutoCaller Software's export results table (SDS) or export table data function to export a tab delimited text file(s) containing the manual genotype calls for each assay. Open the file(s) in a spreadsheet program (e.g. Microsoft Excel®) and follow these steps:

1. For each assay, highlight and sort the data by the sample name column.
2. For each assay, sequentially copy the columns containing the sample names and manual calls and paste them side-by-side in a new spreadsheet. The data for both assays should be aligned for each sample in rows.
3. Highlight all cells and simultaneously sort on the genotype call columns for each assay. This will cluster samples with the same genotype by rows in the spreadsheet.
4. In an adjacent, new column, type in the predicted true sample genotypes based on the per assay manual calls, according to the chart in Table 2.

The final sample genotype calls and their location in the allelic discrimination plots for the C_11711720C_30 (C/A) and C_11711720D_40 (C/T) assays are shown in Figure 4. Note that the split FAM 'subclusters' seen in the C_11711720C_30 assay correlate with distinct genotypes, whereas the other FAM and VIC clusters seen with each assay are more homogeneous: samples carrying either one or two copies of the reported allele are indistinguishable. Also note that the unreported minor allele homozygous samples in each assay have a fair amount of signal due to mismatch hybridization, but this signal is significantly lower than that seen in samples where the reporter probes match the templates.

Table 4 shows the distribution of the genotypes and allele frequencies for ABCB1 rs2032582 in the Applied Biosystems assay validation population panels. These panels include the African American and Caucasian samples shown in Figures 1–4, as well as Japanese and Chinese samples (data plots not shown). These distributions are similar to those reported by the NIH SNP500 project[4] for the ABCB1 rs2032582 in related populations.

Summary

Genotyping triallelic SNPs with high accuracy and specificity is achievable using paired TaqMan SNP Genotyping Assays and the standard, simple TaqMan SNP Assay experimental workflow. The two-step manual data analysis is straightforward and easy to accomplish. This method enables accurate detection of functionally important triallelic SNPs, including critical DME gene polymorphisms, that were previously difficult to interrogate by standard SNP genotyping technologies designed for biallelic SNPs.

	C_11711720C_30	C_11711720D_40
Genotype	**C/A Assay Clusters**	**C/T Assay Clusters**
CC	Allele 1 (11)	Allele 1 (11)
CT	Allele 1 (11)	Both (12)
AT	Allele 2 (22)	Allele 2 (22)
AA	Allele 2 (22)	Undetermined (00)
CA	Both (12)	Allele 1 (11)
TT	Undetermined (00)	Allele 2 (22)
NTC	Undetermined (00)	Undetermined (00)

Table 2. Genotypes for samples run with the two assays for the ABCB1 rs2032582 triallelic SNP. In each assay, the major allele (C) is reported by the VIC® dye-labeled probe and one minor allele (A or T) in each assay is reported by the FAM™ dye-labeled probe. The cluster positions for each sample using each assay are mapped to the true sample genotypes.

References:

1. Gervasini G, Carrillo JA, Garcia M, San Jose C, Cabanillas A, Benitez J. (2006) Adenosine triphosphate-binding cassette B1 (ABCB1) (multidrug resistance 1) G2677T/A gene polymorphism is associated with high risk of lung cancer. *Cancer* **107:**2850–2857.

2. Brant SR, Panhuysen CI, Nicolae D, Reddy DM, Bonen DK, Karaliukas R, Zhang L, Swanson E, Datta LW, Moran T, Ravenhill G, Duerr RH, Achkar JP, Karban AS, Cho JH. (2003) MDR1 Ala893 polymorphism is associated with inflammatory bowel disease. *Am J Hum Genet.* **73:**1282–1292.

3. Sim SC, Ingelman-Sundberg M. (2006) The human cytochrome P450 Allele Nomenclature Committee Web site: submission criteria, procedures, and objectives. *Methods Mol Biol.* **320:**183–91.

4. Packer BR, Yeager M, Burdett L, Welch R, Beerman M, Qi L, Sicotte H, Staats B, Acharya M, Crenshaw A, Eckert A, Puri V, Gerhard D, Chanock SJ. (2006) SNP500Cancer: a public resource for sequence validation, assay development, and frequency analysis for genetic variation in candidate genes. *Nucleic Acids Res.* **34 Database issue:**D617–D621.

5. De La Vega FM, Dailey D, Ziegle J, Williams J, Madden D, Gilbert DA. (2002) New generation pharmacogenomic tools: a SNP linkage disequilibrium map, validated SNP assay resource, and high-throughput instrumentation system for large-scale genetic studies. *Biotechniques.* **Suppl:**48–50, 52, 54.

Genotype	AfAm	Cauc	Chin	Japn
CC	0.84	0.18	0.24	0.13
CT	0	0.04	0.16	0.13
TT	0	0.04	0	0
CA	0.16	0.49	0.42	0.47
AT	0	0.04	0.07	0.11
AA	0	0.22	0.11	0.16

Allele	AfAm	Cauc	Chin	Japn
C	0.92	0.44	0.53	0.43
T	0	0.07	0.11	0.12
A	0.08	0.49	0.36	0.44

Table 3. ABCB1 rs2032582 triallelic SNP genotype and allele distributions in Applied Biosystems assay validation population panels. Panels consist of 45 unrelated gDNA samples from African American (AfAm), Caucasian (Cauc), Chinese (Chin) and Japanese (Japn) populations[5]. Genotype and allele distributions are provided as the fraction of the 45 total sample genotypes or the 90 total alleles tested in each population panel.

For Research Use Only. Not for use in diagnostic procedures.

Practice of the patented 5' Nuclease Process requires a license from Applied Biosystems Inc. The purchase of TaqMan® Drug Metabolism Genotyping Assays or TaqMan® probes includes an immunity from suit under patents specified in the product insert to use only the amount purchased for the purchaser's own internal research when used with the separate purchase of an Authorized 5' Nuclease Core Kit. No other patent rights are conveyed expressly, by implication, or by estoppel. For further information on purchasing licenses contact the Director of Licensing, Applied Biosystems Inc., 850 Lincoln Centre Drive, Foster City, California 94404, USA.

© Copyright 2008, Applied Biosystems. All rights reserved. Applied Biosystems, GeneAmp, and VIC are registered trademarks and AutoCaller and FAM are trademarks of Applied Biosystems Inc. or its subsidiaries in the U.S. and/or certain other countries. AmpErase and TaqMan are registered trademarks of Roche Molecular Systems, Inc. All other trademarks are the sole property of their respective owners.

Printed in the USA, 07/2008 Publication 135AP01-01

Headquarters
850 Lincoln Centre Drive | Foster City, CA 94404 USA
Phone 650.638.5800 | Toll Free 800.345.5224
www.appliedbiosystems.com

International Sales
For our office locations please call the division headquarters or refer to our Web site at
www.appliedbiosystems.com/about/offices.cfm