Detection of CYP2D6/CYP2D7 hybrid genes using the PharmacoPro™ Array for pharmacogenomics research

Carsten Bruckner, Ram Varma, Neha Garg, and Jeremy Gollub. Thermo Fisher Scientific, 3450 Central Expy, Santa Clara, CA, USA, 95051

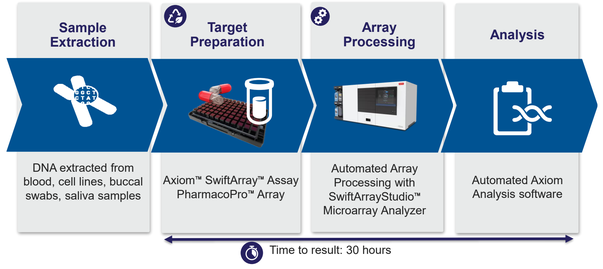
Abstract

CYP2D6 is an important enzyme responsible for drug metabolism in humans, but its activity varies significantly between individuals due to common single nucleotide polymorphisms (SNPs) and structural variations. The gene locus on chromosome 22 includes two highly homologous pseudogenes, CYP2D7 and CYP2D8. Variant CYP2D6 alleles include gene conversions in which CYP2D6 sequence has been replaced by CYP2D7 sequence. Such hybrid genes typically result in loss of enzyme function, and their detection is therefore important for star allele typing in pharmacogenomics.

The Applied Biosystems™ Axiom™ PharmacoPro™ Array enables CYP2D6 hybrid gene detection using around 300 probes interrogating exonic, intronic, and flanking regions of both CYP2D6 and CYP2D7. Signal intensities of these probes are used to estimate copy number (CN) in defined genomic regions. CN changes in these regions are used to infer gene conversion events and enable hybrid gene detection.

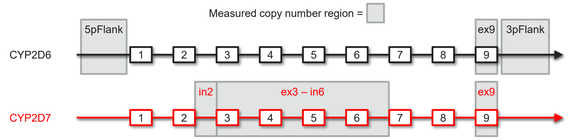
Cell line genomic DNA samples (>850) from diverse populations were assayed on the PharmacoPro array. Copy number events such as deletions, amplifications, and gene conversions were detected with high reproducibility on CYP2D6 and CYP2D7. Results were compared with publicly available sequencing-based star allele genotyping of these samples [1,2]. CYP2D6 whole gene deletions and amplifications were detected with 97.4% concordance to reference star allele genotyping. Both CYP2D7::CYP2D6 and CYP2D6::CYP2D7 hybrids were detected with high accuracy and reproducibility. CYP2D7::CYP2D6 hybrids like *13, in which the 5′ end of CYP2D6 including exon 1 has been replaced with CYP2D7, were detected in five samples from the 1000 Genomes Project. CYP2D6::CYP2D7 hybrids like *36, *68 and *4.013, in which the 3′ end of CYP2D6 including exon 9 has been replaced with CYP2D7, were detected in 182 samples from the 1000 Genomes Project. Furthermore, the system could distinguish between these common CYP2D6::CYP2D7 star alleles based on the extent of the gene conversion event.

Materials and methods



The full workflow from gDNA to report generation is shown above

- Genomic DNA (gDNA) from human cell line, blood, buccal and saliva samples is isothermally amplified using Axiom™ SwiftArray™ Assay.
- Key markers in pseudogenes like CYP2D6 are amplified with the mPCR module in the SwiftArray Assay Reagent Kit.
- Fragments are precipitated, resuspended, & hybridized to the PharmacoPro array.
- Bound target is washed automatically under stringent conditions to remove non-specific background on the SwiftArrayStudio™ Microarray Analyzer.
- After ligation, the arrays are stained and imaged and the data is transferred to Automated Axiom Analysis software for data analysis.



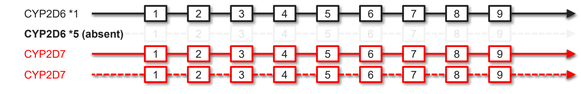
Schematic of CYP2D6 gene (in black) and CYP2D7 (in red), with their 9 exons. Total CN state is measured in three regions for each gene, as indicated by gray boxes. PharmacoPro performs targeted amplification for most of CYP2D6 to achieve high SNP calling accuracy. Targeted amplification precludes CN measurement in CYP2D6 exons 1-8.

Results

Relationship between Coriell annotated star alleles and PharmacoPro-measured CN events

(a) CYP2D6 *5 deletion event

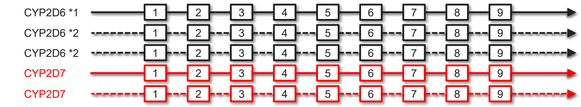
Measured CN State	5pFlank		intron2	exon3-intron6		exon9	3pFlank
CYP2D6	1					1	1
CYP2D7			2	2		2	



Each row of the simplified schematic of CYP2D6 and CYP2D7 represents one copy of the gene. All three CYP2D6 CN regions report CN_State = 1 for 1000Genomes sample HG01890, consistent with a whole-gene deletion event (*5 allele). All three CYP2D7 CN regions report a CN_State = 2. All star allele calls in this section are from Coriell Star Allele Search [1,2] annotated star alleles for associated sample.

(b) CYP2D6 *2x2 duplication event

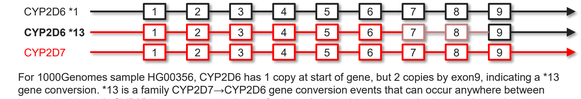
Measured CN State	5pFlank		intron2	exon3-intron6		exon9	3pFlank
CYP2D6	3					3	3
CYP2D7			2	2		2	



All three CYP2D6 CN regions report CN_State = 3 for 1000Genomes samples NA18857, consistent with a whole-gene duplication of the *2 allele. All three CYP2D7 CN regions report a CN_State = 2.

(c) CYP2D7::CYP2D6 hybrid allele *13

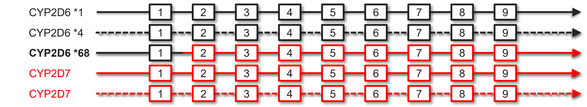
Measured CN State	5pFlank		intron2	exon3-intron6		exon9	3pFlank
CYP2D6	1					2	2
CYP2D7			2	2		1	



For 1000Genomes sample HG00356, CYP2D6 has 1 copy at start of gene, but 2 copies by exon9, indicating a *13 gene conversion. *13 is a family CYP2D7→CYP2D6 gene conversion events that can occur anywhere between intron1 and intron8. CYP2D7 reports a second copy for intron2–intron6 but not exon9, also consistent that extra copy of CYP2D7 sequence being present in CYP2D6 *13 allele.

(d) CYP2D6::CYP2D7 hybrid allele *68

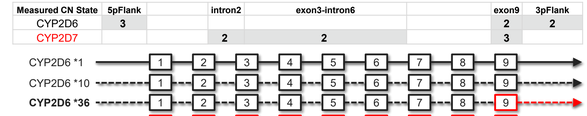
Measured CN State	5pFlank		intron2	exon3-intron6		exon9	3pFlank
CYP2D6	3					2	2
CYP2D7			3	3		3	



For 1000Genomes sample NA10855, CYP2D6 reports 3 copies at start of gene, but only 2 copies at exon9, indicating a CYP2D6→CYP2D7 gene conversion event like *68. *68 switches to CYP2D7 sequence in intron1. CYP2D7 reports an extra copy in each measured region, also consistent that extra copy of CYP2D7 sequence being present in CYP2D6 *68 allele.

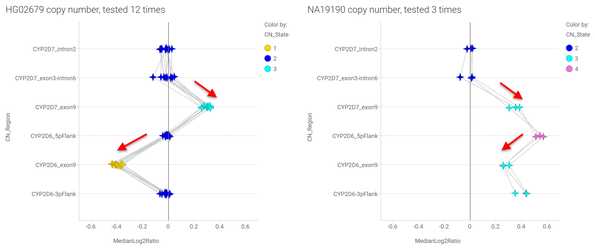
(e) CYP2D6::CYP2D7 hybrid allele *36

Measured CN State	5pFlank		intron2	exon3-intron6		exon9	3pFlank
CYP2D6	3					2	2
CYP2D7			2	2		3	



For 1000Genomes sample NA18956, CYP2D6 reports 3 copies at start of gene, but only 2 copies at exon9, consistent with a CYP2D6→CYP2D7 gene conversion event like *36. *36 switches to CYP2D7 sequence at or within exon9. CYP2D7 reports a duplication of exon9, consistent with an extra copy of CYP2D7 exon9 sequence in the CYP2D6*36 allele. Measuring CYP2D7 regions helps discriminate among various possible CYP2D6 hybrid alleles.

Reported CYP2D6 and CYP2D7 CN states for 2 discordant samples



HG02679 is annotated by Coriell star allele search as CYP2D6 *10/*40, but exon9 CN changes in CYP2D6 and CYP2D7 vs upstream indicate sample is probably *36/*40.

NA19190 is annotated by Coriell star allele search as CYP2D6 *1x3/*10, but exon9 CN changes in CYP2D6 and CYP2D7 vs upstream indicate sample is probably *1x3/*36.

CNV calling performance by CYP2D6 structural variation group

Sample group	Example Coriell CYP2D6 Star Allele call	Expected copy number by CYP2D6 allele group: 2D6 CN	2D6::2D7 CN	2D7::2D6 CN	No. tested 1000 Genomes samples (distinct)	Concordance of gene and hybrid gene CN prediction
1	*5/*5	0	0	0	2 (NA19317)	100%
2	*1/*5	1	0	0	231 (68)	100%
3	*1/*1	2	0	0	1661 (548)	98.9%
4	*1/*1x2	3	0	0	148 (43)	95.3%
5	*4x2/*4x2	4	0	0	6 (3)	50%
6	*36/*36	0	2	0	3 (NA20287)	100%
7	*5/*36+*10	1	1	0	19 (12)	84%
8	*1/*36+*10	2	1	0	352 (132)	98.6%
9	*36+*10/*36+*10	2	2	0	62 (33)	97%
10	*36+*10/*36+*36+*10	2	3	0	13 (3)	92%
11	*1+*90/*36+*10	3	1	0	2 (NA18642)	100%
12	*1/*13	1	0	1	2 (HG00356)	100%
13	*13+*2/*1	2	0	1	5 (3)	100%
14	*13/*68+*4	1	1	1	8 (NA19982)	0%
					overall concordance	98.1%

CYP2D6_5pFlank and CYP2D6_exon9 measured copy number is compared to reference Coriell annotated star allele calls, and comparison is concordant if CN calls are consistent with annotated star allele call. Some discordances are due to likely errors in reference star allele calls.

Conclusions

These highly accurate results demonstrate the ability to infer complex structural variants from copy number changes detectable using a hybridization-based assay. Structural variant detection and SNP genotyping in one assay on the PharmacoPro array can be used to efficiently and precisely call star alleles for pharmacogenomics research applications.

References

1. https://www.coriell.org/StarAllele/search
2. Gharani et al. Star allele search: a pharmacogenetic annotation database and user-friendly search tool of publicly available 1000 Genomes Project biospecimens. *BMC Genomics* (2024) 25:116 https://doi.org/10.1186/s12864-024-09994-6

Trademarks/licensing/labels

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

Science at a scan
Scan the QR code to learn more about our genetic technologies and offerings.

Visit us at **booth 736**

Learn more at thermofisher.com/ashg2025