CRISPR enrichment of *CYP2D6-2D7-2D8* loci allows for phased haplotyping and accurate detection of complex *CYP2D6* structural variation

Amy J Turner[1], Ashley D Derezinski[1], Andrea Gaedigk[2], Mark E Berres[3], David B Gregornik[4], Ulrich Broeckel[1] and Gunter Scharer[1]

[1]RPRD Diagnostics LLC, Milwaukee, WI; [2]Children's Mercy Research Institute, Kansas City, MO; [3]UWBC, University of Wisconsin, Madison, WI; [4]Children's Minnesota, Minneapolis, MN

Background

- The highly polymorphic *CYP2D6* gene is one of the most commonly clinically tested pharmacogenes contributing to the metabolism of >20% of common clinical drugs[1].
- A large portion of enzyme variability can be explained by single nucleotide polymorphisms (SNPs), as well as structural variants (SVs) such as copy number variation (CNV) and hybrid rearrangements with the highly similar *CYP2D7* pseudogene. The latter are known to occur in different configurations and frequencies across populations[1-3].
- Clinically relevant complex SV/CNVs, especially those with *CYP2D6* and *CYP2D7*-derived hybrid sequences and/or multiple gene copies, are notoriously difficult to accurately detect and characterize which may lead to incorrect activity assignments. This may impact drug dosing recommendations, particularly in underrepresented populations[1-10].

Methods

- We developed a method using *CYP2D6-CYP2D7-CYP2D8* locus targeted guide RNAs (gRNAs) along with CRISPR-Cas9 technology and PCR-free single molecule long-read nanopore sequencing to accurately determine diplotypes (phased haplotypes) regardless of SV/CNV complexity.
- We evaluated the approach using previously characterized samples with a variety of SV/CNVs including complex *CYP2D6* CNVs and *CYP2D6::CYP2D7* hybrids. High molecular weight (HMW) DNA obtained from multiple sample types including lymphoblastoid cell lines (LCL), blood, saliva, and liver tissue were all evaluated.
- PCR-free single molecule libraries were prepared using targeted gRNAs and Cas9 Sequencing Kit, then sequenced using Oxford Nanopore MinION flow cells.
- Using standard and custom-build data analysis programs, data was base-called, filtered, aligned (standard and custom references), sorted, indexed and coverage assessed and visualized over a ~52Kb region covering *CYP2D6, CYP2D7*, and *CYP2D8*.

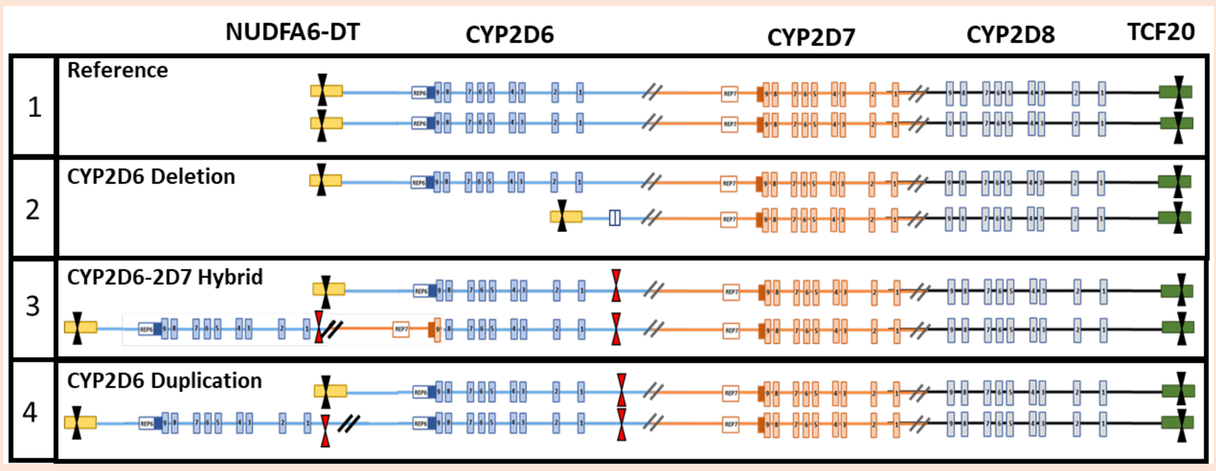

Figure 1. Visualization of target cut sites. Target locations for designed gRNAs to capture the full *CYP2D6-2D7-2D8* loci for sequencing in continuous long reads. (1) reference *CYP2D6-2D7-2D8* loci, (2) *CYP2D6*5* full gene deletion, (3) *CYP2D6*36+*10* arrangement containing a full *CYP2D6* copy plus a hybrid gene, and (4) *CYP2D6* duplication of two identical gene copies.

Results

Figure 2. Visualization of *CYP2D6-2D7-2D8 Loci* Structural Variation. (A) Custom reference track for *CYP2D6::CYP2D7* hybrid duplication alignment. (A1) Sample without SV/CNV in target region. (A2) Sample with one allele without SV/CNV and one allele with a tandem rearrangement. (A3) Sample with one allele without SV/CNV and one allele with a hybrid and tandem rearrangement. (B). Custom reference track for *CYP2D6* gene duplication alignment. (B1) Sample without SV/CNV. (B4) Sample with one allele without SV/CNV and one allele with gene deletion. (B5) Sample with one allele without SV/CNV and one allele with hybrid gene with a tandem arrangement. (B6) Sample with one allele without SV/CNV and one allele with a 2-copy gene duplication[11].

Conclusions

- We enriched and analyzed the entire *CYP2D6-2D7-2D8* loci using PCR-free single molecule long-read sequencing from HMW DNA extracted from LCL, blood, saliva, and liver tissue.
- This novel methodology accurately determined, with 100% concordance, the SNP level genotypes and complex diplotypes with SVs such as *CYP2D6*2/*36+*10*, unconstrained by amplicon size or sample structural complexity.
- Additionally we identified 3 novel *CYP2D6* suballeles **1.058, *1.059, and *2.033.*
- This approach yields fully phased diplotypes across the entire loci. In addition, SV breakpoints can be defined which offers unique advantages overcoming the limitations of current clinical testing platforms.

References

1. Ayano G (2016) Psychotropic Medications Metabolized by Cytochromes P450 (CYP) 2D6 Enzyme and Relevant Drug Interactions. Clin Pharmacol Biopharm 5: 162.
2. Saravanakumar A, Sadighi A, Ryu R, Akhlaghi F. Physicochemical Properties, Biotransformation, and Transport Pathways of Established and Newly Approved Medications: A Systematic Review of the Top 200 Most Prescribed Drugs vs. the FDA-Approved Drugs Between 2005 and 2016. Clin Pharmacokinet. 2019;58(10):1281-1294. 3. Scantamburlo G, Tziolia K, Zopf M, Bernardinelli E, Soyal SM, Civello DA, Vanoni S, Dossena S, Patsch W, Patrinos GP, Paulmichl M, Nofziger C. Allele Drop Out Conferred by a Frequent CYP2D6 Genetic Variation For Commonly Used CYP2D6*3 Genotyping Assays. Cell Physiol Biochem. 2017;43(6):2297-2309.
4. Gaedigk A, Sangkuhl K, Whirl-Carrillo M, et al. The Evolution of PharmVar. Clin Pharmacol Ther. 2019;105(1):29-32. doi:10.1002/cpt. 1275
5. 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. 2018;103(3):399-401.
6. Nofziger C, Turner AJ, Sangkuhl K, et al. PharmVar GeneFocus: CYP2D6. Clin Pharmacol Ther. 2020;107(1):154-170.
7. Steijins LSW, Van Der Weide J. Ultrarapid drug metabolism: PCR-based detection of CYP2D6 gene duplication, Clinical Chemistry, Volume 44, Issue 5,1 May 1998, Pages 914-917.
8. Gaedigk A, Blum M, Gaedigk R, Eichelbaum M, Meyer UA. Deletion of the entire cytochrome P450 CYP2D6 gene as a cause of impaired drug metabolism in poor metabolizes of the debrisoquine/spartina polymorphism. Am J Hum Genet. 1991;48(5):943-950.
9. Gaedigk A. Complexities of CYP2D6 gene analysis and interpretation. Int Rev Psychiatry. 2013;25(5):534-553.
10. Del Tredici AL, Malhotra A, Dedek M, Espin F, Roach D, Zhu GD, Voland J, Moreno TA. Frequency of CYP2D6 Alleles Including Structural Variants in the United States. Front Pharmacol. 2018 Apr 5;9:305.
11. Turner A, Derezinski A, Gaedigk A, Berres ME, Gregornik DB, Brown K, Scharer G and Broeckel U. Characterization of Complex Structural Variation in the CYP2D6-CYP2D7-CYP2D8 Gene Loci Using Single-molecule Long-read Sequencing. (https://www.frontiersin.org/articles/10.3389/fphar.2023.1195778/abstract)