A Cost-effective Workflow for Massively Parallel Sequencing of Drug Metabolizing Enzymes

Ryan Gutierrez, BS*; Kari Graham, BA; Kyleen Elwick, BS; Carrie Mayes, BS; Bobby LaRue, Ph.D

Department of Forensic Science, Sam Houston State University, Huntsville, TX 77340

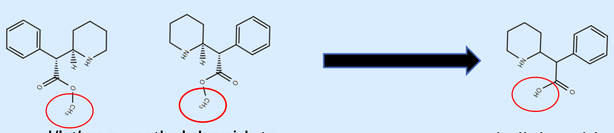
MEMBER THE TEXAS STATE UNIVERSITY SYSTEM

ABSTRACT

Mutations in drug metabolizing enzymes can lead to varying responses to similar doses of pharmaceutical compounds. These mutations can have serious implications ranging from requiring an adjustment of drug regimens to adverse reactions that otherwise would be unanticipated by health care practitioners. In some instances, these adverse reactions can lead to temporary disorientation or even death. In forensic settings, this could play a role in molecular autopsies. In this study a robust method was developed for the use of massively parallel sequencing to identify polymorphisms in drug metabolizing enzymes.

INTRODUCTION

Attention-Deficit/Hyperactivity Disorder (ADHD) is a common diagnosis that affects an estimated 5% of the world's population[1]. Methylphenidate (MPH), a frequently used psychostimulant in the United States, is administered orally as a racemic mixture of d and l threo-enantiomers to treat ADHD. A variable response rate among those who take methylphenidate has been reported in literature with increases in adverse side effects leading to a decreased quality of life in some patients [2, 3]. Methylphenidate is metabolized by Carboxylesterase enzyme 1 (CES1). This enzyme is encoded by the *CES1* gene. This 30.5 kilobase gene is located on the q arm of chromosome 16 and recent studies have found multiple SNPs and CNVs that have been correlated to outcome of MPH dose. To increase the likelihood of identifying sequencing variants that impacted MPH metabolism the entire *CES1* gene was sequenced using massively parallel sequencing. Target enrichment was accomplished using long amplicon polymerase chain reaction, with two approximately equal length overlapping amplicons being amplified. This target enrichment has been used successfully with Sanger Sequencing and with the mitochondrial genome [2, 4]. Target enrichment was optimized and resulting variants were sequenced using the Illumina MiSeq FGx.



d/l-*threo*-methylphenidate ritalinic acid

MATERIALS AND METHODS

Collection
- Saliva - DNA OG-500® Saliva Kits (Oragene, Ottawa, Ontario, Canada) N=173

Extraction
- QIAcube® (Qiagen, Hilden, Germany) with QIAamp DNA Mini Kit®

Quantification
- Qubit 2.0™ fluorometer (Invitrogen, Waltham, MA, USA)

Enrichment

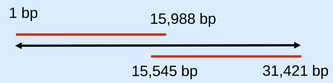


Figure 1: Amplicon overlay for *CES1*

- TaKaRa Long Amplicon Polymerase Chain Reaction®
- Custom Primer Design with Primer-BLAST (NCBI)
- PCR Optimized for *CES1:* 2 primer sets tested, annealing temperature, extension time, cycle number

Table 1: Optimized PCR Conditions

Initial Denaturation	94°C for 2 minutes
Denaturation	98°C for 10 seconds
Annealing	60°C for 1 minute (32x)
Extension amp 1	68°C for 20 minutes (32x)
Extension amp 2	68°C for 18 minutes (32x)
Final Extension	72°C for 10 minutes
Hold	4°C

Library Preparation and Sequencing
- MiSeq FGx with Nextera XT library Prep (Illumina, San Diego, California) resequencing- targeted amplicon workflow
- MiSeq Reagent Kit v3 600 cycle chemistry, 86 samples per run

Analysis
- Miseq Control Software, Real time analysis, MiSeq Reporter
- Generated Bam/Bai files visually compared to VCF files using Integrated Genomics Viewer
- Results were then compared to Variation Viewer (NCBI)

RESULTS & DISCUSSION

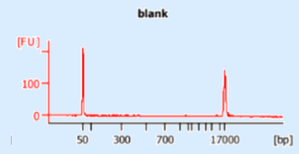


Figure 2

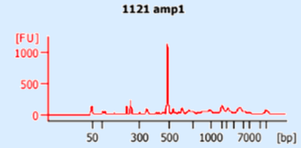


Figure 3

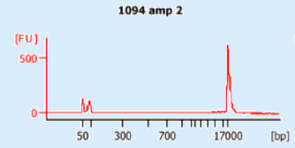


Figure 4

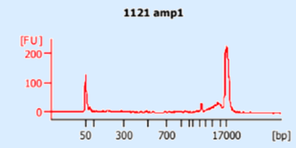


Figure 5

Figures 2-5: Bioanalyzer 2100 results with Figure 3 displaying an initial attempt at enrichment. Figures 4 and 5 display successful enrichment post optimization

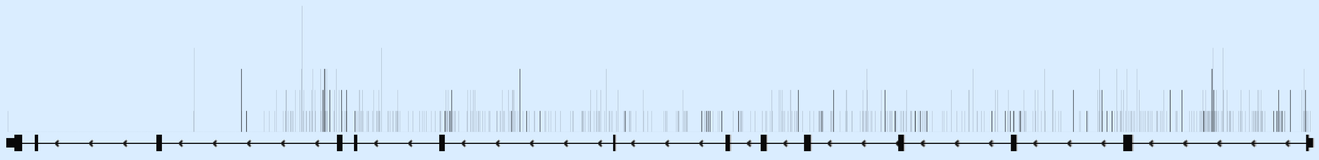


Figure 6: Variants sequenced in the sample population, grouped to the nearest 10 bp for clarity

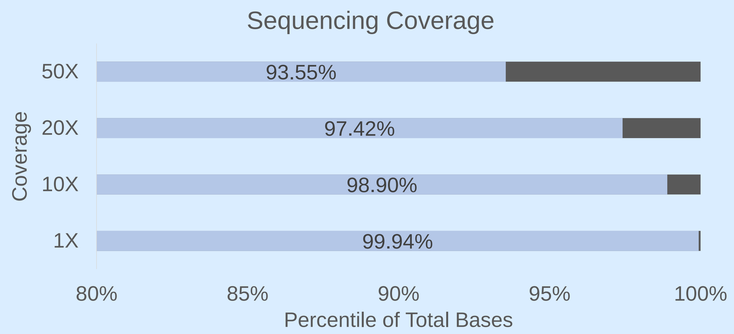
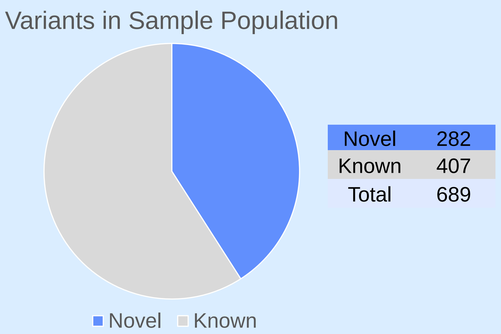


Figure 7: Sequencing Coverage of all successfully amplified samples (n=171)



Novel	282
Known	407
Total	689

Figure 8: Summary of sequenced variants in sample population as compared to Variation Viewer (NCBI)

REFERENCES and ACKNOWLEDGEMENTS

The Authors would like to acknowledge Essentia Healthcare for funding as well as collection of saliva samples for sequencing. The authors would also like to acknowledge the Department of Forensic Science for additional financial support of this research.

[1] G. Polanczyk, M.S. de Lima, B.L. Horta, J. Biederman, L.A. Rohde, The Worldwide Prevalence of ADHD: A Systematic Review and Metaregression Analysis, American Journal of Psychiatry 164(6) (2007) 942-948.
[2] H.-J. Zhu, K.S. Patrick, H.-J. Yuan, J.-S. Wang, J.L. Donovan, C L. DeVane, R. Malcolm, J.A. Johnson, G.L. Youngblood, D.H. Sweet, T.Y. Langaee, J.S. Markowitz, Two CES1 Gene Mutations Lead to Dysfunctional Carboxylesterase 1 Activity in Man: Clinical Significance and Molecular Basis, American Journal of Human Genetics 82(6) (2008) 1241-1248.
[3] C. Stage, G. Jürgens, L.S. Guski, R. Thomsen, D. Bjerre, L. Ferrero-Miliani, Y.K. Lyauk, H.B. Rasmussen, K. Dalhoff, The impact of CES1 genotypes on the pharmacokinetics of methylphenidate in healthy Danish subjects, British Journal of Clinical Pharmacology 83(7) (2017) 1506-1514.
[4] J.L. King, B.L. LaRue, N.M. Novroski, M. Stoljarova, S.B. Seo, X. Zeng, D.H. Warshauer, C.P. Davis, W. Parson, A. Sajantila, B. Budowle, High-quality and high-throughput massively parallel sequencing of the human mitochondrial genome using the Illumina MiSeq, Forensic Science International: Genetics 12 (2014) 128-135.
[6] Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden T (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction.BMC Bioinformatics. 13:134.
[7] Helga Thorvaldsdóttir, James T. Robinson, Jill P. Mesirov. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in Bioinformatics 14, 178-192 (2013).
[8] Variation Viewer. NCBI. https://www.ncbi.nlm.nih.gov/variation/view/overview/

CONCLUSIONS

- This workflow was a cost-effective and successful method for method for sequencing *CES1*
- Preliminary results show multiple novel variants in sample population
- Additional tertiary analysis is needed to correlate sequenced SNPs to clinical outcomes
- Alignment problems because of highly homologous pseudogenes may effect tertiary analysis
- Conserved drops in coverage resulted from the use of LAPCR