Characterisation of *CYP2D6*, *CYP2B6*, and *CYP2A6* pharmacogenetic variation in African populations and development of the StellarPGx diplotype calling algorithm.

View session detail

Author Block: **D. Twesigomwe**[1,2], Z. Lombard[2], S. Hazelhurst[1,3], the Wits-H3Africa/GSK ADME Collaboration; [1]Sydney Brenner Inst. for Molecular BioSci., Faculty of Hlth.Sci., Univ. of the Witwatersrand, Johannesburg, South Africa, [2]Div. of Human Genetics, Natl. Hlth.Lab. Service and Sch. of Pathology, Faculty of Hlth.Sci., Univ. of the Witwatersrand, Johannesburg, South Africa, [3]Sch. of Electrical and Information Engineering, Univ. of the Witwatersrand, Johannesburg, South Africa

Background and aims: Genetic variation is in part responsible for the variability in drug response across populations. However, the full catalog of pharmacogenetic variants, and their distribution, are yet to be established, in particular for African populations. This study therefore aimed to characterize variation in three core drug metabolism pharmacogenes, namely *CYP2D6*, *CYP2B6,* and *CYP2A6*, across diverse and/or previously underrepresented African populations. **Methods**: Given the challenges of diplotyping these hypervariable genes, we developed a novel bioinformatics pipeline (StellarPGx) to facilitate star allele (haplotype) detection from short-read whole genome sequence (WGS) data. This was followed by benchmarking based on simulated data and 109 real-world WGS datasets from participants with gold-standard diplotypes reported by the CDC's Genetic Testing Reference Materials Coordination Program. We then used StellarPGx and other existing tools to call *CYP2D6*, *CYP2B6*, and *CYP2A6* star alleles from high-coverage WGS data from 961 African participants and over 2000 genomes (for comparison) from global populations. **Results**: In our benchmarking analysis, StellarPGx (99%) and Cyrius (98%) had the highest concordance to ground truth *CYP2D6* diplotypes. From the genomic mining, we present frequencies across sub-Saharan Africa (SSA) for star alleles in *CYP2D6* (e.g. **17*, **29*, and **5*—frequency of 20%, 10%, and 8%, respectively), *CYP2B6* (e.g. **6* and **18*—frequency of 33% and 10%, respectively) and *CYP2A6* (e.g. **4*, **9*, and **46*—frequency of 3%, 8%, and 6% respectively), compared to other global populations. StellarPGx identified novel African-ancestry star alleles in *CYP2D6* (n=27), *CYP2B6* (n=18), and *CYP2A6* (n=31); seven of these alleles were validated via targeted Single-Molecule Real-Time (SMRT) resequencing. In addition, collaboration with experts from the Pharmacogene Variation Consortium (PharmVar) has enabled the validation of a further 12 novel *CYP2D6* star alleles. Our phenotype predictions for *CYP2D6* and *CYP2B6* indicate that the landscape of drug metabolizer phenotypes is non-uniform across SSA and differs to a large extent from phenotype distributions in global populations. **Conclusion**: Our findings underscore the need for investigating pharmacogene variation in the African context to reliably inform clinical pharmacogenomics implementation in Africa and the African diaspora. One African population cannot be used as a proxy for the whole continent. We recommend using StellarPGx for high-coverage WGS-based identification of known and novel pharmacogene haplotypes across all global populations.