

HiFi Long Read Genome Sequencing After Negative Clinical Exome Sequencing In a Consecutive Cohort Increases Diagnostic Yields

Sara E. Kalla¹, Stuti Joshi¹, Kenny Chen¹, Arthur Bandeira de Mello Garcia¹, Greta Pitsava¹, Krista Bluske¹, Nida Karra¹, Sarah L. Stenton¹, Sami Belhadj¹, Rachid Karam¹, Seth I. Berger¹

Clinical workflows for genetic diagnosis of rare disease are becoming more comprehensive supported by multiple professional guidelines which recommend exomes and genomes as first line testing for a wide range of indications. Current clinical testing is mostly based around short-read sequencing, but there is significant emerging evidence about increased diagnostic yields from newer long-read sequencing (LRS) approaches. Often, these studies are biased by selection criteria enriching for cases with high pretest probabilities of a missing diagnosis. To assess this increased yield in an unselected cohort, we used long-read HiFi whole genome sequencing in a consecutive cohort with negative exomes from our large clinical reference laboratory.

Consecutive cases that completed Ambry Genetics ExomeReveal Testing (Exome with supplementary RNA analysis) were selected if the clinical exome test was reported as negative and sufficient residual sample remained for LRS. Deidentified blood samples underwent high molecular weight DNA isolation, library preparation, and were sequenced on a PacBio Revio sequencer to target 20x whole genome coverage. Data was processed through the PacBio WDL v3 and interpreted on geneeyx.

Among 35 cases analyzed, clinically relevant findings were identified in 4 cases. Known DMPK repeat expansions were identified in a pair of siblings who underwent exome testing due developmental differences disproportionate to the myotonic dystrophy diagnosis. The long-read data also identified an X-Y translocation with unbalancing leading to duplication overlapping a Xq28 triplosensitivity region which was classified as a variant of uncertain significance. In another case, a pathogenic mitochondrial variant was identified by LRS and was confirmed to have been identified by mitochondrial genome analysis after case selection. Lastly, a deep intronic variant in *FBN1* with predicted pseudoexon inclusion was identified in a case with clinical Marfan's syndrome and classified as likely pathogenic.

In exome-negative cases, reflex long-read genome analysis approached 10% additional yield of clinically relevant findings, some of which could be identified by short-read genomes. Further work is needed to integrate more features of long-read data, such as haplotype resolution, phasing, and methylation, to fully assess the full impact of including this technology systematically into the clinical evaluation for rare diseases.