Title: Long Read Genome Sequencing for Cerebellar Ataxia

Authors: Niki Maki, Kirsten Blanco, Rebekah Barrick, Erica Smith, Ivan de Dios, Seth Berger, Emmanuèle Délot, Krista Bluske, Greta Pitsava, Soma Das, Christopher M. Gomez, Eric Vilain, Changrui Xiao

Objective: To assess the diagnostic yield of long-read genome sequencing (LR-GS) in ataxia patients with prior non-diagnostic genetic testing.

Background: Hereditary cerebellar ataxias are genetically diverse disorders characterized by progressive cerebellar atrophy. Despite advances in conventional methods of genetic testing, such as exome sequencing, repeat-expansion testing, and targeted gene panels, many ataxia cases remain genetically undiagnosed. The diagnostic yield of next-generation sequencing and repeat-expansion panels remains limited, and multiple tests are often needed. LR-GS can detect structural variants, including tandem repeats, more robustly than short-read approaches and may consolidate testing and increase diagnostic yield.

Design/Methods: We attempted to perform LR-GS on 100 previously extracted DNA samples from patients with negative prior clinical testing for cerebellar ataxia. 59 samples had adequate quality for LR-GS. Inclusion was not limited by prior test type. We used the standard PacBio pipeline for variant calling and annotation.

Results: All 59 patients had negative ataxia repeat-expansion analysis, and a subset had prior exome testing. 6 yielded a diagnostic or likely diagnostic finding. One candidate variant was identified that require further validation. In one case, LR-GS assisted in the interpretation of two previously identified variants in SYNE1 by establishing their phase as in trans.

Conclusions: LR-GS can accurately identify disease causing variants in ataxia. While high-molecular-weight DNA extraction is recommended for LR-WGS, over half of previously extracted DNA samples were of sufficient quality, suggesting feasibility of reanalyzing previously stored specimens. A potentially diagnostic finding was identified in 8 (14%) of cases, however in 7 of those 8 the finding would likely have been identified on short-read genome sequencing.