Title: The limits of Mendelian assumptions in genomic diagnostics: evidence from long-read whole genome sequencing and exome analysis

Authors: Erica D Smith (Ambry), Greta Pitsava (UCI), Krista Bluske (Ambry), Changrui Xiao (UCI), Rebekah Barrick (CHOC), Ivan De Dios (UCI), Kirsten Blanco (UCI), Elizabeth Chao (UCI), Kelly Radtke (Ambry), Eric Vilain (UCI), Emmanuèle C. Délot (UCI), Seth Berger (Ambry), Rachid Karam (Ambry)

Abstract: Analysis of diagnostic exome and genome sequencing often leverages assumptions of Mendelian inheritance (i.e. rarity in the population and/or family-based inheritance filtering) to prioritize variants and support scalability. These assumptions can lead to missed diagnostic opportunities, however, when causative dominant disease variants are transmitted from asymptomatic parents or occur at higher-than-expected population frequencies. This was observed in our research-grade analysis of long read Whole Genome Sequencing (lrWGS) as part of the UCI -GREGoR program. lrWGS was performed on 25 participants whose genetic diagnosis was not previously identified on exome or short-read Whole Genome Sequencing (srWGS). New findings were identified in six cases, half of which (three) were detected but not reported as diagnostic findings on previous srWGS or exome assays. These variants were likely filtered from previous analysis due to being inherited from “healthy” parents (*PPP2R5D* c.589G>C; p.Glu197Gln and *FGF8* c.356C>T; p.Thr119Met) or being found at higher-than-expected frequency in GnomAD v4 (74 alleles: *GIGYF1* c.332del; p.Leu111fs). These cases are not considered solved because the candidate variants do not follow Mendelian patterns.

To determine how often diagnostic variants are inherited from parents described as unaffected in a clinical setting, we performed a retrospective analysis of 13,000 consecutive diagnostic exome cases at a clinical laboratory. Inherited Pathogenic (P) or Likely Pathogenic (LP) variants were examined because they are protected from filtering based on inheritance or population frequency. Most variants interpreted as P/LP for autosomal dominant conditions were *de novo* in probands and were not included in this analysis. Of 222 inherited P/LP variants for autosomal dominant conditions, 118 (53%) were transmitted from a reportedly healthy parent, and 11 (5%) were transmitted from a parent with confirmed or suspected mosaicism. These inherited variants were mostly associated with conditions known to have incomplete penetrance, variable expressivity, or later onset, or were explanatory for only part of the proband’s phenotype.

These examples underscore the importance of gene-disease specific penetrance curation and consideration of variable expressivity during diagnostic case analysis. We conclude that caution should be used when ruling out variants based on family inheritance patterns or setting broad population thresholds for analysis in diagnostic genetic testing.