The limits of Mendelian assumptions in genomic diagnostics:

evidence from long-read whole genome sequencing and exome analysis

Erica D Smith[1] *(esmith@ambrygen.com)*, Greta Pitsava[2], Krista Bluske[1], Changrui Xiao[2], Rebekah Barrick[2], Ivan De Dios[2], Kirsten Blanco[2], Elizabeth Chao[2], Kelly Radtke[1], Eric Vilain[2], Emmanuèle C. Délot[2], Seth Berger[1], Rachid Karam[1]

[1]Ambry Genetics, 1 Enterprise, Aliso Viejo, CA; [2]Institute for Clinical and Translational Science, University of California, Irvine, CA

Background

- Diagnostic analysis of rare disease using exomes or genomes has historically relied on **assumptions of Mendelian inheritance** to reduce the number of variants for manual review
- Family-based inheritance filtering may eliminate low-penetrance, autosomal dominant diagnostic variants when parents are reportedly unaffected

AIM: To quantify how often diagnostic variants violate Mendelian patterns

Methods

- **Cohort 1: Research whole genome sequencing (WGS) analyses.** Specimens from patients with rare diseases were gathered as part of the GREGoR consortium. Short read WGS (srWGS) was analyzed using standard Mendelian filtering settings. Thorough manual analysis of long read WGS (lrWGS) was performed using GeneYX ***without pre-set filters based on Mendelian assumptions.***
- **Cohort 2: Diagnostic exome sequencing** was performed at Ambry Genetics using a proprietary analysis platform; ***Pathogenic and Likely Pathogenic variants are protected from filtering*** and are reported if the patient phenotype matches. The inheritance of reported Pathogenic (P) or Likely Pathogenic (LP) variants for Autosomal Dominant (AD) conditions was examined in exome trios.

Cohort 1: Candidate diagnoses in 12% (3/25) of lrWGS cases violate Mendelian assumptions

Case	Gene (c.; p.)	Inheritance	Condition	Alleles in gnomAD	Evidence
1	***PPP2R5D*** (c.589G>C; p.Glu197Gln)	Healthy father	Developmental and epileptic encephalopathy	0	2 P/LP variants at same aa, missense constraint
2	***FGF8**** (c.356C>T; p.Thr119Met)	Healthy father	Holoprosencephaly	2	Founder variant for holoprosencephaly PMID: 29584859
3	***GIGYF1**** (c.332del; p.Leu111fs)	*de novo*	Autism Spectrum disorder	74	Recurrent *de novo*, functional evidence PMID: 35917186

Table 1: 25 cases were negative after srWGS. 6 candidate findings were identified on lrWGS, 3 of which were also detected on srWGS but were deprioritized during analysis due to being inconsistent with inheritance pattern.
(*) Incomplete penetrance and/or variable expressivity

Cohort 2: Clinical exome cohort shows 62% of inherited P/LP diagnoses are from a healthy parent (Mendelian violation)

Figure 1: Study design

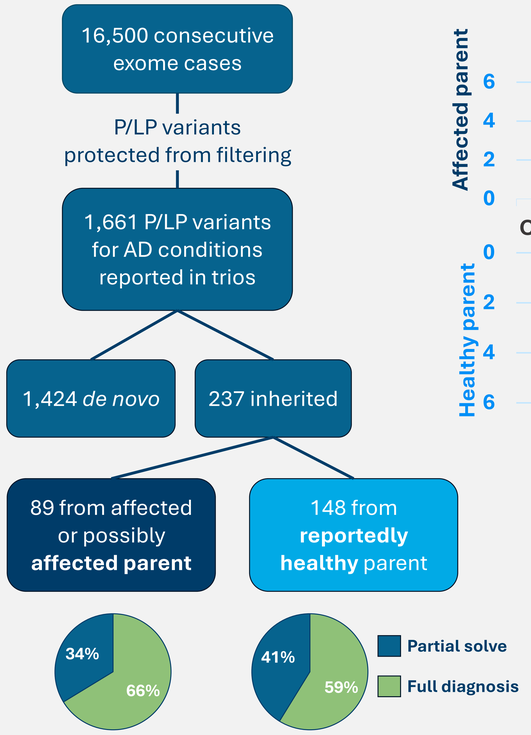
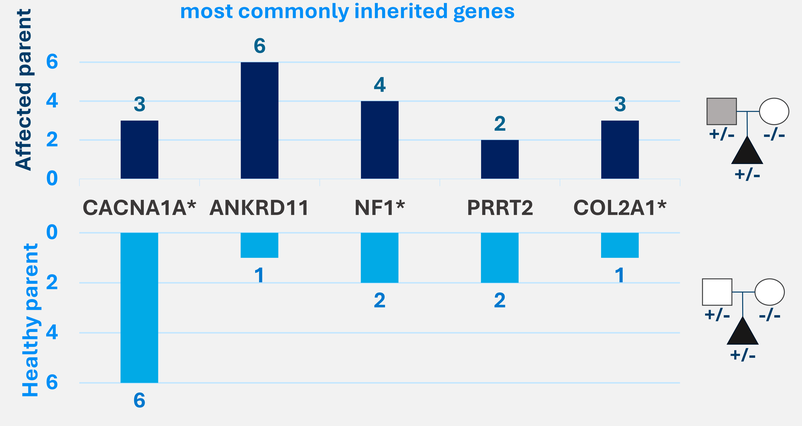


Figure 2: Trio P/LP variant counts in the most commonly inherited genes



Key Observations

- 62% (148/237) of inherited P/LP are from parents described by clinician as healthy, *i.e.* Mendelian violation (Fig 1)
- 41% (64/161) of P/LP variants inherited from healthy parents were "partial solves", *i.e.* explains only part of the patient's phenotype (Fig 1)
- Genes with most inherited P/LPs frequently exhibit variable expressivity and/or incomplete penetrance(*) (Fig 2)

Take Home Points

- **50% (3/6) of candidate findings** identified in 25 cases by lrWGS were also detected but filtered on srWGS because they **do not follow Mendelian patterns** (Table 1)
- Reported parental affected status is not reliable for **family-inheritance based variant filtering**, even for the full diagnosis (Figure 1).
- Case-level analysis of exome or genome sequencing should account for **gene-disease specific variable expressivity and penetrance** to optimize diagnoses (Figure 2)