Exome Sequencing Covers >99% of Mutations Identified on Targeted Next Generation Sequencing Panels

Elizabeth Chao[1,2], Jefferey Chen[1], Hsiao-Mei Lu[1], Wenbo Mu[1], Huy Vuong[1], Holly LaDuca[1], Kelly D. Farwell[1], Layla Shahmirzadi[1], Sha Tang[1]

[1]Ambry Genetics, Department of Clinical Diagnostics, Aliso Viejo, CA 92656; [2] University of California, Irvine, Department of Pediatrics, Division of Genetics and Metabolism, Irvine, CA 92868

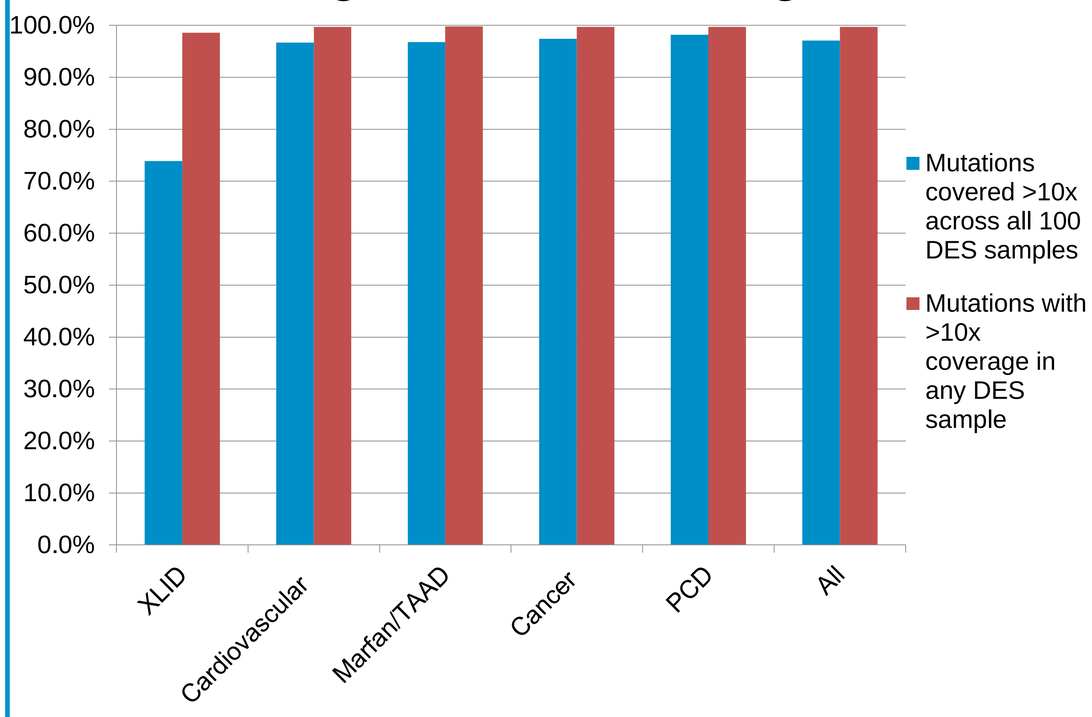
BACKGROUND

- With the expanded availability of next generation sequencing (NGS)-based clinical genetic tests, clinicians must weigh the superior coverage of targeted gene panels with the greater number of genes included in diagnostic exome sequencing (DES) when considering their first-tier testing approach.[1]
- This decision may be particularly challenging for diseases with significant genetic and phenotypic heterogeneity.
- To date, few studies have examined the analytic sensitivity of DES using position-specific basepair coverage.
- Here, we aim to predict the analytic sensitivity of DES using mutations identified on targeted NGS panels as a reference.

METHODS

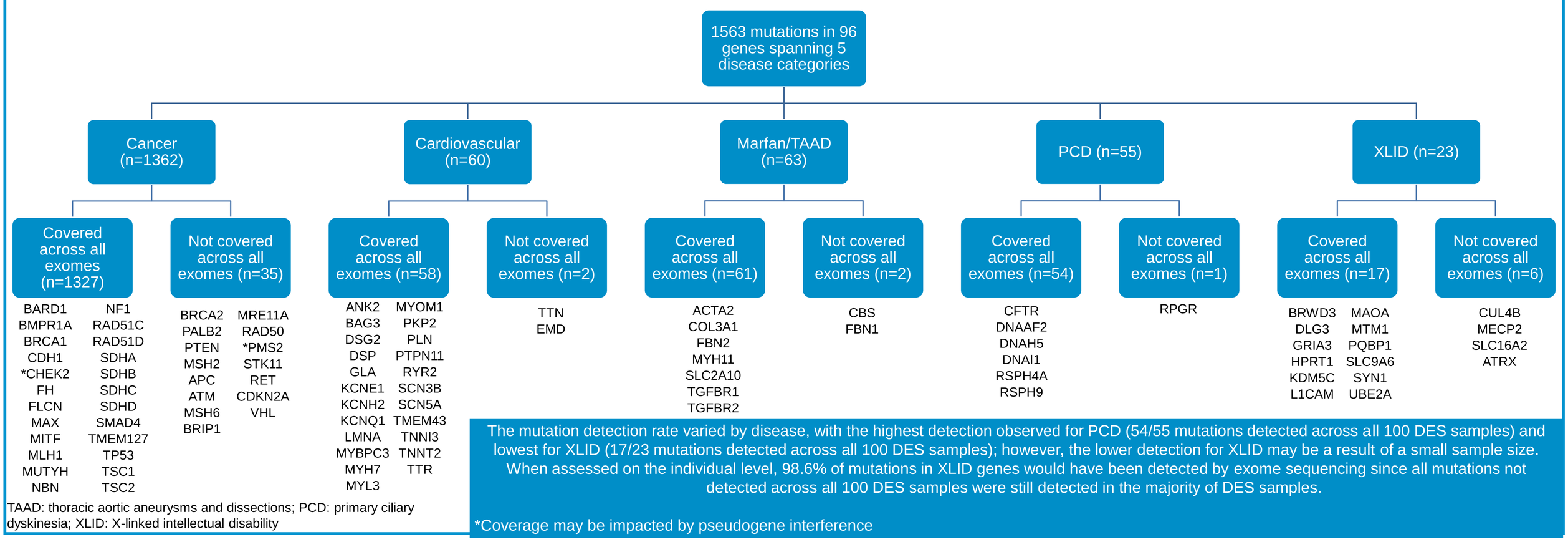
- Queried internal laboratory database for all pathogenic and likely pathogenic variants ("mutations") detected on targeted NGS multi-gene panel testing at our clinical diagnostic laboratory.
- 1563 different mutations identified on targeted NGS multi-gene panel testing were included in this analysis, representing 96 genes implicated in 5 disease categories (Figure 2).
- Single nucleotide substitutions were the most common type of mutation included in this analysis (n=691, 44.2%), followed by small deletions (n=486, 31.1%), intronic mutations (n=185, 11.8%), small duplications (n=137, 8.8%), and insertions and indels (each 2%). The lengthiest variants assessed were a 40-nucleotide deletion in *BRCA1*, and a 20-nucleotide duplication in *BARD1*.
- All mutations were identified by NGS and confirmed by Sanger sequencing. Corresponding nucleotide positions for these mutations were interrogated in data from 100 randomly-selected clinical DES samples to quantify the sequence coverage at each position.
- DES samples were prepared as previously described.[2]
- Mutations were interpreted as 'detected' if exome coverage at the respective nucleotide positon was ≥10x. Coverage at the flanking nucleotides was averaged for insertions, and for indels and deletions, coverage was recorded as the minimum of the first and last nucleotides.

Figure 1. DES coverage



Considering that coverage was assessed among 100 individual DES samples for each mutation (156,300 individual assessments), a total of 99.7% (n=155,772) of mutations would likely have been detected on DES. For 97.1% of mutations (n=1517), coverage at the respective nucleotide positions was ≥10x across all 100 DES samples.

Figure 2. Disease and gene-specific coverage on DES

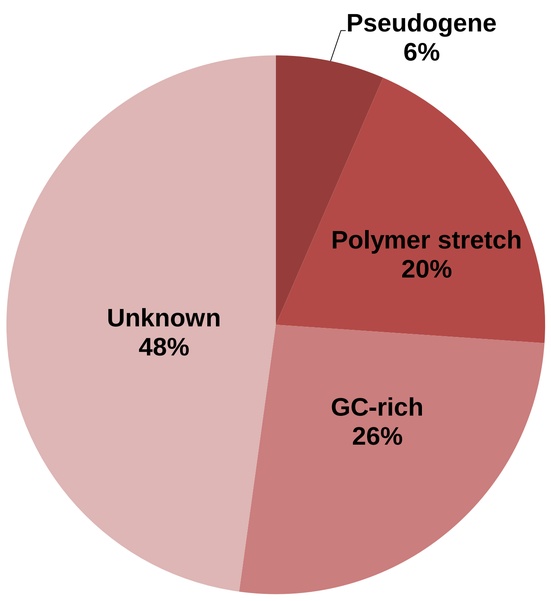


TAAD: thoracic aortic aneurysms and dissections; PCD: primary ciliary dyskinesia; XLID: X-linked intellectual disability

The mutation detection rate varied by disease, with the highest detection observed for PCD (54/55 mutations detected across all 100 DES samples) and lowest for XLID (17/23 mutations detected across all 100 DES samples); however, the lower detection for XLID may be a result of a small sample size. When assessed on the individual level, 98.6% of mutations in XLID genes would have been detected by exome sequencing since all mutations not detected across all 100 DES samples were still detected in the majority of DES samples.

*Coverage may be impacted by pseudogene interference

Figure 3. Potential explanations for incomplete coverage



For the 46 mutations that were not covered ≥10x across all 100 exomes, the number of samples with adequate coverage ranged from 36 to 99.

TAKE-HOME POINTS

- Despite current estimates that 90-95% exome-wide coverage is achieved with exome sequencing,[3] results from this position-specific comparative analysis limited to disease-causing mutations demonstrate that exome sequencing is expected to perform well (>99.5%) for a range of inherited diseases.
- This data suggests that the use of exome sequencing may achieve similar results/diagnostic yield when compared to panel based tests and may be an appropriate option to consider when indicated.

REFERENCES

1. Rehm, H.L. (2013). Disease-targeted sequencing: a cornerstone in the clinic. *Nature reviews Genetics* 14, 295-300.
2. Farwell, K.D., Shahmirzadi, L., El-Khechen, D., Powis, Z., Chao, E.C., Davis, B.T., Baxter, R.M., Zeng, W., Mroske, C., Parra, M.C., et al. (2014). Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. *Genet Med.* 2014 Nov 6
3. Biesecker, L.G., and Green, R.C. (2014). Diagnostic clinical genome and exome sequencing. *N Engl J* Med 370, 2418-2425.