Tracing the dark matter:
Prevalence of intragenic CNVs in Mendelian disease genes
R Truty, S Lincoln, M Kennemer, J Paul, E Gafni, S Fay, E Olivares, R Nussbaum, S Aradhya
Invitae Corporation, San Francisco, CA
Disclosure statement: All authors are employees and stockholders of Invitae Corporation.

Introduction
The prevalence of intragenic deletions and duplications (copy number variants; CNVs) has not been investigated extensively. Their clinical significance is well-characterized for only a handful of Mendelian disorders. We simultaneously detected sequence variants and CNVs by high-depth NGS and describe their occurrence in a large number of disease genes.
- Subsets of 1,507 genes were analyzed in 143,000 individuals referred for gene panel testing for hereditary cancer, cardiovascular, neurological, pediatric, and rare disorders
- We describe the prevalence and structure of intragenic CNVs in the context of their classification status, gene distribution, and mutational mechanism
- We also examined CNVs in genes unrelated to the presenting clinical phenotypes to understand the baseline prevalence and structure of these events in human disease genes

CNVs in clinical testing
- 2,096 pathogenic CNVs in 230 genes; 9% of all pathogenic variants
- Pathogenic CNVs mostly partial-gene events in dominant LOF genes*

Clinical CNVs: structure and interpretation

Clinical CNVs: inheritance and mutational mechanism*

Baseline CNV observations
- 4,215 CNVs in 601 genes; average of one CNV per 3,679 genes analyzed
- Predominantly large duplications in recessive and non-LOF genes*

Baseline CNVs: structure

Baseline CNVs: inheritance and mutational mechanism

Results

Baseline CNVs: structure

Baseline CNVs: inheritance and mutational mechanism

Both rare and recurrent variants are common
- Recurrent variants dominate some genes (PMP22, SMN1)
- Most genes show a diverse set of rare CNVs (DMD, BRCA2)
- Recurrent and rare CNVs found in some genes (BRCA1, MSH2)

Rare CNVs play important role in disease
- 32% of clinically reportable CNVs were observed only once
- 60% were observed ≤ 5 times

Many genes have very few CNVs
- Of the 384 genes in which CNVs were observed, 51% had only 1 CNV observed
- 18% of CNVs were observed in genes where ≤ 5 CNVs were found

Conclusions
- This study is among the largest to describe intragenic CNVs in human disease. Overall 9% of pathogenic variants are CNVs
- CNVs contribute significantly to positive diagnoses across clinical areas: 10% of individuals with a clinically significant result had a CNV
- Only testing for a recurrent CNV or genes known to have high CNV prevalence can miss 18%-60% of clinically relevant CNVs.
- Baseline CNVs display characteristics compatible with likely benign outcomes or in some cases point to carrier status for pathogenic variants in recessive disorders
- Our data provide insight into the baseline prevalence of intragenic CNVs in disease genes, helping shed more light on potential patterns of CNV variation expected in a healthy population

All clinical areas showed CNVs at rates of 5-35% of positive variants, with higher prevalence in oncology, neurology, and pediatrics. High CNV rates in neurology are dominated by SMN1, PMP22, and DMD; CNVs contribute to 6% of pathogenic results even after excluding these three genes

*LOF = loss of function mutational mechanism
** AD = autosomal dominant; AR = recessive