Copy number variant calling on a 176 condition expanded carrier screening panel reveals impact of *HBB* deletions

Kyle A. Beauchamp PhD, K. Eerik Kaseniit MEng, Peter Grauman BS, Gregory J. Hogan PhD, Kevin R. Haas PhD, Genevieve M. Gould PhD, Kenny K. Wong MS CGC, Gabriel Lazarin MS CGC, Eric A. Evans PhD, Dale Muzzey PhD

South San Francisco, California

Disclosure

All authors are current or former employees of Counsyl, Inc.

Introduction

Expanded carrier screening (ECS) identifies couples whose future children are at increased risk of Mendelian conditions and may be performed using either targeted genotyping (TG) or next generation sequencing (NGS). Historically, ECS panels have focused on deleterious SNPs and indels but have been performed with limited or no copy number variant (CNV) calling. Using the Modeled Fetal Disease Risk[1,2], here we evaluate the performance of hypothetical and commercial ECS panels. We also evaluate the impact of deletion CNVs on two ECS panels with 94 conditions and 176 conditions, respectively.

Modeled fetal disease risk

Previously, carrier frequency and carrier couple frequency have been used to quantify the sensitivity of ECS panels. However, to serve this purpose more effectively for diseases with complex inheritance (e.g., fragile X syndrome), we recently introduced[1,2] the modeled fetal disease risk (MFDR), which is the probability that a random fetus will be affected by one of the panel diseases. The modeled fetal disease risk allows comparisons of ECS panel sensitivity even for diseases with complex inheritance.

Lessons from hypothetical panels

To assess the sensitivity of various ECS approaches, we compared the modeled fetal disease risk captured by hypothetical panels containing up to 94 "Severe" and "Profound" conditions[3]. We first considered an NGS panel that excludes several "special case" diseases (fragile X syndrome, 21-hydroxylase-deficient congenital adrenal hyperplasia, alpha thalassemia, and spinal muscular atrophy) that are technically challenging to probe. We then considered the effect of adding special cases and panel-wide (i.e., non-founder) copy number (CNV) calling. We finally considered "best-possible" TG panels with a fixed number of optimally-selected variants, both with and without the special cases. The disease risk of each hypothetical panel shows that neglecting special cases and exon-wide coverage overlooks 10% to 55% of affected fetuses. Furthermore, non-founder CNVs contribute approximately 4 affected fetuses per 100,000 — roughly equivalent to the contribution of the 50 least-prevalent diseases on the 94 condition panel.

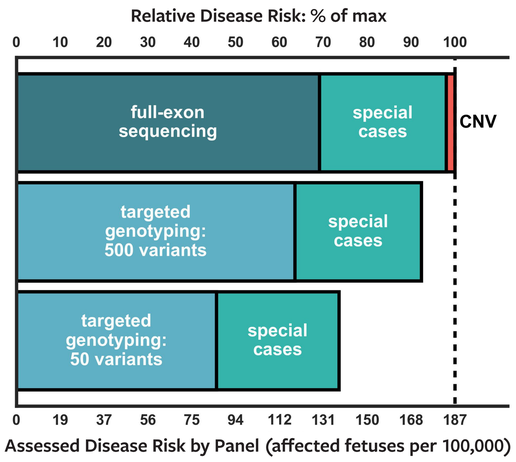

Figure 1: Modeled fetal disease risk (per 100,000 births) and percent of total risk is shown for hypothetical TG and NGS versions of the 94 condition ECS panel. Non-founder deletion CNVs contribute an additional 4 affecteds per 100,000 — approximately 2% of total risk.

High-prevalence genes dominate disease risk

A common question is how to best improve the sensitivity of an ECS panel. While adding more genes always increases the assessed disease risk, typically the most prevalent diseases contribute over half of the disease risk. Thus, improving ECS panels will likely require both increasing detection rate for existing diseases (such as via panel-wide CNV calling) and adding additional conditions.

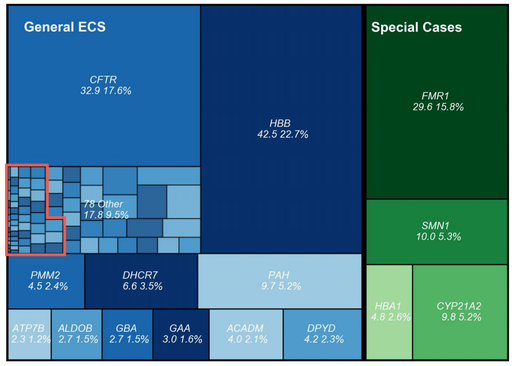

Figure 2: Modeled fetal disease risk (per 100,000 births) and percent of total risk is shown for each condition on the 94 condition panel. The red box shows the approximate number of single-gene conditions required to achieve a disease risk comparable to panel-wide deletion CNVs.

Modeled fetal risk of a 176 disease panel

Based on the previous observations, we developed an enhanced ECS panel with 176 diseases and panel-wide deletion calling. Here we report[4] modeled fetal disease risk estimates for this new panel, based on allele frequencies observed in 15,177 patients; the addition of new diseases (e.g., DMD) and technical features (panel-wide CNVs) lead to a substantially increased disease risk (323 per 100,000; approximately 1 in 300 births).

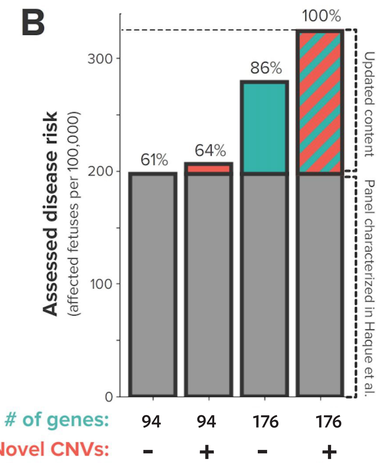

Figure 3: Modeled fetal disease risk is shown for hypothetical subsets of a new 176 disease ECS panel.

Per-gene contributors of copy number variants

Next, we examined a cohort of 27,446 patients to assess the per-gene impact of panel-wide copy number variant calling. Consistent with its known high frequency of CNVs, DMD showed the highest number of CNV carriers. However, CNVs in other genes (e.g., *HBB*) contributed substantially. Approximately ⅓ of the *HBB* deletions were observed in African Americans, among whom the deletion carrier frequency was approximately 0.8% — consistent with estimates from ExAC[5]. Considered as a whole, deletions comprise the third most common pathogenic *HBB* variant among African Americans.

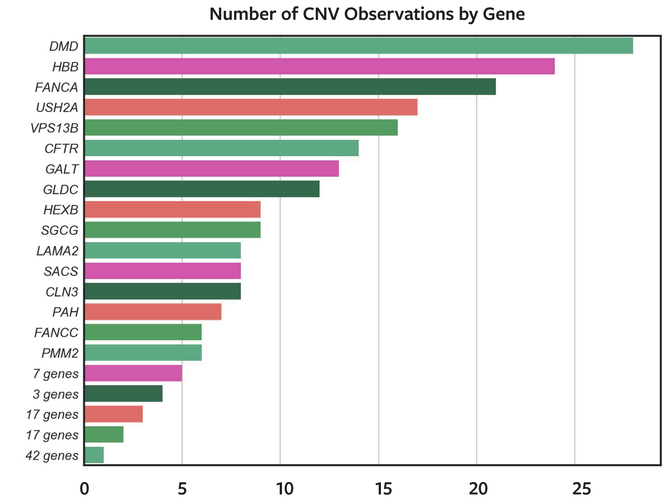

Figure 4: The number of observed pathogenic CNVs is listed for each gene. The bottom five rows represent the number of genes that have a particular number of observed pathogenic CNVs (e.g., for each of seven different genes, there were five patients observed to have pathogenic CNVs)

Conclusions

Modeled fetal disease risk allows systematic comparison of ECS panels. We previously predicted additional genes and copy number deletions as avenues for enhancing the clinical sensitivity of ECS. Using clinical reports from at least 15,177 patients, we assess these improvements on a 176 disease ECS to show a substantial increase in modeled fetal disease risk.

Methods

405,195 patients seeking ECS (Counsyl Family Prep Screen) between Jan. 2012 and Dec. 2016 for reason of "Carrier Testing" were anonymized and included in the disease risk analysis on the 94 disease panel; 56,267 of these samples were used for panel-wide copy number analysis. Results for self-reported ethnicities were reweighted based on US census data. For the 176-disease panel, we analyzed cohorts of 15,177 and 27,446 patients. Large deletions were called on most diseases except for technically challenging genes (e.g., *FMR1*). Large duplications were called on *CFTR* and *DMD*, as well as select special cases. Copy number variants were clinically interpreted and reviewed through our standard variant curation protocol.

REFERENCES 1. Haque, IS et al. Modeled Fetal Risk of Genetic Diseases Identified by Expanded Carrier Screening. JAMA. 2016;316(7):734-742. **2.** Beauchamp, KA et al. Systematic Design and Comparison of Expanded Carrier Screening Panels. GIM (2017; in Press). **3.** Lazarin, GA. et al. Systematic classification of disease severity for evaluation of expanded carrier screening panels. PLoS ONE 2014; 9(12): e114391. **4.** Hogan, GJ et al. Development and validation of an expanded carrier screen that optimizes sensitivity via full-exon sequencing and panel-wide copy-number-variant identification. bioRxiv, 2017. **5.** Lek, M et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature, 2016.

View all posters and research at:
research.counsyl.com