Design of scalable gene panels for carrier screening

Counsyl

Nikita Mehta, MS; Dale Muzzey, PhD; Matt Leggett, BS; Kenny K. Wong, MS; Gabriel A. Lazarin, MS; Eric A. Evans, PhD; Imran S. Haque, PhD

South San Francisco, CA

Background

For every 2,000 births worldwide, nearly three children have a severe autosomal-recessive or an X-linked disorder, and only ~10% of all inherited diseases may be effectively managed with treatment.[1] Expanded carrier screening (ECS) can help mitigate the impact of such diseases by either decreasing the number of affected pregnancies or improving care for an affected child via early diagnosis. The number of diseases on an ECS panel affects both its economic viability and clinical utility; thus, panel expansions must be done judiciously. Here we report a principled method by which disease genes can be ranked to guide ECS panel expansion.

Methods

We reviewed and ranked genes obtained from commercial ECS panels and Online Mendelian Inheritance in Man (Figure 1 and Table 1). Then, screening protocols were optimized for workflow and assay cost by consideration of homology, assay compatibility, and gene size (Figure 2).

Conclusion

Collectively, the entire process described above enables a general and principled selection of severe diseases for a clinically and economically viable ECS panel that identifies the most at-risk couples.

All Counsyl posters available online at research.counsyl.com

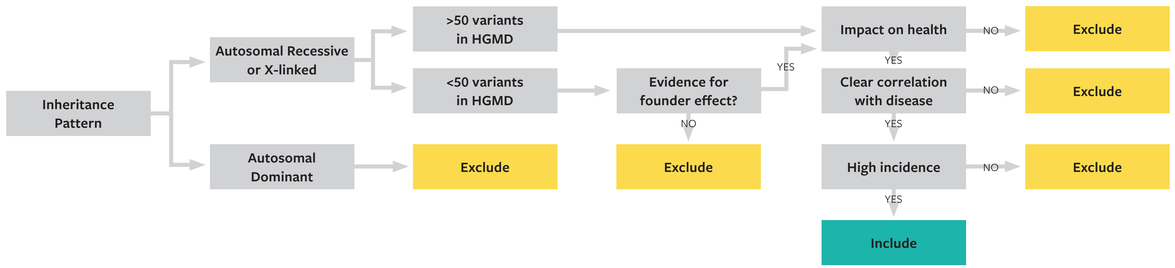


Figure 1

Gene selection algorithm. Inclusion criteria for consideration of the genes are illustrated. A quantitative proxy, based on the number and frequency of each gene's annotated variants, was used to determine clinical impact. For each associated disease phenotype, at least one board-certified genetic counselor assessed the following three factors: impact on health (e.g., shortened lifespan, intellectual disability, physical malformations), clear correlation with gene and symptoms, and high incidence (generally or within certain populations).

Disease	Gene	Characteristics	Relative risk
dystrophinopathies	*DMD*	• Severe[2] • 1/3,500 males (includes de novo cases)[3] • Adds detection to muscular dystrophies overall • No known founder effect • Not on NBS, but on an existing panel	1
X-linked combined variable immunodeficiency (including severe combined immunodeficiency)	*IL2RG*	• Severe[2] • 1/50,000-100,0003 • Does not add detection rate to a disease on current panel • No known founder effect • Present on NBS and an existing panel	2
maple syrup urine disease Ia	*BCKDHA*	• Severe[2] • 1/185,000 combined[3] • MSUD Ib on current panel • Founder effect in Mennonite population[2] • Present on NBS panel and an existing panel	3
oculocutaneous albinism 1	*TYR*	• Moderate[2] • 1/40,0003 • Does not add detection rate to a disease on current panel • May be more common in Japan and India[3,4] • Not on NBS, but on an existing panel	4
Mulibrey nanism	*TRIM37*	• Moderate[2] • Low prevalence (110 cases described as of 2004)[5] • Does not add detection rate to a disease on current panel • Probable founder effect in Finland (85 cases reported) • Not on NBS, but on an existing panel	5

Table 1

Diseases that were included based on the gene selection algorithm were further ranked based on: 1) disease severity, 2) disease incidence, 3) added detection rate to a specific disease phenotype already on the ECS panel, 4) added detection rate for genetic disease in a specific ethnic group, and 5) presence on the newborn screen (NBS) and/or an existing commercial ECS panel. Example rankings are listed below.

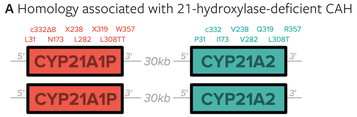
Figure 2

Two factors constrain how many genes comprise the final ECS panel: 1) the difficulty of assessing carrier status in the desired genes and 2) the aggregate number of positions being probed.

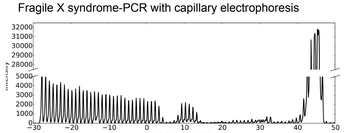
A Variant identification in genes can be challenging for a variety of reasons, including the presence of homologous regions (e.g., the CYP21A2 gene underlying 21-hydroxylase-deficient congenital adrenal hyperplasia (CAH) has a 99% identical pseudogene, CYP21A1P) or repeat sequences (e.g., CGG repeat in FMR1). Either custom NGS solutions or non-NGS assays (e.g., triplet-primed PCR followed by capillary electrophoresis for fragile X syndrome) are needed.

B The size of the genome interrogated is a major limiting factor in panel design, especially for tests performing full-exon NGS. Coding-sequencing length varies greatly from gene-to-gene. Below, the genes considered for inclusion on a commercial ECS panel are organized by coding-sequence length, which depicts that a single large gene can consume 40x as much NGS bandwidth as a smaller gene.

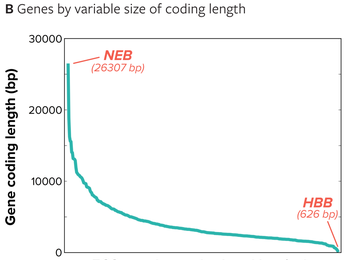
A Homology associated with 21-hydroxylase-deficient CAH



Fragile X syndrome-PCR with capillary electrophoresis



B Genes by variable size of coding length



CITATIONS 1| World Health Organization (http://apps.who.int/iris/bitstream/10665/41846/1/WHO_TRS_865.pdf), Control of Hereditary Diseases, 1996 2 | Lazarin et al., PLoS One, 2014 (PMID: 25494330) 3 | GeneReviews 4 | Ankala et al., Hum Mutat, 2015 (PMID: 25323826) 5 | Karlberg et al., N Engl J Med, 2004 (PMID: 15590968)