


Clinical experience with carrier screening in a general population: support for a comprehensive pan-ethnic approach

Maggie Westemeyer, MS, CGC¹, Jennifer Saucier, MS, CGC¹, Jody Wallace, MS, CGC¹, Sarah A. Prins, PhD¹, Aparna Shetty, PhD¹, Meenakshi Malhotra, PhD¹, Zachary P. Demko, PhD¹, Christine M. Eng, MD, FACMGG², Louis Weckstein, MD³, Robert Boostanfar, MD, FACOG⁴, Matthew Rabinowitz, PhD¹, Peter Benn, PhD⁵, Dianne Keen-Kim, PhD, FACMGG¹ and Paul Billings, MD, PhD ¹

Purpose: To present results from a large cohort of individuals receiving expanded carrier screening (CS) in the United States.

Methods: Single-gene disorder carrier status for 381,014 individuals was determined using next-generation sequencing (NGS) based CS for up to 274 genes. Detection rates were compared with literature-reported values derived from disease prevalence and carrier frequencies. Combined theoretical affected pregnancy rates for the 274 screened disorders were calculated.

Results: For Ashkenazi Jewish (AJ) diseases, 81.6% (4434/5435) of carriers identified did not report AJ ancestry. For cystic fibrosis, 44.0% (6260/14,229) of carriers identified had a variant not on the standard genotyping panel. Individuals at risk of being a silent spinal muscular atrophy carrier, not detectable by standard screening, comprised 1/39 (8763/344,407) individuals. For fragile X syndrome, compared with standard premutation screening, AGG

interruption analysis modified risk in 83.2% (1128/1356) premutation carriers. Assuming random pairing across the study population, approximately 1/175 pregnancies would be affected by a disorder in the 274-gene screening panel.

Conclusion: Compared with standard screening, NGS-based CS provides additional information that may impact reproductive choices. Pan-ethnic CS leads to substantially increased identification of at-risk couples. These data support offering NGS-based CS to all reproductive-aged women.

Genetics in Medicine (2020) 22:1320–1328; <https://doi.org/10.1038/s41436-020-0807-4>

Keywords: carrier screening; single-gene disorders; pathogenic variants; carrier frequencies; pan-ethnic screening

INTRODUCTION

The landscape of prenatal and preconception carrier screening (CS) is rapidly changing. Compared with traditional genotyping methods, modern genetic screening technology using next-generation sequencing (NGS) has rapidly expanded disease screening options.¹ Traditional prenatal and preconception CS targeted specific diseases with high prevalence, particularly in defined subpopulations. At present, CS for eight disorders (excluding cystic fibrosis [CF]) are recommended by the American College of Medical Genetics and Genomics (ACMG) and the American College of Obstetricians and Gynecologists (ACOG) for individuals of Ashkenazi Jewish (AJ) heritage.^{2,3} These professional guidelines also recommend pan-ethnic screening for CF and spinal muscular atrophy (SMA) and ACOG recommends pan-ethnic screening for hemoglobinopathies.^{4–6} Though these guidelines target some of the most prevalent genetic disorders, they do

not identify carriers for many common conditions of importance to the diverse subpopulations present in contemporary societies. Moreover, CS based on ethnicity and race is often untenable due to complex population admixture, multiethnic and multiracial populations, and missing or incorrect ethnic/racial information.⁷

Modern genetic screening technology using NGS overcomes these limitations by expanding the number of diseases covered and applies the testing to whole populations.⁸ Exome and genome sequencing may further expand the scope of variants identified in the CS panels.⁹ Recognizing these benefits, ACOG and ACMG have endorsed offering expanded CS to couples who are considering pregnancy or are already pregnant, regardless of ethnicity.^{10,11} Several studies have demonstrated that CS identifies an increased number of fetuses at risk for severe phenotypes compared with that based on traditional professional screening guidelines.^{12–14} Currently available

¹Natera, Inc., San Carlos, CA, USA; ²Baylor Miraca Genetics Laboratories, Houston, TX, USA; ³Reproductive Science Center of the San Francisco Bay Area, San Ramon, CA, USA; ⁴Huntington Reproductive Center Fertility, Encino, CA, USA; ⁵Department of Genetics and Genome Sciences, University of Connecticut Health Center, Farmington, CT, USA. Correspondence: Paul Billings (pbillings@natera.com)

These authors contributed equally: Maggie Westemeyer, Jennifer Saucier

Submitted 6 January 2020; revised 31 March 2020; accepted: 1 April 2020

Published online: 5 May 2020

expanded CS panels are extremely heterogeneous in panel size, with the number of diseases being screened ranging from 41 to 1556.¹⁵ A joint statement from the ACMG, ACOG, the National Society for Genetic Counselors (NSGC), Perinatal Quality Foundation (PQF), and the Society for Maternal–Fetal Medicine (SMFM) acknowledged the utility of more comprehensive, pan-ethnic CS and provides guidelines for the inclusion of disorders but does not give specifics regarding individual genes.¹⁶

We present here NGS-based CS results for more than 380,000 individuals from a pan-ethnic population. The CS panels contained up to 274 genetic disorders that met ACMG, ACOG, NSGC, PQF, and SMFM guidelines¹⁶ and included a number of screening enhancements. Carrier frequencies obtained from the NGS-based CS were compared with the literature-reported carrier frequencies. We predicted the rate of at-risk couples and of affected pregnancies in randomly paired or ethnically matched couples based on the observed carrier frequencies.

Carrier frequencies of specific genetic disorders that are guideline-recommended (AJ disorder panel) or commonly screened (CF, hemoglobinopathies, SMA, fragile X) were examined in greater detail. For CF, we compared detection by NGS-based screening with that by the ACMG 23 variant panel of the *CFTR* gene (CF23).¹⁷ For hemoglobinopathies, we identified silent carriers of α -thalassemia ($-\alpha/\alpha$) or carriers of mild/silent ($\beta+$) *HBB* variants, which traditional hematological tests do not always detect, but can be important in certain scenarios for risk assessment and prenatal diagnosis.^{4,18,19} Furthermore, traditional CS for SMA cannot detect silent (2 + 0) carriers, i.e., individuals with two copies of *SMN1* on one chromosome (duplication allele) and zero copies on the other (deletion allele).⁵ Here we report on screening for *SMN1* copy number and the single-nucleotide polymorphism (SNP) associated with an *SMN1* duplication haplotype whose presence indicates a risk of silent (2 + 0) SMA carrier status. For fragile X syndrome, premutation carriers at risk of having affected children can have additional risk refinement with AGG interruption testing.²⁰ We refined risk assessment for carriers by providing AGG interruption testing for the subset of premutation carriers with following AGG interruption testing.²¹ Finally, we also present the carrier frequency for Duchenne muscular dystrophy (DMD) and DMD-related dystrophinopathies, a group of related X-linked disorders with limited carrier frequency data published.

MATERIALS AND METHODS

Study cohort

We performed a retrospective analysis of de-identified data from 381,014 individuals with CS tests ordered by clinicians in the United States from 1ST May 2015 through 31ST December 2018. Referral sources included obstetricians, perinatologists, reproductive endocrinologists, geneticists, and genetic counselors. Cohort demographics, including race/ethnicity, were collected from information provided by the clinician or clinic on the requisition form. Retrospective

review of cases used in this study was exempt from institutional review board review (E&I Review, exemption protocol #19040–01). As this study was based on analysis of de-identified data, informed consent was not requested. Data from patients who, at the time of testing, had declined research participation were excluded.

CS testing platforms

Sequence analysis including NGS and copy-number analysis was performed for CS panel sizes ranging from 1 gene to 274 genes (Table S1). Sequencing included all exons and ± 10 base pairs (bp) into the introns for the majority of the genes. Testing was carried out at Natera, Inc. (San Carlos, CA), Baylor Miraca Genetics Laboratories (Houston, TX), and/or the Icahn School of Medicine at Mount Sinai (New York, NY). There were six panels options (4, 14, 27, 106, 137, and 274 genes) and à la carte single-gene testing (DMD, CF, and SMA). All included disorders involve a cognitive or physical disability, the need for postnatal surgical or medical intervention, or a detrimental effect on quality of life; and are disorders where prenatal intervention could improve perinatal outcomes and delivery management or where prenatal education could prepare parents for special needs after birth.¹⁰ All testing therefore met the screening criteria recommended in the 2015 ACMG, ACOG, NSGC, PQF, SMFM joint statement.¹⁴ Males were not screened for X-linked genes.

A hybridization capture based methodology (SureSelectTM QXT, Agilent, Santa Clara, CA) was used, followed by paired-end sequencing using either the NextSeq or HiSeq2500 system (Illumina, San Diego, CA) with alignment to GRCh37 (hg19) to identify variants. All variants were classified using ACMG/Association for Molecular Pathology (AMP) guidelines.²² Only pathogenic and likely pathogenic variants were reported; classification was based on the understanding of gene variants at the time of reporting. Confirmation of variants was performed using Sanger sequencing (BigDye[®] Direct, Thermo-Fisher Scientific, Waltham, MA). Copy-number variants were identified by either quantitative polymerase chain reaction (qPCR), an appropriate multiplex ligation probe assay (MLPA, MRC-Holland, Amsterdam, Netherlands), or NGS.

The *FMRI* gene (fragile X syndrome) was evaluated using AmpliDye[®] PCR reagents (Asuragen, Austin, TX) and capillary electrophoresis on ABI3730xl instruments (Thermo-Fisher Scientific, Waltham, MA). Triplet repeat expansions (CGG trinucleotide repeats) were counted using either GeneMarker[®] (Softgenetics, State College, PA) or an in-house developed tool. Expansions within the premutation size range of 55–90 CGG repeats were reflexively tested for AGG interruptions at Asuragen (Austin, TX, CLIA #45D1069375) to refine the risk of CGG repeat length expansion to pathogenic variant (>200 CGGs). The *SMN1* g.27134T>G variant was assessed in conjunction with *SMN1* copy number to provide a risk estimate of SMA silent (2 + 0) carrier status.²³ *SMN2* copy number was not assessed.

Data and statistical analyses

For each gene, the observed carrier frequency was calculated and compared with the published, expected carrier frequencies established by various approaches (Appendix 1). A Chi-squared goodness-of-fit test was used to test the null hypothesis that the proportion of positive results for a specific gene in the patient cohort was equal to the expected proportion as estimated from literature reports of disease incidence and prevalence. We employed the Benjamini–Yekutieli multiple testing correction without assumptions on the joint distribution of the test statistics and bounded the false discovery rate below 5%. For ease of interpretation, we computed adjusted *p* values using the Benjamini–Yekutieli multiple testing correction, which can then individually be used to determine the result of each hypothesis test by comparing the adjusted *p* value to the significance level of 0.05. For genes in which the observed and expected carrier frequencies were both $\leq 1/500$ ($n = 113$), the differences between observed and expected were not considered meaningful and were excluded in the comparisons of rates.

To estimate the proportion of affected pregnancies expected to occur in our screening population, we first calculated the observed carrier frequency for each disorder (including those with frequencies $< 1/500$) by dividing the number of positive carrier results (pathogenic or likely pathogenic variants) by the total number tested. We assumed no prior knowledge of risk status for any of the conditions. Only those with one copy of *SMN1* (carriers) were considered to calculate expected number of SMA affected pregnancies; at-risk individuals for silent carrier status with *SMN1* (2 + 0) genotypes were not included in this calculation. For α -thalassemia, individuals with the genotype $\alpha\text{-}/\alpha\alpha$ (silent carriers) were excluded. Those with $\alpha\alpha\text{-}/\alpha\text{-}/\alpha\text{-}$ and $\alpha\text{-}/\alpha\text{-}$ genotypes were included in calculating the carrier frequency. Inclusion of $\alpha\text{-}/\alpha\text{-}$ genotype may have overestimated the combined projected affected pregnancy rate.

This calculation was repeated for ethnicity-based subcohorts, and the projected numbers of affected pregnancies were calculated for couples within ethnicities and randomly coupled across all ethnicities. For autosomal recessive diseases, the projected affected pregnancy rate = at-risk couple rate \times 0.25, where at-risk couple rate = (observed general or ethnic-specific population carrier rate)². For X-linked recessive diseases, the projected affected pregnancy rate = the observed carrier rate \times 0.25 (0.5 risk of passing on the pathogenic variant by the mother \times 0.5 risk of a male fetus). For fragile X syndrome, the projected affected pregnancy rate = the observed premutation sized *FMR1* CGG repeat values \times the risk for expansion to a pathogenic variant \times 0.5 (risk of passing on the pathogenic variant to the offspring).²⁴ The projected number of affected pregnancies for each gene was summed to determine the total number of affected pregnancies.

RESULTS

Population demographics

Reported race/ethnicity and demographics are summarized for all 381,014 individuals screened in Table S2. Overall,

148,828 (39.06%) individuals in our cohort were documented as Caucasian, 62,626 (16.44%) Hispanic, 52,454 (13.77%) African American, and the remaining 117,106 (30.74%) were either of other races/ethnicities or did not provide information. As most clinicians only ordered screening on the male partner if the female had a positive result for an autosomal recessive disorder, our cohort was mostly women (339,739, 89.17%). Seventy-three percent of the cohort were between the ages of 19 and 35 years.

CS analysis

The most commonly screened genes from the CS panel (Table S3) were *CFTR* (374,911; 98.39% of all individuals screened), *SMN1* (344,407; 90.39%), *FMR1* (290,745; 76.30%), and *DMD* (288,268; 75.65%). The carrier frequency data set of CF, SMA, fragile X, and *DMD* obtained from our data set are discussed in greater detail below, along with the ACMG/ACOG recommended AJ disorder panel^{2,3} and hemoglobinopathies. Among 60,052 individuals screened with the largest 274 gene panel, 38,300/60,052 (63.78%) were positive for one or more disorders. Of those positive cases, 22,116/38,300 (57.74%) were positive for one disorder, 11,201/38,300 (29.25%) for two disorders, 3772/38,300 (9.85%) for three disorders, 953/38,300 (2.49%) for four disorders, and 258/38,300 (0.67%) for five or more disorders.

Of all the genes screened, 117/274 (42.70%) had observed carrier rates that were significantly different from expected, as estimated from literature reports of disease incidence or carrier screening (Table S3, Appendix 1). The disorders with the greatest proportional difference between the observed and expected carrier frequencies were familial Mediterranean fever, *MEFV*; *PEX6*-related Zellweger syndrome spectrum, *PEX6* (both with higher observed carrier frequencies than expected); and spondylothoracic dysostosis, *MESP2*; galactokinase deficiency, *GALK1* (both with lower observed carrier frequencies than expected, Fig. 1).

The observed carrier frequencies of the eight ACOG/ACMG-recommended AJ disorders (Tay–Sachs disease, Gaucher disease, familial dysautonomia, Canavan disease, Bloom syndrome, mucopolidosis IV, Fanconi anemia group C, and Niemann–Pick disease type A; excluding CF)^{2,3} were analyzed based on ethnicity as reported on test requisition forms. Among individuals that listed only AJ ancestry, carrier frequencies for these disorders ranged from 1/18 (Gaucher disease) to 1/125 (Niemann–Pick type A). The corresponding carrier frequencies for the eight disorders for the full screened population ranged from 1/123 (Gaucher disease) to 1/985 (mucopolidosis IV, Table S4). Only 1001/5435 (18.42%) carriers of a variant in any one of these eight genes reported AJ ancestry, while 4434/5435 (81.58%) of cases did not report AJ ancestry. Most of the carriers (range: 65–98%) for these disorders were of non-AJ ancestry (Fig. 2).

Cystic fibrosis

Of the 374,911 individuals screened for *CFTR*, 14,229 (3.80%, 1/26 carrier frequency) were observed to have a

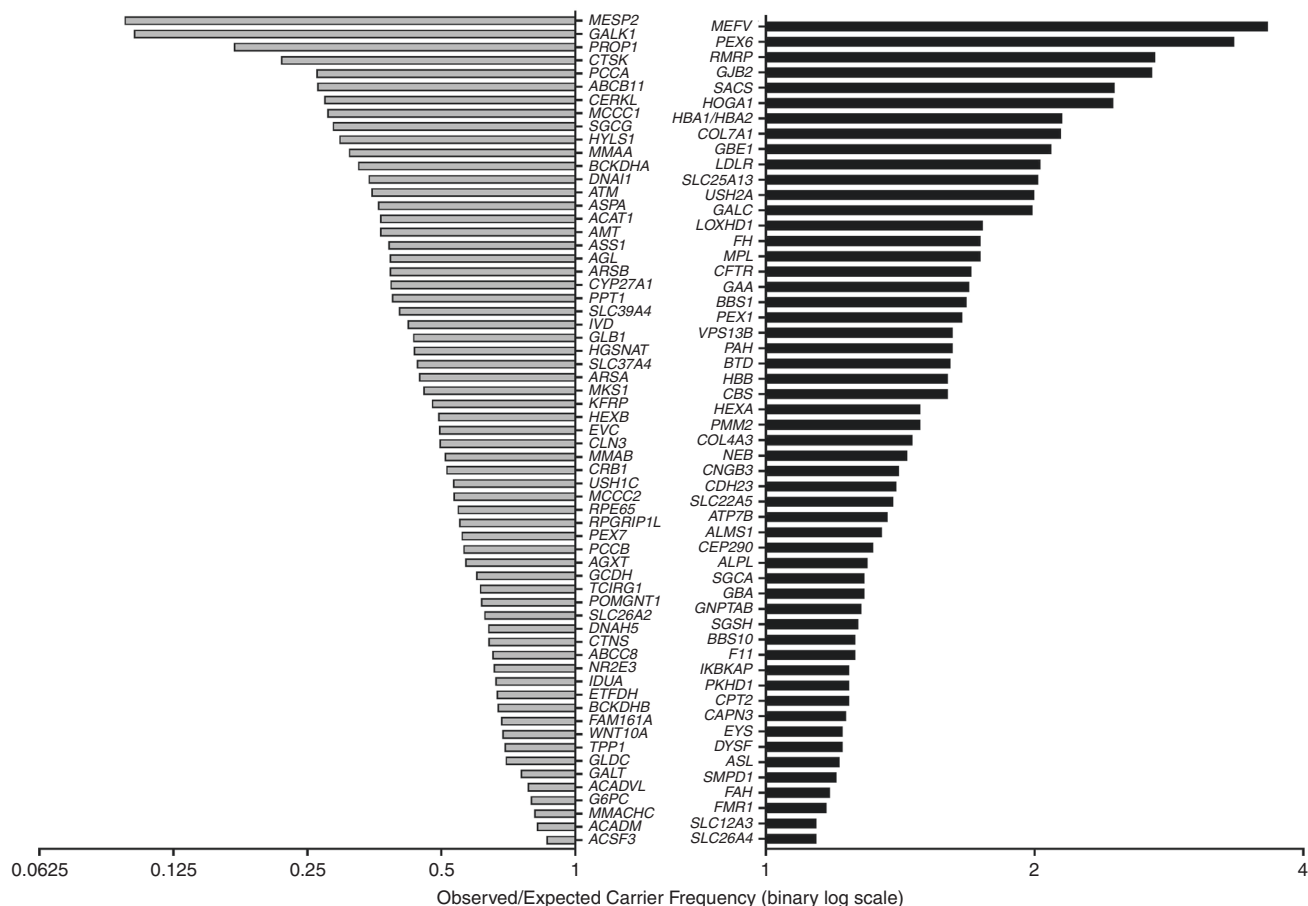


Fig. 1 Disorders where the observed carrier frequency was significantly higher or lower than expected. The relative differences in observed carrier frequencies of the 117 single-gene disorders that were significantly higher or lower than expected from prior literature based on disease prevalence and frequencies are presented in descending order.

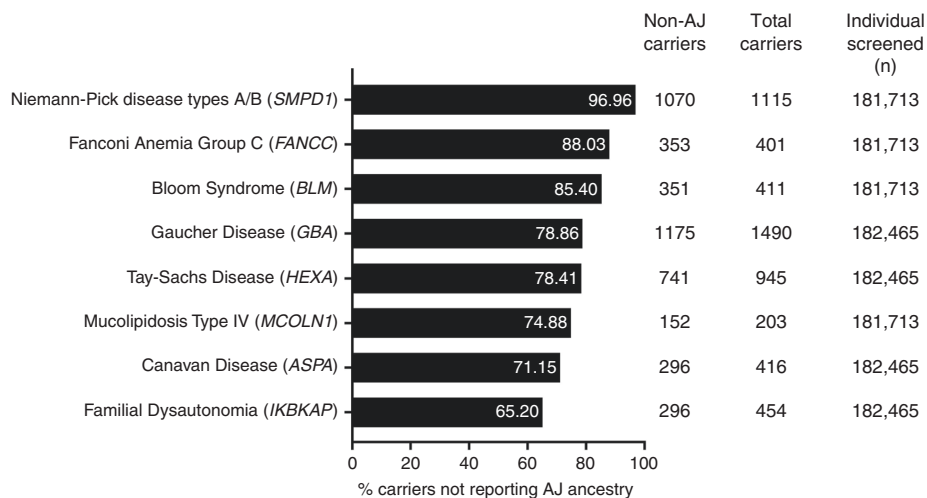


Fig. 2 Percentage of carriers of an Ashkenazi Jewish (AJ) disorder not reporting AJ ancestry. Proportion of carriers that did not report AJ ancestry or were of unknown ancestry among total carriers for the eight “Jewish genetic disorders”.

pathogenic or likely pathogenic *CFTR* variant, with 245 (1.72%, 1/1530) individuals having more than one variant identified, in either the *cis* or *trans* configuration. Among variant carrier positive individuals, 504 different *CFTR*

variants were detected (some seen in multiple individuals), of which 481 (95.43%) were not included in the standard targeted genotyping panel (CF23).¹⁷ In addition, 6260 individuals (43.99% of carriers) had *CFTR* variants that

would not have been detected by the standard CF23 screen.¹⁷ We examined the common *CFTR* variants by racial/ethnic groups (Table 1) and observed that among East Asians and South East Asians, four of the five most common *CFTR* variants detected by NGS would not be detected by the CF23 screen. Among African Americans, three of the five most common *CFTR* variants detected by NGS would not be detected by the CF23 screen (Table 1).

Spinal muscular atrophy

Of 344,407 individuals screened for SMA, 14,606 (4.24%, 1/24) were carriers or at-risk silent carriers; of these, 7 individuals had no copies of *SMN1*, 5836 (1.69%, 1/59) were carriers with a single copy of *SMN1*, and another 8763 (2.54%, 1/39) were at risk for being silent (2 + 0) carriers with two copies of *SMN1* based on the presence of the g.27134T>G SNP.²³ Those at risk for being silent carriers varied by race/ethnicity, ranging from 1/8 (11.82%) in African Americans to 1/1540 (0.06%) in East Asians (Table S5).

Fragile X syndrome

Of 290,745 women screened for *FMRI* CGG repeat length, 33 (0.011%, 1/8810) had >200 CGGs (pathogenic variant carriers); 1356 (0.47%, 1/214) had 55–200 CGGs (premutation carriers). AGG interruption testing was performed in premutation carriers with 55–90 CGG repeats to further evaluate CGG expansion risk.^{20,24} Of the 1356 premutation carriers, 1259 were reflexed to AGG interruption testing and 1128 (89.59%) had an adjusted risk for expansion to a pathogenic variant (>200 CGGs): 1013 (80.46% of the total reflexed) had decreased risk, 115 (9.13%) had increased risk, and 131 (10.41%) had no change in risk for CGG repeat expansion to a full pathogenic variant. Change in risk assessment after AGG testing by racial/ethnic group was not meaningful due to insufficient data (Table S6). The number of AGGs detected ranged from none to five: 163 (12.94% of the total reflexed) individuals did not have any AGG interruptions; 500 (39.71%) had one, 566 (44.95%) had two, 25 (1.98%) had

Table 1 Most common *CFTR* variants (cystic fibrosis) by race/ethnicity detected by expanded carrier screening.

| Racial or ethnic group | Sum of variants ^a | Most common variants ^b | Number reported | % of total | Detectable by CF23 |
|------------------------|------------------------------|-----------------------------------|-----------------|------------|--------------------|
| Caucasian | 7311 | c.1521_1523delCTT (p.F508del) | 3865 | 52.87 | Yes |
| | | c.350G>A (p.R117H) | 557 | 7.62 | Yes |
| | | c.3154T>G (p.F1052V) | 226 | 3.09 | No |
| | | c.3454G>C (p.D1152H) | 151 | 2.07 | No |
| | | c.1624G>T (p.G542*) | 139 | 1.90 | Yes |
| Hispanic | 1880 | c.1521_1523delCTT (p.F508del) | 561 | 29.84 | Yes |
| | | c.3454G>C (p.D1152H) | 119 | 6.33 | No |
| | | c.3154T>G (p.F1052V) | 100 | 5.32 | No |
| | | c.1624G>T (p.G542*) | 68 | 3.62 | Yes |
| | | c.350G>A (p.R117H) | 64 | 3.40 | Yes |
| African American | 1271 | c.1521_1523delCTT (p.F508del) | 296 | 23.29 | Yes |
| | | c.2988+1G>A | 121 | 9.52 | Yes |
| | | c.1865G>A (p.G622D) | 86 | 6.77 | No |
| | | c.1853T>C (p.I618T) | 42 | 3.30 | No |
| | | c.3297C>A (p.F1099L) | 32 | 2.52 | No |
| East Asian | 134 | c.1865G>A (p.G622D) | 20 | 14.93 | No |
| | | c.3209G>A (p.R1070Q) | 18 | 13.43 | No |
| | | c.3205G>A (p.G1069R) | 15 | 11.19 | No |
| | | c.1521_1523delCTT (p.F508del) | 10 | 7.46 | Yes |
| | | c.2909G>A (p.G970D) | 7 | 5.22 | No |
| South East Asian | 187 | c.1865G>A (p.G622D) | 46 | 24.60 | No |
| | | c.3209G>A (p.R1070Q) | 33 | 17.65 | No |
| | | c.1521_1523delCTT (p.F508del) | 24 | 12.83 | Yes |
| | | c.1558G>A (p.V520I) | 15 | 8.02 | No |
| | | c.1367T>C (p.V456A) | 14 | 7.49 | No |
| Ashkenazi Jewish | 255 | c.3846G>A (p.W1282*) | 74 | 29.02 | Yes |
| | | c.1521_1523delCTT (p.F508del) | 56 | 22.96 | Yes |
| | | c.3454G>C (p.D1152H) | 34 | 13.33 | No |
| | | c.3154T>G (p.F1052V) | 18 | 7.06 | No |
| | | c.1624G>T (p.G542*) | 8 | 3.14 | Yes |

^aSum of all variants that were detected in *CFTR* carriers.

^bThe variants noted here reflect classification as likely pathogenic or pathogenic at the time of manuscript submission. Classifications may change over time.

three, four (0.31%) had four, and one (0.07%) case had five AGG interruptions.

DMD-related dystrophinopathies

A total of 288,268 women were screened for carrier status for DMD-related dystrophinopathies (DMD, Becker muscular dystrophy, and cardiomyopathy). Of these, 402 (0.14%, 1/717) were positive for a pathogenic (212/402, 52.74%) or likely pathogenic variant (189/402, 47.01%, no results were obtained from 1 sample). Analysis of carrier frequencies of *DMD* variants by race/ethnicity is presented in Table S7. Of the 402 carriers, 78 (19.40%) reported a family history of a male with muscular dystrophy or a female carrier of a *DMD* variant, 123 (30.60%) reported no family history of muscular dystrophy, and 201 (50.00%) cases did not provide any family history information. Family history data were not analyzed for negative samples. Variants detected included large deletions (246, 61.19%) or duplications of at least one exon (62, 15.42%), and other variants such as single-nucleotide variants (SNVs), indels (short insertions and deletions <72 bp), and splice site variants (94, 23.38%).

Hemoglobinopathies

Of the 182,465 individuals screened for *HBA1/2* and *HBB*, 15,694 (8.60%) screened positive for α -thalassemia and 5972 (3.27%) screened positive for a β -hemoglobinopathy. Of those individuals positive for *HBA1/2*, 14,064 (89.61%) were silent carriers ($-\alpha/\alpha\alpha$) and 1630 (10.39%) were classified as carriers or affected with hemoglobin H disease (typically $-\alpha/-\alpha$, $-\alpha/\alpha$, $-/-\alpha$). Among β -hemoglobinopathy carriers, 358 (6.01%) had a mild/silent (β^+) variant, while the remainder (93.99%) had either a major (β^0) or structural *HBB* variant. The analysis of carriers by race/ethnicity and type, including silent carriers status, is presented in Fig. 3.

Projected at-risk couples and affected theoretical pregnancies

We analyzed the observed carrier frequencies in the overall testing population to calculate the projected at-risk carrier couple frequencies and number of theoretical pregnancies expected to be affected by one of the 274 diseases (Table 2). Assuming random couple-pairing across all ethnicities in the population, and no prior knowledge of risk status, 1 in 44 (2.3%) couples would be expected to be at risk and 1 in 175 (0.6%) theoretical pregnancies were projected to be affected by one of the CS diseases. If couple-pairing within each racial/ethnic group is assumed, then the at-risk couple rate would range from 1 in 17 (AJ) to 1 in 59 (Hispanic) and the rate of affected pregnancies would range from 1 in 55 (AJ) to 1 in 192 (Hispanic). Smaller panels would result in identification of substantially fewer affected pregnancies. For example, based on our data set, only 1 in 874 pregnancies from random couple-pairings with a genetic disorder would be detectable by the four-gene CS panel, and 1 in 620 pregnancies with a genetic disorder be detectable by the 27-gene CS panel; less

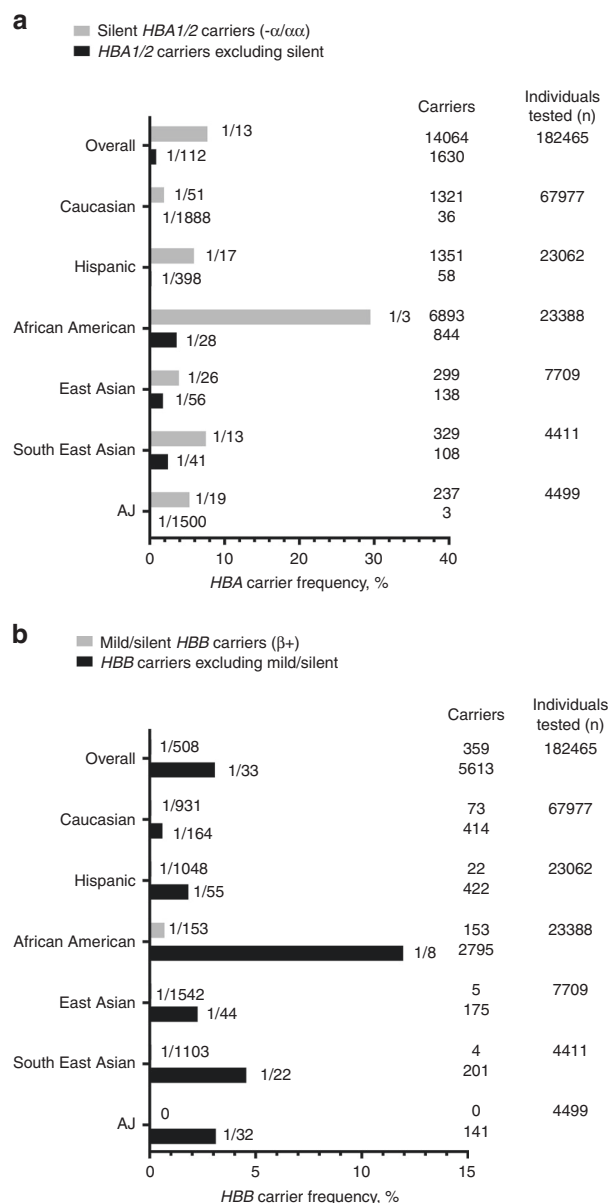


Fig. 3 Observed α -thalassemia and β -hemoglobinopathies carrier frequencies by race/ethnicities. (a,b) Carrier frequencies observed for (a) α -thalassemia with silent status and (b) β -hemoglobinopathy with mild/silent status from next-generation sequencing (NGS)-based carrier screening (CS) are presented. AJ Ashkenazi Jewish.

than a quarter and a third as many, respectively, than if the 274-gene panel were used.

DISCUSSION

This study presents CS results from the pan-ethnic US population. To the best of our knowledge, this cohort of more than 380,000 tests is currently the largest such data set available. Testing was based on panels of up to 274 genes meeting clinical utility criteria outlined in the 2015 ACMG, ACOG, NSGC, PQF, and SMFM joint statement.¹⁶ We calculated that, based on random mating, the 274-gene panel would identify a risk for a genetic disorder in the offspring of

Table 2 Combined at-risk couple rate and affected pregnancies expected across and within racial/ethnic groups by 1 of the 274 disorders.

| | At-risk couple rate (%) | Affected pregnancies rate (%) ^a |
|------------------------------------|-------------------------|--|
| General population random coupling | 1/44 (2.28) | 1/175 (0.58) |
| Racial/ethnic group | | |
| Caucasian | 1/32 (3.16) | 1/108 (0.93) |
| Hispanic | 1/59 (1.70) | 1/192 (0.52) |
| African American | 1/23 (4.42) | 1/84 (1.19) |
| East Asian | 1/35 (2.88) | 1/128 (0.78) |
| South East Asian | 1/34 (2.96) | 1/123 (0.81) |
| Ashkenazi Jewish | 1/17 (5.96) | 1/55 (1.81) |

^aFor α -thalassemia, individuals with the genotype $\alpha\alpha$ (silent carriers) were excluded. Those with $\alpha\alpha$ -, α -/ α - and α -/- genotypes were included in calculating the carrier frequency. Inclusion of the α -/ α - genotype may have overestimated the combined projected affected pregnancy rate.

1 in 44 (2.3%) couples. Rates were higher for most within-race/ethnicity couples. These estimates excluded those with risks associated with *SMN1* (2 + 0) and *HBA1/2* ($-\alpha/\alpha$) silent carrier status.

NGS-based CS for CF has the advantage of identifying substantially more pathogenic or potentially pathogenic variants compared with a CF23 panel. The CF23 panel was designed to identify *CFTR* variants with >0.1% frequency in a cohort of predominantly non-Hispanic Caucasian CF patients.¹⁷ Our results show NGS-based CS is particularly advantageous for CF screening in East and South East Asians and African Americans because some of the most common variants in these groups are not included in the CF23 panel.^{25,26} Likewise, our approach to SMA screening also allows for the identification of considerably more at-risk pregnancies, particularly among African Americans where identification of the g.27134T>G variant (linked to two copies of *SMN1*) facilitates identification of those at risk to be silent (2 + 0) carriers. We screened for disorders typically associated with AJ ancestry in non-AJ ethnicities, and for variants in genes with frequent disease occurrence but without standard panels, such as for DMD-related dystrophinopathies and hemoglobinopathies. The Supplementary Discussion provides additional description for each of these disease-specific screening enhancements.

This study is based on a cohort of mostly women choosing CS from many referral centers in the United States. Information on race/ethnicity was provided for 85% of the individuals tested. Although the distribution appears to approximate the demographics for the United States,²⁷ some racial/ethnic groups may be underreported due to ambiguous classification definitions, complex heritage or reluctance to divulge racial/ethnicity-based information. As test referral is based on individual patient preference, we cannot exclude ascertainment bias based on increased screening of women with a family history of a disorder or modified a priori risk

due to carrier testing in a relative. Approximately 20% (78/402) of women receiving DMD screening did report a positive family history but generally this type of information was unavailable for other conditions screened. As individual disorders are rare, and carrier screening for many is relatively recent, we believe these biases are unlikely to change the overall implications of our observed carrier estimates.

A central question in CS is the optimal content of test panels.^{10,11,16} The development of large panel, >200 gene tests, selected per guidelines from the ACMG, ACOG, NSGC, PQF, and SMFM joint statement,¹⁶ and offered to all individuals, recognizes the diversity in contemporary populations and circumvents the need for determinations based on complex, and often inaccurate, ancestry information. Similar to a previous study, we observed relatively high frequencies of variants for the eight AJ genetic disorders among those referred with no documented AJ ancestry.¹⁴ These observations illustrate the utility of applying current panels to all individuals. A 1/100 carrier frequency has been suggested as a cutoff guide when considering diseases to include on an expanded CS panel.¹⁰ However, if such a cutoff were to be selected for a total population, this could be a disadvantage to subpopulations where a particular disorder is more common. It is also important to consider X-linked diseases, such as fragile X and DMD, using carrier frequency inclusion criteria designed around the projected affected pregnancies rates of the full CS panel.^{10,11,28} For likely pathogenic variants (especially for combinations in autosomal recessive disorders), there is some uncertainty regarding phenotype. Panel design also must account for the different needs of various populations considering screening and should respect patient autonomy. Indeed, ACOG acknowledges that screening approaches should consider the personal values of the patient regarding appropriate counseling.¹⁰

Although not the purpose of CS, this testing can potentially facilitate health benefits for carrier parents (e.g., early identification of risk for primary ovarian insufficiency and fragile X-associated tremor/ataxia syndrome in individuals with expanded *FMRI* repeats, cardiomyopathy, or other DMD-related symptoms in carriers of dystrophin gene variants). Current standards of care require disclosure of such information; some individuals may not wish to receive this type of information and they may decline testing to avoid receiving this. These ethical challenges emphasize the importance of defined policies endorsed by professional organizations, carrier screening education, and individualized counseling.

Our overall assessment of various panel designs indicates clear benefits associated with larger panels; it is important to note that the relative benefit varies by race/ethnicity (Table 2). CS using more selective panels may be appropriate for some individuals. However, routinely providing limited panels to patients based on race/ethnicity raises problems with equitable provision of services. Access to a selection of panel sizes provides clinicians and patients with choices. Our experience with offering different size panels showed that

80.0% of test orders were for 27 or fewer genes (Table S1). The ordering preference for larger CS panels (>100 genes) at our laboratory was greatest among reproductive endocrinologists who offer CS to individuals considering fertility treatment, in vitro fertilization (IVF), and/or gamete donation. This suggests that a one-size-fits-all approach to CS may not be optimal.

Cost is a further consideration. In this study, the data was compiled from different sized panels with variable costs, and with different molecular genetic technologies used for some disorders. Therefore, the cost per disorder screened could not be calculated. Economic modeling for a population representative of the United States indicated that, in contrast to genotyping, NGS-based CS could improve clinical outcomes and lower total health-care costs.^{29,30} A formal cost-benefit analysis for individual disease components in CS is complex, as it is based on decisions as to which diseases need to be screened, which is dependent on disease severity, frequency, and acceptability to the population. The incremental cost of adding disorders to an NGS-based CS panel can be small in terms laboratory expenses but there are additional costs associated with counseling and further testing and these need to be factored into the total cost analysis.

There are components in current prenatal screening guidelines that are not covered by the CS described in this study.⁴ This includes the recommendation for measuring serum hexosaminidase A levels or leukocytes for Tay-Sachs carrier testing in general populations for the possibility of identifying rare variants. It is unclear whether CS approaches are satisfactorily identifying these cases. A recent study demonstrated that for both AJ and non-AJ populations, clinical performance of NGS-based CS was better than enzymatic methods to screen for Tay-Sachs.³¹ Additional studies are needed to clarify whether the enzyme studies are still needed for rare variants. Similarly, screening for hemoglobinopathies has traditionally involved a complete blood count, red blood cell indices, reflex hemoglobin electrophoresis, and other testing if there is a suspicion of a hemoglobinopathy. It is assumed, but not yet demonstrated, that molecular genetic approaches can replace these traditional screening methods. Future professional guidance statements may need to consider whether the net overall gains in detection using contemporary molecular methods justify the transition from traditional screening technologies.

Conclusion

This study of 381,014 individuals screened for pathogenic variants with up to 274 genes evaluated reaffirms the view that the traditional CS approach based on ethnicity and family history merits re-examination, considering an increasingly diverse and multicultural population. NGS-based CS has the potential to impact the lives of at-risk couples in a way that current prenatal CS guidelines and traditional carrier screening have failed to achieve, i.e., by identifying large numbers of at-risk pregnancies. This, in turn, increases reproductive

options, opportunities for early intervention, and therapeutic management.³²

SUPPLEMENTARY INFORMATION

The online version of this article (<https://doi.org/10.1038/s41436-020-0807-4>) contains supplementary material, which is available to authorized users.

ACKNOWLEDGEMENTS

This study was funded by Natera, Inc. We would like to acknowledge Gabriel McNeill for statistical analysis performed, as well as Sheetal Parmar and Ramesh Hariharan, all from Natera, Inc., for intellectual contributions in the development of this manuscript.

DISCLOSURE

M.W., J.S., J.W., A.S., M.M., Z.P.D., M.R. and D.K.-K. are current employees and equity holders of Natera, Inc. S.A.P. is currently employed by GRAIL, Inc. P.B. is a paid consultant to Natera and holds stock options. C.M.E. is an employee of Baylor University. R.B. is a consultant and receives research support from Ferring Pharmaceuticals. L.W. declares no conflicts of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Hallam S, Nelson H, Greger V, et al. Validation for clinical use of, and initial clinical experience with, a novel approach to population-based carrier screening using high-throughput, next-generation DNA sequencing. *J Mol Diagn*. 2014;16:180–189.
- ACOG. ACOG committee opinion no. 442: preconception and prenatal carrier screening for genetic diseases in individuals of Eastern European Jewish descent. *Obstet Gynecol*. 2009;114:950–953.
- Gross SJ, Pletcher BA, Monaghan KG, Professional P, Guidelines C. Carrier screening in individuals of Ashkenazi Jewish descent. *Genet Med*. 2008;10:54–56.
- ACOG. Committee opinion no. 691: carrier screening for genetic conditions. *Obstet Gynecol*. 2017;129:e41–e55.
- Prior TW. Carrier screening for spinal muscular atrophy. *Genet Med*. 2008;10:840–842.
- Grody WW, Cutting GR, Klinger KW, et al. Laboratory standards and guidelines for population-based cystic fibrosis carrier screening. *Genet Med*. 2001;3:149–154.
- Perez AD, Hirschman C. The changing racial and ethnic composition of the US population: emerging American identities. *Popul Dev Rev*. 2009; 35:1–51.
- Kraft SA, Duenas D, Wilfond BS, Goddard KAB. The evolving landscape of expanded carrier screening: challenges and opportunities. *Genet Med*. 2019;21:790–797.
- Himes P, Kauffman TL, Muessig KR, et al. Genome sequencing and carrier testing: decisions on categorization and whether to disclose results of carrier testing. *Genet Med*. 2017;19:803–808.
- ACOG. American College of Obstetricians and Gynecologists committee opinion: number 690, March 2017. Carrier screening in the age of genomic medicine. *Obstet Gynecol*. 2017;129:595–596.
- Grody WW, Thompson BH, Gregg AR, et al. ACMG position statement on prenatal/preconception expanded carrier screening. *Genet Med*. 2013;15:482–483.
- Ben-Shachar R, Svenson A, Goldberg JD, Muzzey D. A data-driven evaluation of the size and content of expanded carrier screening panels. *Genet Med*. 2019;21:1931–1939.

13. Haque IS, Lazarin GA, Kang HP, Evans EA, Goldberg JD, Wapner RJ. Modeled fetal risk of genetic diseases identified by expanded carrier screening. *JAMA*. 2016;316:734–742.
14. Lazarin GA, Haque IS, Nazareth S, et al. An empirical estimate of carrier frequencies for 400+ causal Mendelian variants: results from an ethnically diverse clinical sample of 23,453 individuals. *Genet Med*. 2013;15:178–186.
15. Chokoshvili D, Vears D, Borry P. Expanded carrier screening for monogenic disorders: where are we now? *Prenat Diagn*. 2018;38:59–66.
16. Edwards JG, Feldman G, Goldberg J, et al. Expanded carrier screening in reproductive medicine—points to consider: a joint statement of the American College of Medical Genetics and Genomics, American College of Obstetricians and Gynecologists, National Society of Genetic Counselors, Perinatal Quality Foundation, and Society for Maternal–Fetal Medicine. *Obstet Gynecol*. 2015;125:653–662.
17. Watson MS, Cutting GR, Desnick RJ, et al. Cystic fibrosis population carrier screening: 2004 revision of American College of Medical Genetics mutation panel. *Genet Med*. 2004;6:387–391.
18. Origa R. Beta-thalassemia. *GeneReviews*. <https://www.ncbi.nlm.nih.gov/books/NBK1426/>. Accessed 25 July 2019.
19. Origa RMP. α -thalassemia. *GeneReviews*. <https://www.ncbi.nlm.nih.gov/books/NBK1435/>. Accessed 25 July 2019.
20. Yrigollen CM, Durbin-Johnson B, Gane L, et al. AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome. *Genet Med*. 2012;14:729–736.
21. Mila M, Alvarez-Mora MI, Madrigal I, Rodriguez-Revenga L. Fragile X syndrome: an overview and update of the FMR1 gene. *Clin Genet*. 2018; 93:197–205.
22. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405–424.
23. Luo M, Liu L, Peter I, et al. An Ashkenazi Jewish SMN1 haplotype specific to duplication alleles improves pan-ethnic carrier screening for spinal muscular atrophy. *Genet Med*. 2014;16:149–156.
24. Nolin SL, Brown WT, Glucksman A, et al. Expansion of the fragile X CGG repeat in females with premutation or intermediate alleles. *Am J Hum Genet*. 2003;72:454–464.
25. Beauchamp KA, Johansen Taber KA, Grauman PV, et al. Sequencing as a first-line methodology for cystic fibrosis carrier screening. *Genet Med*. 2019;21:2569–2576.
26. Palomaki GE, FitzSimmons SC, Haddow JE. Clinical sensitivity of prenatal screening for cystic fibrosis via CFTR carrier testing in a United States panethnic population. *Genet Med*. 2004;6:405–414.
27. US Census Bureau. Quick facts: United States. 2018. <https://www.census.gov/quickfacts/fact/table/US/RHI125218>. Accessed 10 October 2019.
28. Stevens B, Krstic N, Jones M, Murphy L, Hoskovec J. Finding middle ground in constructing a clinically useful expanded carrier screening panel. *Obstet Gynecol*. 2017;130:279–284.
29. Azimi M, Schmaus K, Greger V, Neitzel D, Rochelle R, Dinh T. Carrier screening by next-generation sequencing: health benefits and cost effectiveness. *Mol Genet Genomic Med*. 2016;4:292–302.
30. Beauchamp KA, Johansen Taber KA, Muzzey D. Clinical impact and cost-effectiveness of a 176-condition expanded carrier screen. *Genet Med*. 2019;21:1948–1957.
31. Cecchi AC, Vengoechea ES, Kaseniit KE, et al. Screening for Tay–Sachs disease carriers by full-exon sequencing with novel variant interpretation outperforms enzyme testing in a pan-ethnic cohort. *Mol Genet Genomic Med*. 2019;7:1–12.
32. Langlois S, Benn P, Wilkins-Haug L. Current controversies in prenatal diagnosis 4: pre-conception expanded carrier screening should replace all current prenatal screening for specific single gene disorders. *Prenat Diagn*. 2015;35:23–28.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, and provide a link to the Creative Commons license. You do not have permission under this license to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2020