

RESEARCH

Open Access



Comprehensive analysis of NGS-based expanded carrier screening and follow-up in southern and southwestern China: results from 3024 Chinese individuals

Qinlin Huang^{1†}, Juan Wen^{1†}, Hongyun Zhang¹, Yanling Teng¹, Wen Zhang¹, Huimin Zhu¹, Desheng Liang¹, Lingqian Wu^{1,2*} and Zhuo Li^{1*}

Abstract

Background This study aimed to screen southern and southwestern Chinese individuals using expanded carrier screening (ECS), which explores the carrier status of recessively inherited diseases in southern and southwestern China, evaluates the clinical effectiveness of ECS application, and helps recognize high-risk fetuses that may have genetic disorders early in pregnancy, to provide better reproductive guidance.

Methods ECS for 220 diseases based on next-generation sequencing was performed on 3024 southern and southwestern Chinese individuals (1512 couples). Carrier status was analyzed; genes and loci with high frequencies of variants and on high-risk couples (ARCs) were focused to evaluate the clinical utility of our ECS technology and provide them precise fertility guidance.

Results In total, Pathogenic/likely pathogenic(P/LP) variants were found in 1885 individuals, so the carrier frequency was 62.3%, and 23.2% of the individuals were carriers of multiple diseases. Furthermore, 2837 variants were detected, and the average number of P/LP variants carried per subject was 0.938. Additionally, 128 ARCs carried P/LP variants of the same gene, and the theoretical incidence rate in their offspring was as high as 2.12%.

Conclusion This study validated the application of our ECS technique for carrier screening in southern China, identifying carrier status and providing accurate carrier frequencies for hundreds of genetic diseases.

Keywords Carrier screening, Carrier frequencies, Single-gene disorders, Pathogenic variants, Recessive disease

[†]Qinlin Huang and Juan Wen contributed equally to this work.

*Correspondence:

Lingqian Wu
wulingqian@sklmg.edu.cn

Zhuo Li
lizhuo@sklmg.edu.cn

¹Center for Medical Genetics, Hunan Key Laboratory of Medical Genetics, MOE Key Lab of Rare Pediatric Diseases, School of Life Sciences, Central South University, Changsha 410000, China

²Laboratory of Molecular Genetics, Hunan Jiahui Genetics Hospital, Changsha, Hunan, China



© The Author(s) 2024. **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, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence 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 licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

Single-gene disorders are one of the most important causes of birth defects, and data show that the prevalence of single-gene genetic disorders is as high as 1/250 [1]; moreover, they account for 20% of the infant mortality and 18% of the pediatric hospitalization rates [2]. Among the Single-gene disorders, recessive monogenic diseases have received special attention because of their unpredictable characteristics. This is because if the parents are carriers of the same autosomal recessive condition or if the mother is a carrier of the X-chromosome recessive pathogenic variants, even if both parents do not have a clinical phenotype and there is no relevant family history of the genetic disease, there is still a higher risk (25%) of having offspring affected by the genetic disease. Moreover, the combined prevalence of recessive genetic disorders is 1/454, accounting for 61% of all monogenic genetic disorders [3]. Therefore, preventive screening, especially preconception/prenatal carrier screening (CS), is a preventive tool to target the source of birth defects, which can effectively prevent the first occurrence of genetic disorders, and thus prevent the occurrence of birth defects more effectively [4].

Carrier screening, a screening technique that involves genetic testing of prospective parents or couples in the early stages of pregnancy, can be traced back to 1971, when Kaback et al. performed CS for Tay-Sachs Disease in Jews, a study that ultimately led to a significant decrease in the incidence of Tay-Sachs Disease in Jews [5]. Since then, other recessive monogenic diseases have continued to be among the diseases for which CS is performed [6–8]. However, the number of diseases for which carriers were screened at that time was limited because of the inadequacy of testing technology, and it was difficult to conducting large-scale screening of disease carriers [9, 10].

With the development of sequencing technology at the beginning of the 21st century, it became possible to simultaneously screen carriers for multiple diseases. In 2009 and 2011, CS was expanded for the first time from the detection of a single disease to simultaneous screening of approximately 100 and 448 recessive monogenic diseases. Since then, single-disease CS has been replaced by the higher throughput and more cost-effective expanded carrier screening (ECS) [11, 12]. According to Schofield and Beauchamp, ECS reduces the birth rate of affected children more effectively and is more cost-effective than CS, as it screens individuals for several diseases [13, 14]. In recent years, there has been increasing interest in the use of ECS for population-based and preconception/prenatal CS. In 2013, Lazarin et al. screened 23,453 samples from a variety of ethnicities for carriers of more than 400 Mendelian genetic disorders, identifying 24% of the individuals as carriers of at least one of the

108 disorders [15]. In 2018, Zhao et al. conducted ECS testing of 10,476 Chinese couples with 11 common multiple recessive diseases 27.49% of which were found to be carriers of at least one of the 11 diseases. Another 255 couples (2.43%) were considered to be high-risk couples (ARCs) with a higher risk of having children [16].

There has been controversy regarding panel selection for ECS [17], and in Chau et al.'s reanalysis of exome and genome sequencing data of 1543 southern Chinese individuals, incremental detection was performed using the ECS panel. As a result, the increase in the threshold of carrier detection was nonlinearly related to the detection rate of high-risk couples, especially when the carrier condition of the included genes was extended from 1/200 to 1/1000, and the increase in the detection rate of high-risk couples became very weak [18]. In 2018, Chan et al. compared the selection of ECS panels from 16 laboratories in different regions; ultimately, only three disease disorders were simultaneously included in the panel of 16 laboratories, which showed that the selection of panels varies greatly among different laboratories [19]. In 2021, the American College of Medical Genetics and Genomics (ACMG) standardized the disease panel selection for ECS and made recommendations for panel selection for different populations, especially for all pregnant and planning to be pregnant populations recommending to provide carriers $\geq 1/200$ (including the X-linked inheritance patterns) of the disease [20]. Our study selected 220 panels of severely hyperprevalent diseases. To the best of our knowledge, this is the largest disease-scale ECS study conducted in China.

The incidence of birth defects in China is high and geographically specific [21], especially in previous studies, and it was found that the incidence of some diseases, such as thalassemia, was much higher in southern China [22]. However, the frequency of carriers in southern Chinese populations is still not well known, the differences between southern and southwestern carriers are not well defined, and previous studies on ECS have often lacked follow-up to assess the effectiveness of ECS application. Our study screened 1,512 couples in southern China who were in preparation for or during early pregnancy for 220 single-gene disorders with high morbidity, serious disability, and fatality and followed them up at any time to improve the population carrier frequency in southern China and to provide timely reproductive guidance and necessary medical interventions for ARCs. Ultimately, we expect to evaluate the effectiveness of the clinical application of ECS and provide experience and reference for its clinical application in China to effectively prevent and reduce birth defects.

Materials and methods

Subjects

A total of 3024 samples (1512 couples) were enrolled in this study from October 31, 2022, to June 3, 2023. Inclusion criteria for these couples required that they were either couples preparing for pregnancy or couples in early pregnancy ($\leq 13+6$ weeks) who signed informed consent and were followed up. It is worth noting that for generalizability and generality of this study, we excluded these couples: (a) couples who had conceived through assisted reproduction and couples who had conceived from non-autologous gametes; (b) couples with phenotypic abnormalities in one or both spouses, with a high degree of suspicion or diagnosis of hereditary disorders; and (c) pregnant women or couples who had undergone organ transplants, allogeneic transfusions within one year, or immunotherapy. We took a cautious approach in couples with a history of adverse pregnancy outcomes or a family history of genetic diseases. We advised these couples to undergo genetic diagnosis and screening for the etiology of the disease in their family lineage. If they insisted on participating, we enrolled them and fully informed them of the scope and limitations of this study, and that this study was not a substitute for genetic testing. Our subjects were recruited from various regions in southern and southwestern China, most of whom were from Changsha, Hunan Province (1352 samples) and Qianxinan, Guizhou Province (1122 samples); the exact sources of the samples are available in Supplementary Fig. S1. A questionnaire about ethnicity, attitudes toward ECS, and other relevant issues was provided to each subject at the pre-study participation stage, but this was not mandatory. All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of the Center for Medical Genetics, Central South University, Hunan, China (2021-1-26).

Panel design

After a preliminary study of regional morbidity in China and comprehensive consideration of economic effects, we selected 220 diseases for the panel (Supplementary Table S1). The selected diseases included those with high morbidity, serious disability, and fatality (220 monogenic diseases in female subjects and 194 monogenic diseases in male subjects. The main difference is that because we recruited healthy couples, male subjects were not tested for X-linked diseases). At the same time, our selection followed the ACMG's ECS panel recommendations for pregnant and planning to become pregnant people combined with the high prevalence of single-gene disorders in the Chinese region [20]. These diseases can be categorized by system as genetic-metabolic, ocular-auditory,

neuromuscular, dermatologic, hematologic, immunologic, urologic, endocrinologic, skeletal, syndromic, and other systems. As can be seen in the clinic, many diseases are screened in the metabolic system, which is the most prevalent system of genetic diseases. Detailed information on these diseases is available in Supplementary Table S1, with all OMIM numbers. Notably, the *GJB2*(NM_004004.6):c.109G>A(p.V37I) locus was reported only in the lineage of infertile females at the time of completion of the information collection form. The *UGT1A1*(NM_000463.3):c.211G>A(p.G71R) locus was not identified.

Next-generation sequencing (NGS) and detection of specific diseases

Target region capture combined with high-throughput sequencing

Our study used target region capture combined with NGS: the detection region was the coding sequence of the target gene and its neighboring ± 10 bp intronic regions, and some intronic variants that were judged to be causative in ClinVar and HGMD databases. The types of variants detected included single nucleotide variants, insertion/deletion variants within 10 bp, and partial gene deletion/duplication variants (including deletions/duplications at the exon level of the DMD gene, deletion of exon 7 of the SMN1 gene, and three deletions in *HBA1/HBA2* -SEA, - $\alpha 3.7$, and - $\alpha 4.2$). Testing for X-linked inheritance pattern genes was restricted to female participants.

Genomic DNA was extracted from peripheral blood samples using the MagaBio Plus General Genomic DNA Purification Kit II (Bioer Technology, Hangzhou, China) according to standard extraction methods. After fragmentation of the extracted DNA using a DNA library construction kit based on enzyme fragmentation (one-step method) (MatriDX, Changsha, Hunan), hybridization was performed through end repair, splicing, and PCR amplification. Pre-library was captured. The probe was captured by liquid hybridization with a pre-library using the Human Peripheral Blood Single Gene Genetic Disease Test Kit (Probe-Hybridization Capture Method) (MatriDX, Changsha, Hunan, China) and enriched with DNA from the target region to construct a target library. Exons were sequenced on the DNBSEQ-T7 platform (MGI-TECH, Shenzhen, China) in 100 bp paired-end reads, and variants were confirmed by Sanger sequencing (Tsingke Biotechnology Co., Ltd., Beijing, China).

Triplet-primed PCR (TP-PCR)

The DNA samples extracted from female subjects were amplified using Fragile X Syndrome CGG Repeat Count Assay Kit (MatriDX, Changsha, Hunan) and the products were stored at -20°C for further processing. A 3500 Dx Genetic Analyzer (Thermo Fisher Scientific, Waltham,

MA, USA) was used to detect the amplification of short fragment repeats by capillary electrophoresis and determine the number of short fragment repeats.

Long-range PCR(LR-PCR)

To detect the inversion of intron 1 or 22 of the *F8* gene, we performed LR-PCR amplification using the Hemophilia A *F8* Gene Test Kit (MatriDX, Changsha, Hunan). The amplification products were subjected to gel electrophoresis to detect inversions using an HE-120 high-throughput horizontal electrophoresis apparatus (Tanon, Shanghai, China).

Data and statistical analysis

Fastp and Bamdst software were used for the automated quality control, filtering, correction, and preprocessing of fastq documents. The third-party softwares GATK (<https://software.broadinstitute.org/gatk/>) and HaplotypeCall were used for variant calling. Candidate variants validated by high-throughput sequencing and special PCR techniques were selected for more profound annotation. To describe the candidate variants, the Guideline and standard of ACMG/Association for Molecular Pathology (AMP) [23] were used as a reference. Variant annotation and interpretation were conducted using the ANNOVAR software. Furthermore, clean reads are mapped to the human genome with the SMN2 gene masked. The read depths are then normalized using the median depth across the capture regions. The aggregate copy number of SMN1 and SMN2 is determined by analyzing the normalized read depth of SMN1, while the paralog-specific copy numbers are calculated by identifying and assessing the differentiating sites between SMN1 and SMN2. For HBA1 and HBA2, copy number estimation is achieved through a probe design that extends beyond the CDS regions, allowing for precise classification of α -thalassemia deletion types.

For the top 10 most common diseases and genes, we predicted the expected carrier counts and expected carrier frequency by reading and analyzing past literature on the gene (see Supplement for references). We attempted to determine whether the top 10 most common disease carrier frequencies in our study significantly differed from the carrier frequencies reported in the literature. A Chi-squared goodness-of-fit test was used to test the null hypothesis that the positive disease-specific carrier frequency in our study was equal to the expected positive carrier frequency reported in the literature. In addition, we employed the Benjamini-Hochberg multiple testing correction to correct the P-value. The adjusted P-values were compared with a significance level of 0.05 as a means to determining whether they were significantly different.

Principles and procedures of Questionnaire and follow-up

As mentioned above, we provided a questionnaire that will not be traced to individuals, through which we can obtain basic information about the subjects and their attitudes toward the ECS. After the start of the program, for the purpose of regular follow-up, we categorized the participants into three groups according to their status: nonpregnant, pregnant, and postpartum. For nonpregnant couples, follow-up visits were conducted every six months and focused on the planned mode of pregnancy and whether the original mode of pregnancy would change based on the ECS results. Families who were in early pregnancy at enrollment or had conceived during the program were followed up twice at 24–28 weeks of gestation and 42 days after the birth of the newborn, with clinical information about the pregnancy (results of obstetric and other genetic testing), pregnancy outcomes, and newborn screening results. For ARCs during pregnancy, increased the number of follow-up visits to ensure that they were aware of the risks of childbearing and advise them to undergo relevant prenatal diagnoses or seek guidance on reproduction.

Result

Population demographics

The mean age of the couples who participated in the program was 30.53 years, and the median age was 30 years. Of the 1,512 southern and southwestern Chinese couples, the majority came from Hunan and Guizhou, China, with 700 and 561 couples, respectively. The sample sizes of Guangxi and Guangdong were smaller (209 and 42 pairs, respectively); therefore, our study had a total of 996 pairs of southern couples and 561 pairs of southwestern couples (Supplementary Fig. S1). Since our questionnaire was not mandatory for all participants to complete, we obtained only 2,890 (95.57%) questionnaire results, the details of which are available in Supplementary Table S4. Among the couples who provided information from the questionnaire, excluding three who did not want to disclose their ethnicity, 16 different Chinese ethnic groups were included. Han individuals had the largest number of subjects with 2,501 (82.71%), whereas ethnic minority individuals accounted for a relatively small number of 386 (12.76%). However, it is worth noting that although there were fewer subjects from ethnic minorities, there was a sizable proportion of ethnic minorities among the subjects in the Guizhou and Guangxi regions. There were 212 (54.92%) and 69 (17.88%) participants, respectively, and the minority samples in these regions accounted for a large proportion of the total sample of participants in their regions (18.89% and 16.51%, respectively).

Table 1 Estimated burden of carriers in the study cohort

Number of disease(s) carried in a person	Number of samples	Carrier frequency (%)
0	1139	37.7
1	1182	39.1
2	535	17.7
3	136	4.5
4	23	0.8
5	9	0.3

The Table shows the estimated burden of carriers in the Southern and Southwestern Chinese population. Approximately 62.3% ($n=1885$) of individuals were carrier of at least one recessive condition

Table 2 Estimated burden of couples with different carrier status in the study cohort

Number of carriers in a couple	Number of samples	Carrier frequency (%)
0	211	14.0
1	721	47.7
2	580	38.4

The table shows the number of subjects who are carriers in a couple

Evaluation of Quality Metrics

In our cohort, all samples were subjected to capture enrichment and sequencing of the exonic and nearby regions of target genes. In all samples, the average mapped_reads was 15,811,070, and the mapped_reads rate was over 99.7%. The average sequencing depth of the target region was 504.81, the average target region coverage (1×coverage) was 99.94%, and the proportion of target region depth $\geq 20X$ sites was 99.76% on average. Detailed QC information is provided in Supplementary Table S2.

Carrier status of recessively inherited diseases in the Southern and Southwestern Chinese individuals

Among the 220 recessive genes, 931 were classified as pathogenic or likely pathogenic (P/LP; Supplementary Table S3). Candidate variants were found in 1885 subjects, and 2837 P/LP variants were reported, with an average of 0.938 P/LP variants per subject (no chimeras were found in our subjects). More than half of our subjects (62.3%) were carriers of at least one P/LP variant of a recessively inherited disease, that is, the P/LP variant carrier frequency was 62.3%. It was found that 703 (23.2%) were carriers of multiple recessively inherited diseases, of which 535 (17.7%) were carriers of two recessively inherited diseases, 136 (4.5%) were carriers of three recessively inherited diseases, 23 (0.8%) were carriers of four recessively inherited diseases, and nine (0.3%) were carriers of five recessively inherited diseases. (Table 1). Since our subjects were couples, we also counted the carrier status of the P/LP variants in each couple. Almost half of these couples were '1+0' carriers (one of the couples was a carrier and the other was a negative individual) (47.7%, $n=721$ couples), 211 couples were normal (both were negative individuals) (14.4%), and 580 were '1+1' type carriers (both couples were carriers of recessive P/LP variants) (38.4%) (Table 2).

The carrier frequencies for the top 15 most common conditions are listed in Table 3. In the Southern and Southwestern Chinese populations, the carrier frequencies of 10 conditions exceeded 2%, and two conditions exceeded 5%, including Deafness, Autosomal Recessive1A (DFNB1A) ($n=643$, 21.3%), and α -thalassemia ($n=224$, 7.4%). *GJB2*-associated DFNB1A is a recessive disease with the highest carrier frequency among all diseases, far exceeding that of other diseases. It was also three times higher than that in α -thalassemia ($n=224$,

Table 3 Top 15 conditions with the highest carrier rates in the Southern and Southwestern Chinese population

Condition	Gene	Count	Frequency (%)
Autosomal Recessive Deafness 1 A	<i>GJB2</i>	643	21.26
Alpha-thalassemia	<i>HBA1/HBA2</i>	224	7.41
Wilson Disease	<i>ATP7B</i>	94	3.11
Autosomal recessive progressive external ophthalmoplegia	<i>POLG</i>	81	2.68
Glucosephosphate dehydrogenase deficiency	<i>G6PD</i>	78	2.58
Phenylketonuria	<i>PAH</i>	75	2.48
Krabbe Disease	<i>GALC</i>	74	2.45
Autosomal Recessive Deafness 4, with Enlarged Vestibular Aqueduct	<i>SLC26A4</i>	73	2.41
Beta-thalassemia	<i>HBB</i>	64	2.12
Citrullinemia, type II, neonatal-onset	<i>SLC25A13</i>	62	2.05
classic congenital adrenal hyperplasia due to 21-hydroxylase deficiency	<i>CYP21A2</i>	59	1.95
Spinal Muscular Atrophy	<i>SMN1</i>	57	1.88
Systemic primary carnitine deficiency disease	<i>SLC22A5</i>	56	1.85
Usher syndrome type 2	<i>USH2A</i>	49	1.62
2-methylbutyrylglycinuria	<i>ACADSB</i>	43	1.42
Glycogen Storage Disease Type II	<i>GAA</i>	43	1.42

7.9%) and seven times higher than that in Wilson’s disease ($n=96$, 3.4%). It is worth noting that nine of the 15 recessively inherited diseases with the highest carriage frequencies were metabolic system diseases; as such, metabolic system diseases should be considered essential in the ECS panel.

The top 10 most common variants in the southern and southwestern Chinese populations and the percentage of these variants in the total variants of their genes are shown in Fig. 1. The most common variant in our cohort was *GJB2* c.109G>A p.V37I, which was associated with *DFNB1A* and was found in 557 subjects (18.4%). A total of eight variants had a carrier frequency higher than 1%, with two of the *HBA1/HBA2* having a carrier frequency higher than 2%, and both were in the top five of the carrier frequency, $-\alpha3.7$ (3.1%) and $-\text{SEA}$ (2.3%), respectively. Among the 10 most common variants, the following accounted for the majority (>50%) of all variants detected in this gene: *GJB2* c.109 ($n=557$, 86.6%), *POLG* c.2890 C>T ($n=79$, 97.5%), *SMN1* exon-7del ($n=57$, 100%), *GALC* c.1901T>C ($n=55$, 74.3%), *SLC25A13* c.852_855del ($n=40$, 64.5%). In addition, the two previously mentioned variation loci *HBA1/HBA2* $-\alpha3.7$ and *HBA1/HBA2* $-\text{SEA}$ together also accounted for the majority in *HBA1/HBA2* ($n=163$, 72.8%).

Analysis of test results for specific diseases

In our cohort, some specific diseases could not be screened for variations by NGS. We identified these variations by TP-PCR combined with capillary electrophoresis or LR-PCR combined with gel electrophoresis; as such, these results have been further separately analysed. Of the 3024 individuals screened, none carried intron 1 or intron 22 variants of the *F8* gene. A total of 1512 women were screened for *FMRI* CGG repeat length; none of them had >200 CGGs (pathogenic variant carriers), two (0.13%) had 55–200 CGGs (premutation carriers), and 12 (0.79%) had intermediate (45–55 CGGs) (Supplementary Table S5). For women carrying premutation mutations, not every pregnancy occurs with an increase in the copy number of CGG triplet repeat sequences, but all premutation mutations are at risk of a large increase in copy number to become full mutations during inheritance; therefore, we provided risk alerts for these two subjects.

Comparative analysis of carrying frequencies between the Southern and Southwestern regions of China

We categorized 3024 subjects into 1122 from southwestern China and 1902 from southern China, according to their detailed geographic locations. We analyzed the detected variants by category and the results are presented in Table 4. First, the carrier frequency and number of variants per subject in southwestern China were

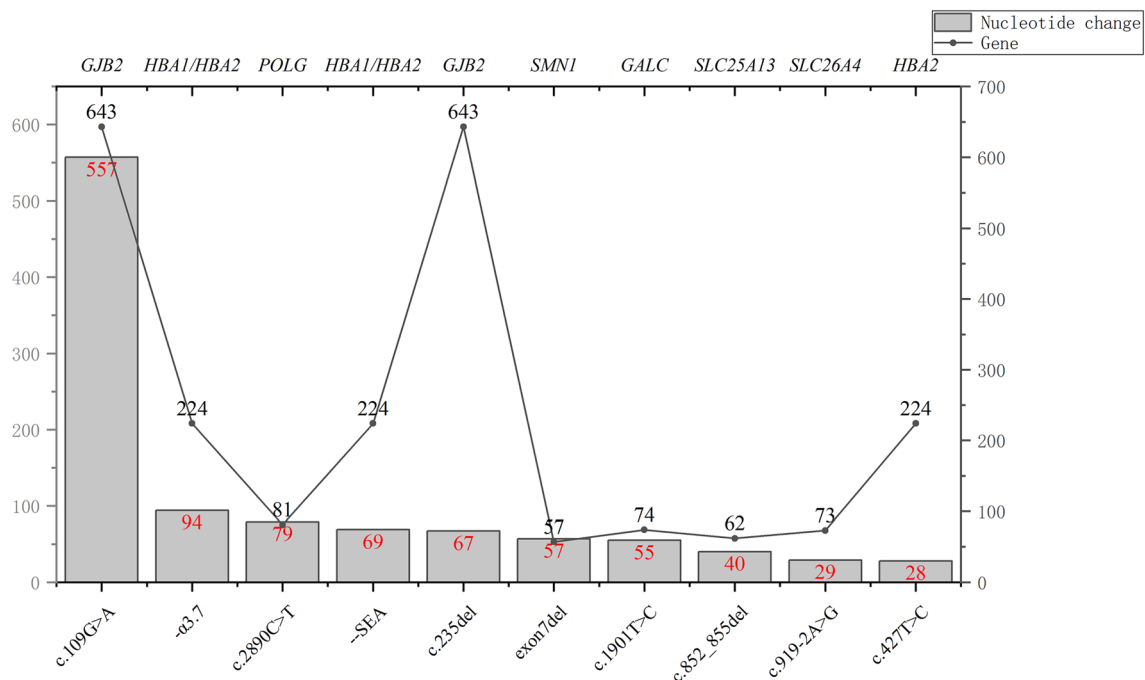


Fig. 1 Top 10 most common P/LP variants and the total number of variants in the corresponding genes in the Southern and southwestern Chinese populations. The data in the bar graph show the number of the top 10 most common P/LP variants, and the data in the dotted line graph show the total number of variants in the corresponding genes for these variants. The lower X-axis represents the top 10 most common P/LP variants, and the upper X-axis indicates the genes corresponding to the variants

Table 4 Top 10 conditions with the highest carrier rates in southern China and southwestern China, respectively. Diseases with large differences in carrier frequencies between southern and southwestern China were targeted using different colors

Southwestern				Southern			
Condition	Gene	Count	Frequency (%)	Condition	Gene	Count	Frequency (%)
Autosomal Recessive Deafness 1A	<i>GJB2</i>	250	0.222816	Autosomal Recessive Deafness 1A	<i>GJB2</i>	393	0.206625
α -thalassemia	<i>HBA1/HBA2</i>	101	0.090018	α -thalassemia	<i>HBA1/HBA2</i>	123	0.064669
Wilson Disease	<i>ATP7B</i>	47	0.041889	Autosomal recessive progressive external ophthalmoplegia	<i>POLG</i>	61	0.032072
Krabbe Disease	<i>GALC</i>	30	0.026738	Glucosephosphate dehydrogenase deficiency	<i>G6PD</i>	49	0.025762
Glucosephosphate dehydrogenase deficiency	<i>G6PD</i>	29	0.025847	Phenylketonuria	<i>PAH</i>	48	0.025237
β -thalassemia	<i>HBB</i>	28	0.024955	Autosomal Recessive Deafness 4, with Enlarged Vestibular Aqueduct	<i>SLC26A4</i>	48	0.025237
Phenylketonuria	<i>PAH</i>	27	0.024064	Wilson Disease	<i>ATP7B</i>	47	0.024711
Autosomal Recessive Deafness 4, with Enlarged Vestibular Aqueduct	<i>SLC26A4</i>	25	0.022282	classic congenital adrenal hyperplasia due to 21-hydroxylase deficiency	<i>CYP21A2</i>	45	0.023659
Citrullinemia, type II, neonatal-onset	<i>SLC25A13</i>	24	0.02139	Krabbe Disease	<i>GALC</i>	44	0.023134
2-methylbutyrylglycinuria	<i>ACADSB</i>	21	0.018717	Citrullinemia, type II, neonatal-onset	<i>SLC25A13</i>	38	0.019979

slightly higher than in southern China; however, they were almost the same. In terms of the top 10 most common conditions, there were subtle differences between the southern and southwestern regions. In these 10 conditions, the sum of the frequencies in the southwestern part of China was not much different from that in the southern part, which is 51.9% and 47.1%, respectively, suggesting that recessively inherited diseases carried by the subjects in both the southern and southwestern regions of China are not dispersed. Although the top two diseases carried in both regions were *GJB2*-associated DFNB1A and *HBA1/HBA2*-associated Alpha-thalassemia, the frequency of both in the southwest was approximately 2% higher than in the south. Of interest were *ATP7B*-associated Wilson's disease and *HBB*-associated β -thalassemia, which were present at considerably higher frequencies in the southwestern subjects than in the southern subjects. In contrast, individuals in Southern China had significantly higher frequencies of mutation carriage for *POLG*-associated Autosomal recessive progressive external ophthalmoplegia and *CYP21A2*-associated classical congenital adrenal hyperplasia due to 21-hydroxylase deficiency than those in southwestern China. This suggests that the carrier statuses in southern and southwestern China have their own characteristics, with certain genes having significantly higher carrier frequencies than in the overall population.

Comparison to previously published carrier frequencies

We compared the carrier frequencies of the top 10 most common diseases in southwestern China and southern China obtained from our CS with those from previous CS studies, and the results are shown in Table 5 (the literature relied upon to derive the expected carrier frequencies of the Top 10 most common conditions can be found in Additional File S2). The expected carrier frequencies were derived from the combined predictions of previous CS studies, and after a chi-squared goodness-of-fit test, the observed carrier frequencies were statistically different from the expected carrier frequencies for the six diseases. The disease with the largest difference in carrier frequency was *HBA1/HBA2*-associated α -thalassemia, which nearly doubled in frequency. In addition, eight of the top 10 most common diseases listed in Table 4 had observed carrier frequencies greater than the expected carrier frequencies. Only two diseases, *GJB2*-associated DFNB1A and *G6PD*-associated glucose-phosphate dehydrogenase deficiency, had observed carrier frequencies that were slightly lower than the expected carrier frequencies, and there was no statistically significant difference between the observed and expected carrier frequencies of *GJB2*.

At-risk couples and affected theoretical pregnancies

When both individuals carry variants of the same recessive gene, the likelihood of giving birth to a child with a recessive disease is as high as 25%. A total of 210 (6.94%)

Table 5 Top 10 conditions with the highest observed carrier frequencies with corresponding observed carrier frequency and expected carrier frequency

Condition	Gene	Expected carrier frequency	Observed carrier frequency	Statistically significant carrier frequency difference
Autosomal Recessive Deafness 1 A	<i>GJB2</i>	22.71	21.26	NO
Alpha-thalassemia	<i>HBA1/HBA2</i>	4.27	7.41	YES
Wilson Disease	<i>ATP7B</i>	1.75	3.11	YES
Autosomal recessive progressive external ophthalmoplegia	<i>POLG</i>	1.30	2.68	YES
Glucosephosphate dehydrogenase deficiency	<i>G6PD</i>	3.70	2.58	YES
Phenylketonuria	<i>PAH</i>	1.87	2.48	YES
Krabbe Disease	<i>GALC</i>	1.56	2.45	YES
Autosomal Recessive Deafness 4, with Enlarged Vestibular Aqueduct	<i>SLC26A4</i>	2.22	2.41	NO
Beta-thalassemia	<i>HBB</i>	1.79	2.12	NO
Citrullinemia, type II, neonatal-onset	<i>SLC25A13</i>	1.61	2.05	NO

The literature used to predict expected carrier frequency is listed in Supplementary Data S1

The Chi-squared goodness-of-fit test was used to test the null hypothesis that the observed carrier frequency for a specific gene in the patient cohort was equal to the expected carrier frequency

individuals from all the southern subjects carried the same autosomal recessive genes as their partners. However, we excluded some couples from the ARCs because of the relatively smaller adverse risk to the offspring when they jointly carried the *GJB2* c.109G>A variant, as described in the methods. Thus, there were 51 pairs of autosomal recessive ARCs (3.44%), of which 18 pairs were deafness-associated (*GJB2*), 19 pairs were thalassemia (*HBA1/HBA2* or *HBB*)-associated, and 15 pairs were other systems (cases 22G00284093 and 22G00284094 were both deafness-associated ARCs and other systems ARCs). In addition, there were 77 female subjects carrying X-linked recessive gene variations who were at risk of causing the birth of an affected son (1/4) (cases 23G00040615 and 23G00040616 were both thalassemia-associated ARCs and X-linked recessive disease-associated ARCs). Thus, our ARCs totaled 127 ARCs (51 autosomal recessive and 77 X-linked recessive pairs, with one pair having both autosomal recessive and X-linked recessive disease-associated ARCs). Of the 127 ARCs, 47 (19 autosomal recessive and 28 X-linked recessive) were southwestern couples, and 80 (32 autosomal recessive and 49 X-linked recessive) were southern couples (one couple had both autosomal recessive and X-linked recessive disease-associated ARCs) (Supplementary Table S6). The frequencies of ARCs in southern and southwestern China were 8.38% and 8.41%, respectively, indicating that there was little difference in the frequencies of ARCs between southern and southwestern China. In all the subjects, the theoretical number of offspring affected without intervention was 2.10%.

Follow-up results and evaluation of the effectiveness of ECS application

After all the reports were distributed to the subjects, we followed up with 1512 couples in three groups, preparation for pregnancy, pregnancy, and postpartum, to study the effectiveness of ECS. We received 1,422 follow-ups, with 1,031, 263, and 128 couples in the three groups, respectively, for a follow-up rate of 94.1%. Among the nonpregnant couples, 409 (39.7%) had no current pregnancy plans, 597 (57.9%) chose natural conception to have offspring, and 25 (2.4%) switched from natural conception to preimplantation genetic testing for monogenic (PGT-M). Of the pregnant couples, 245 chose to continue the pregnancy, 10 couples had spontaneous abortions, and eight couples chose induced abortions (due to the detection of chromosomal disorders or monogenic disorders). Of the 14 pregnant couples with fetal abnormalities (detected by ultrasound, non-invasive, and other prenatal tests), nine couples were tested for genetic abnormalities, and four couples had positive results. Among the postnatal couples, 64 (50%) underwent postnatal genetic testing, of which seven had positive results. Fortunately, six of these couples had newborns who were carriers of the causative gene rather than patients, and the remaining couple had a newborn with *GJB2*, c.109G>A purity who passed the hearing screening.

Discussion

This study presents the ECS results for 220 diseases from 3,024 individuals (1512 couples) in southern and southwestern China. This was the first ECS study in which the groups were divided into southern and southwestern China for comparison. Thus, this data provides the first report of carrier frequencies for many rare diseases across southern and southwestern China, which have

high frequencies of recessively inherited diseases and is an important resource for guiding the diagnosis, treatment, and prevention of Mendelian disease.

Interesting information, such as history of abnormal pregnancy, knowledge, and attitude toward ECS, was obtained through a non-compulsory questionnaire (Supplementary Table S1). Out of the 2890 questionnaires received, 2178 had no history of unexplained spontaneous abortion and/or fetal anomalies, 467 (15.44%) had a history of one adverse pregnancy, and 245 (8.10%) had a history of two or more adverse pregnancies, which shows that there was not an insignificant number of adverse pregnancies among our samples (23.54%). Despite the high percentage of 2587 (85.55%) who were highly educated, 1042 (34.46%) had not heard of ECS at all, and only 603 (19.94%) were aware of ECS. This result proves that the ECS is rarely mentioned in the Chinese population and lacks further publicity and popularization. Of the 2872 people willing to undergo ECS, 1658 (57.7%) were willing to obtain all information on the carrier status of genetic diseases for both individuals. There were still 385 (13.4%) individuals who only wanted to know their own genetic disease carrier status or whether both individuals were carriers of the same recessive genetic disease. In addition, 829 individuals did not have any opinion on this. Therefore, further investigation is needed before implementing ECS in a clinical setting to ensure the efficiency of CS while respecting subjects' wishes.

In our study, the differences in the carrier frequencies of the top 10 most common diseases were obtained by comparing the carrier rates of the top 10 most common diseases in southern and southwestern China. Overall, the top 10 most common diseases in southern and southwestern China were similar, but individual diseases were geographically specific. *ATP7B*-associated Wilson Disease and *HBB*-associated β -thalassemia were significantly more prevalent in the southwest Chinese study population. *POLG*-associated Autosomal recessive progressive external ophthalmoplegia and *CYP21A2*-associated classical congenital adrenal hyperplasia caused by 21-hydroxylase deficiency was significantly more prevalent in the southern Chinese population. Although our study population was small, it can still serve as a basis for characterizing the disease carriage rates in southern and southwestern China.

We compared the carrier frequencies of common diseases obtained in our study with the expected carrier frequencies estimated in the literature, which were slightly higher than the overall expected carrier frequencies (Table 4). The highest carrier frequency, *GJB2*-associated DFNB1A, was detected at 22.28%, which was not significantly different from the carrier frequency obtained in previous CS studies conducted in Guangdong and Hong Kong, China [18, 19, 24]. The observed

carrier frequencies for the other Southern Chinese pan-ethnic disease-associated genes, *HBA1/HBA2*, *ATP7B*, and *POLG*, were nearly twice as high as the expected carrier frequencies, which may be attributed to the fact that the number of ethnic minorities in the samples we collected was large, and some of them lived in the same village. Lazarin et al. screened 12,915 Northwestern Europe in 2012 through ECS for 400+ diseases, and the diseases that resulted in higher carrier rates were α -1-Antitrypsin deficiency, Cystic fibrosis, DFNB1, Spinal muscular atrophy, and Smith-Lemli-Opitz syndrome [15]. In this study and ours, only two diseases, DFNB1A and SMA, had high carrier frequencies in the target populations, suggesting that there is considerable racial variability in disease carrier rates. In China, Zhao et al. conducted a multicenter CS study of 10,476 Chinese couples for 11 recessive diseases, which resulted in higher carrier frequencies for α -thalassemia, β -thalassemia, Phenylketonuria (PKU), Wilson's disease, DFNB1A [16]. Compared with our study, the carrier frequencies of α -thalassemia, DFNB1A, and Wilson's disease were generally higher, indicating that these three diseases generally have higher carrier frequencies in the Chinese population. However, the carrier frequencies of β -thalassemia and PKU could not be ranked in the top five in this study, so the carrier frequencies of single-gene disorders also varied somewhat among the same ethnicity and different geographic regions. Therefore, the inclusion criteria for the ECS panel should consider the prevalence of monogenic diseases in different races, regions, and populations.

According to the 2021 ACMG recommendations, all couples planning a pregnancy should receive CS for diseases with a carrier frequency of $\geq 1/200$ (including X-linked pattern disorders) [20]. In this guideline, the authors provide 97 autosomal recessive disorders for which screening is recommended. However, there are still other disorders in our study with carrier frequencies significantly higher than 1/200 that are not among the 97 disorders for which the ACMG recommends screening. Importantly, six diseases that were among the top 15 most common conditions in our study were not among the 97 diseases: *G6PD*-associated Glucose-6-phosphate dehydrogenase deficiency; *GALC*-associated Krabbe disease; *SLC25A13*-associated citrullinemia type II, neonatal onset; *SMNI*-associated spinal muscular atrophy, type I; *SLC22A5*-associated deficiency of Systemic primary carnitine deficiency disease; and *ACADSB*-associated 2-methylbutyrylglycinuria. In our study, all diseases had a carrier frequency higher than 1.5% (after rounding). Therefore, in the ECS in southern and southwestern China, different panels should be selected according to local disease carriage characteristics, rather than referring exclusively to the ACMG disease recommendations.

In NGS-based carrier screening studies, the detection of copy number variations (CNV) is often difficult and lacks accuracy. Currently, CNV can be assessed using bioinformatics tools; for example, *SMN1* deletion variation can be predicted by SMAca or SMN copy number callers [25, 26]. Although these optimized bioinformatics tools can help us obtain results beyond the standard pipeline, their accuracy and sensitivity are insufficient; therefore, we jointly used multiple raw-fiducial algorithms and tools to compensate for the possible shortcomings. In addition, we tested other complex diseases, such as fragile X-chromosome disorders and *F8* gene inversion-related diseases, using LR-PCR combined with agarose gel electrophoresis and TP-PCR combined with capillary electrophoresis. This can be a more effective solution for diseases that are difficult to detect by NGS and improve the comprehensiveness and accuracy of CS.

In our study, 127 couples (8.40%) were diagnosed with ARCs: 99 couples in preparation for pregnancy, and 28 couples during pregnancy or postpartum. Among couples who were already pregnant, ten couples carried autosomal recessive P/LP variants of the same gene and 18 couples carried X-linked recessive P/LP variations of the same gene. For all ARCs who were planning to become pregnant but were not currently pregnant, we fully informed them of their pregnancy risks and followed them regularly to determine their pregnancy status. We advised all pregnant ARCs to choose appropriate prenatal screening strategies, such as non-invasive prenatal testing and amniocentesis genetic testing, and followed them up continuously. Among all pregnant ARCs, there was one case of induced abortion due to the presence of a child with thalassemia, and the other ARCs did not undergo prenatal genetic testing; the test results were negative, or the fetus was a carrier of P/LP variants rather than a patient. Surprisingly, during our follow-up, a case involving pregnant parents (family number: JX22G00284121) carrying the *GJB2* c.109G>A variant resulted in a normal fetus during labor and delivery. However, the 5-year-old child failed hearing screening at birth, with a hearing test suggesting a low threshold. This suggests that the previous child likely received the *GJB2* c.109G>A purity variant, highlighting the importance of prenatal diagnosis. It is evident that the ECS may help couples with a history of abnormal pregnancies hypothesize the etiology and further help them with fertility guidance.

Our study has some limitations. First, our study was based on NGS with target sequence capture; therefore, we could not accurately detect variants in deep non-coding regions. Meanwhile, detection sensitivity may be seriously affected by the low probe capture efficiency of NGS for complex sequences (high GC content, complex structure, etc.). Our detection of the number of CGG

repeats in *FMR1* may have had a small error and we were unable to detect AGG disruptions, which could not further reveal the mutation risk of the subjects. Second, the causative genes of 21-hydroxylase deficiency (21-OHD), *CYP21A2* and *CYP21A1P* are true and false genes, and high sequence similarity between them promotes gene conversion. Because of the high sequence similarity between the true and false genes, the probe captured *CYP21A2* along with *CYP21A1P*. Although the experiment had a separate design for the PCR enrichment capture of the *CYP21A2* gene, it still resulted in false-positive and false-negative results. In addition, for the detection of exon seven deletion variants of *SMN1*, we were unable to detect silent (2+0) SMA carriers, that is, individuals with two copies of *SMN1* on one chromosome (duplication allele) and zero copies on the other (deletion allele) [27]. In previous studies, silent (2+0) carriers were predicted using *SMN1* g.27,134 T>G and g.27706_27707delAT SNP loci [28]. However, the frequency of silent (2+0) carriers is low in Asian populations [25], and it has been reported that these two SNP variants were not found in 10 cases of identified silent (2+0) carriers or 18 cases of suspected silent (2+0) carriers [29]. Therefore, we did not adopt the polymorphic loci g.27134T>G and g.27706_27707delAT as markers that could significantly improve the detection of silent (2+0) SMA carriers, because they may have a relatively low proportion in the Chinese population. Finally, the number of provinces and participants included in our study was relatively small, and more provinces and people are necessary in the future studies to enhance the relevance and effectiveness of carrier screening in the identification and management of genetic diseases in China's ethnic minorities.

We received 1422 follow-ups from our participants, resulting in a high follow-up rate (94.1%). Twenty-five couples chose to switch to unnatural conception methods such as PGT-M to conceive their offspring, proving that our ECS study is a good indicator of fertility risk in couples with ARCs. In addition, as a result of participating in our study, nine couples who had fetal abnormalities detected during pregnancy chose to undergo genetic testing, effectively identifying fetuses with possible recessive pure mutations that cause disease. According to our follow-up surveys, not all couples who participated in our extended CS had children with panel-related recessive genetic disorders since their participation in the program to date. To the best of our knowledge, many current CS studies do not include follow-up surveys, which may make CS their effectiveness. A follow-up survey would allow subjects to fully understand and pay attention to their carrier status, thus reducing the risk of ARCs giving birth to fetuses with genetic diseases. It can also be a good way to test the effectiveness of the application of

ECS studies and evaluate various aspects of panel selection and implementation for clinical applications.

Conclusion

This study evaluated the clinical utility of this CS strategy in southern China by screening 3024 individuals in southern and southwestern China for 220 disease carriers using ECS. It was found that more than half of the couples of childbearing age were carriers for any recessive condition (62.3% of the population), with 0.938 P/LP variants per capita, and that approximately 1 in 12 couples had ARCs for at least one recessive disease (1/4 probability of affecting offspring), which can provide a database for the development of panel inclusion and screening criteria for conducting ECS in southern and southwestern China, as well as experience and reference for clinical application.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40246-024-00680-y>.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

We thank all subjects and team members for their participation in this study.

Author contributions

QLH and JW helped in conceptualisation and writing; HYZ, YLT helped in collection of clinical data; QLH, LW, and WZ performed the experiment; QLH, HMZ, and DSL helped in data analysis; ZL and JW helped in design the study; ZL and LQW helped in revise the manuscript. All authors agree to the published version of the manuscript.

Funding

This work was supported by grants from the National Key R&D Program of China (2021YFC1005300, 2022YFC2703702, and 2022YFC2703703), the National Natural Science Foundation of China (82171711 and 82371724), the Hunan Provincial Natural Science Foundation of China (2023JJ30725), the Science and Technology Innovation Program of Hunan Province (2019SK1014), and the Open Research Funds of the State Key Laboratory of Ophthalmology (303060202400382).

Data availability

The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of the Center for Medical Genetics, Central South University, Hunan, China (2021-1-26). Informed consent was obtained from all subjects involved in the study.

Consent for publication

All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Received: 21 May 2024 / Accepted: 27 September 2024

Published online: 08 October 2024

References

1. Evans MI, Wapner RJ, Berkowitz RL. Noninvasive prenatal screening or advanced diagnostic testing: caveat emptor. *Am J Obstet Gynecol*. 2016;215(3):298–305.
2. Kingsmore S. Comprehensive carrier screening and molecular diagnostic testing for recessive childhood diseases. *PLoS Curr* 2012;2(4):e4f9877ab9878ffa9879.
3. Baird PA, Anderson TW, Newcombe HB, Lowry RB. Genetic disorders in children and young adults: a population study. *Am J Hum Genet*. 1988;42(5):677–93.
4. Henneman L, Borry P, Chokoshvili D, Cornel MC, van El CG, Forzano F, Hall A, Howard HC, Janssens S, Kayserili H, et al. Responsible implementation of expanded carrier screening. *Eur J Hum Genet*. 2016;24(6):E1–12.
5. Kaback M, Lim-Steele J, Dabholkar D, Brown D, Levy N, Zeiger K. Tay-Sachs disease—carrier screening, prenatal diagnosis, and the molecular era. An international perspective, 1970 to 1993. The international TSD data Collection Network. *JAMA*. 1993;270(19):2307–15.
6. Genetics, ACo. ACOG committee opinion. Number 298, August 2004. Prenatal and preconceptional carrier screening for genetic diseases in individuals of Eastern European Jewish descent. *Obstetrics and gynecology* 2004, 104(2):425–428.
7. Ioannou L, McClaren BJ, Massie J, Lewis S, Metcalfe SA, Forrest L, Delatycki MB. Population-based carrier screening for cystic fibrosis: a systematic review of 23 years of research. *Genet Med*. 2014;16(3):207–16.
8. Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, Lok S, Plavsic N, Chou JL. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Sci (New York NY)*. 1989;245(4922):1066–73.
9. Committee Opinion No. 691: carrier screening for genetic conditions. *Obstet Gynecol*. 2017;129(3):e41–55.
10. Sparks TN. Expanded carrier screening: counseling and considerations. *Hum Genet*. 2020;139(9):1131–9.
11. Bell CJ, Dinwiddie DL, Miller NA, Hateley SL, Ganusova EE, Mudge J, Langley RJ, Zhang L, Lee CC, Schilkey FD, et al. Carrier testing for severe childhood recessive diseases by Next-Generation sequencing. *Sci Transl Med*. 2011;3(65):14.
12. Srinivasan BS, Evans EA, Flannick J, Patterson AS, Chang CC, Pham T, Young S, Kaushal A, Lee J, Jacobson JL, et al. A universal carrier test for the long tail of mendelian disease. *Reprod Biomed Online*. 2010;21(4):537–51.
13. Beauchamp KA, Taber KAJ, Muzzey D. Clinical impact and cost-effectiveness of a 176-condition expanded carrier screen. *Genet Med*. 2019;21(9):1948–57.
14. Schofield D, Lee EVY, Parmar J, Kelly S, Hobbs M, Laing N, Mumford J, Shrestha R. Economic evaluation of population-based, expanded reproductive carrier screening for genetic diseases in Australia. *Genet Med*. 2023;25(5):14.
15. Lazarin GA, Haque IS, Nazareth S, Iori K, Patterson AS, Jacobson JL, Marshall JR, Seltzer WK, Patrizio P, Evans EA, 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(3):178–86.
16. Zhao SM, Xiang JL, Fan CN, Asan, Shang X, Zhang XH, Chen Y, Zhu BS, Cai WW, Chen SK, et al. Pilot study of expanded carrier screening for 11 recessive diseases in China: results from 10,476 ethnically diverse couples. *Eur J Hum Genet*. 2019;27(2):254–62.
17. Westemeyer M, Saucier J, Wallace J, Prins SA, Shetty A, Malhotra M, Demko ZP, Eng CM, Weckstein L, Boostanfar R, et al. Clinical experience with carrier screening in a general population: support for a comprehensive pan-ethnic approach. *Genet Med*. 2020;22(8):1320–8.
18. Chau JFT, Yu MHC, Chui MMC, Yeung CCW, Kwok AWC, Zhuang XH et al. Comprehensive analysis of recessive carrier status using exome and genome sequencing data in 1543 Southern Chinese. *Npj Genomic Med*. 2022;7(1):9.
19. Chan OYM, Leung TY, Cao Y, Shi MM, Kwan AHW, Chung JPW, Choy KW, Chong SC. Expanded carrier screening using next-generation sequencing of 123 Hong Kong Chinese families: a pilot study. *Hong Kong Med J*. 2021;27(3):177–83.
20. Gregg AR, Aarabi M, Klugman S, Leach NT, Bashford MT, Goldwaser T, Chen E, Sparks TN, Reddi HV, Rajkovic A, et al. Screening for autosomal recessive and X-linked conditions during pregnancy and preconception: a practice

- resource of the American College of Medical Genetics and Genomics (ACMG). *Genet Med.* 2021;23(10):1793–806.
21. Dai L, Zhu J, Liang J, Wang YP, Wang H, Mao M. Birth defects surveillance in China. *World J Pediatr.* 2011;7(4):302–10.
 22. Yang Y, Zhang J. Research Progress on Thalassemia in Southern China -Review. *Zhongguo Shi Yan xue ye xue Za Zhi.* 2017;25(1):276–80.
 23. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, 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(5):405–24.
 24. Xi YP, Chen GQ, Lei CX, Wu JP, Zhang S, Xiao M, Zhang WB, Zhang YP, Sun XX. Expanded carrier screening in Chinese patients seeking the help of assisted reproductive technology. *Mol Genet Genom Med.* 2020;8(9):13.
 25. Chen X, Sanchis-Juan A, French CE, Connell AJ, Delon I, Kingsbury Z, Chawla A, Halpern AL, Taft RJ, Bentley DR, et al. Spinal muscular atrophy diagnosis and carrier screening from genome sequencing data. *Genet Med.* 2020;22(5):945–53.
 26. Lopez-Lopez D, Loucera C, Carmona R, Aquino V, Salgado J, Pabalos S, Miranda M, Alonso A, Dopazo J. *SMN1* copy-number and sequence variant analysis from next-generation sequencing data. *Hum Mutat.* 2020;41(12):2073–7.
 27. Gitlin JM, Fischbeck K, Crawford TO, Cwik V, Fleischman A, Gonye K, Heine D, Hobby K, Kaufmann P, Keiles S, et al. Carrier testing for spinal muscular atrophy. *Genet Med.* 2010;12(10):621–2.
 28. Luo MJ, Liu L, Peter I, Zhu J, Scott SA, Zhao GP, Eversley C, Kornreich R, Desnick RJ, Edelman L. An Ashkenazi jewish *SMN1* haplotype specific to duplication alleles improves pan-ethnic carrier screening for spinal muscular atrophy. *Genet Med.* 2014;16(2):149–56.
 29. Yanyan C, Miaomiao C, Fang S, Yujin Q, Jinli B, Hong WJYH. Familial study of spinal muscular atrophy carriers with *SMN1* (2+0) genotype. *Yi Chuan* 2021;43(2):160–8.

Publisher's note

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