

## Carrier frequency estimation of pathogenic variants of autosomal recessive and X-linked recessive Mendelian disorders using exome sequencing data in 1,642 Thais

Wanna Chetruengchai<sup>1,2</sup>, Prasit Phowthongkum<sup>1,3\*</sup>, Vorasuk Shotelersuk<sup>1,2</sup>

<sup>1</sup>Excellence Center for Genomics and Precision Medicine, King Chulalongkorn Memorial Hospital, the Thai Red Cross Society, Bangkok, 10330, Thailand

<sup>2</sup>Center of Excellence for Medical Genomics, Medical Genomics Cluster, Department of Pediatrics, Faculty of Medicine, Chulalongkorn University, Bangkok, 10330, Thailand

<sup>3</sup>Division of Medical Genetics and Genomics, Department of Medicine, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

\*Corresponding author

E-mail: [Prasit.Ph@chula.ac.th](mailto:Prasit.Ph@chula.ac.th) (PP)

### Abstract

People with autosomal recessive disorders often were born without awareness of the carrier status of their parents. The American College of Medical Genetics and Genomics (ACMG) recommends screening 113 genes known to cause autosomal recessive and X-linked conditions in couples seeking to learn about their risk of having children with these disorders to have an appropriate reproductive plan. Here, we analyzed the exome sequencing data of 1,642 unrelated Thai individuals to identify the pathogenic variant (PV) frequencies in genes recommended by ACMG. The ascertainment bias was controlled by excluding the carriers of the PV in the genes for the conditions that are attributed to their offspring disorders. In the 113 ACMG-recommended genes, 165 PV and likely PVs in 60 genes of 559 exomes (34%, 559/1642) were identified. The carrier rate was increased to 39% when glucose-6-phosphate dehydrogenase (G6PD) was added. The carrier rate was still as high as 14.7% when thalassemia and hemoglobinopathies were excluded. In addition to thalassemia, hemoglobinopathies, and G6PD deficiency, carrier frequencies of >1% were found for Gaucher disease, primary hyperoxaluria, Pendred syndrome, and Wilson disease. Nearly 2% of the couples were at risk of having offsprings with the tested autosomal recessive conditions. The expanded carrier screening focused on common autosomal recessive conditions in Thai seems to be benefit among the study samples.

### Introduction

Although each Mendelian disease is generally uncommon, with >5,000 known heritable disorders, they are accounted for ~20% of infant mortality, ~18% of pediatric hospitalizations, and substantial numbers in adults [1]. At least 2,000 autosomal recessive diseases [2] and 500 X-linked diseases [3] have been identified. The preconception carrier screening for these disorders and proper reproductive option counseling have been shown to be cost-effective measures that reduce the burden of Mendelian

diseases [4-7]. The studies of the carrier rate of these disorders are highly varied with the number of tested genes and the genes included in the studies among different populations with no consensus agreement [8-11]. Until recently, the American College of Medical Genetics and Genomics (ACMG) released a set of 113 genes, both autosomal recessive and X-linked conditions, proposed as a standard list of genes for carrier screening [12]. Because X-linked glucose-6-phosphate dehydrogenase (G6PD) deficiency is the most common genetic disorder in Thailand and Southeast Asia [13], we added G6PD into the list and determined the carrier frequency and variant distribution of 114 recessive genes in the Thai population using the exomes of 1,642 unrelated Thais.

## Results

Using 1,642 exome sequencing data of unaffected parents whose children had rare diseases, we first excluded pathogenic and likely pathogenic variants (PV and LPV, respectively) found in these exomes, which caused the diseases in their children. We excluded 13 PV and LPV in 18 exomes (1%, 18/1642) (S1 Table).

Of the 113 ACMG-recommended genes, 165 PV and LPV in 60 genes of 559 exomes were found (559/1642: 34%). For G6PD, 7 PV and LPV of 127 exomes (female) were identified. In the analysis of a total of 114 genes, 172 PV and LPV in 61 genes of 640 exomes (640/1642: 39%) were identified (Table 1). Although we excluded HBA1, HBA2, and HBB genes causing thalassemia and hemoglobinopathies, the carrier rate of at least one screened disorder was still as high as 14.7% (241/1642) in our cohort. Of the samples, 7.4% (121/1642) carried more than one PV and LPV (Fig 1 and S2 Table).

### Gene carrier rate

In our cohort, all top three gene carrier rates were in hemolytic disorders with positive selection from malaria. The gene with the highest carrier rate is HBB, with 50.2% (321/640) and 19.6% (291/1642) of all carriers and participants, respectively. The second most common carrier gene is G6PD, with 19.8% (127/640) and 7.7% (127/1642) of all carriers and participants, respectively. HBA2 is the third most common gene, with 10.2% (65/640) and 4% (65/1642) of all carriers and participants, respectively.

There were 17 autosomal recessive genes with carrier frequencies of at least 1/250, an expected prevalence at birth of at least 1 in 250,000. As shown in Fig 2, three, five, and nine of these genes are related to hemolytic disorders (G6PD, HBB, and HBA2), inborn errors of metabolism (AGXT, ATP7B, CYP27A1, GBA, and PAH), and in the miscellaneous group (CEP290, CFTR, CLCN1, GJB2, MCPH1, NEB, PKHD1, SLC26A4, and USH2A), respectively.

### Variant carrier rate

The most common PV is HBB; c.79G>A (p.Glu27Lys) responsible for hemoglobin (Hb) E (HbE), representing 45.5% (291/640) and 17.7% (291/1642) of all carriers and participants, respectively. HBA2; c.427T>C (p. Ter143Gln\*31, also known as Hb Constant Spring) is the second most common variant,

representing 9.2% (59/640) and 3.6% (59/1642) of all carriers and participants, respectively. The third most common variant is G6PD; c.961G>A (p.Val321Met) causing G6PD deficiency, with 8.6% (55/640) and 3.4% (55/1642) of all carriers and participants, respectively. The variant carrier rates are shown in S3 Table.

#### Thalassemia and hemoglobinopathies

Although, the most common type of  $\alpha$ -thalassemia mutation is the deletion of one or more of the  $\alpha$ -globin genes, HBA1 and HBA2 [14], we did not attempt to analyze this type of genetic variant due to the inherent limitation of exome sequencing. The carrier rates of nondeletional  $\alpha$ -thalassemia were detected in 5.1% (71/1642; 6 in HBA1 and 65 in HBA2) of our cohort. The most common variant is c.427T>C 3.6% (59/1642) in Hb Constant Spring, followed by Hb Paksé c.429A>T 0.3% (5/1642). The frequency rate of  $\beta$ -thalassemia carriers is 19.6% (321/1642) with c.79G>A (p.Glu27Lys [HbE]) as the most common variant (17.7%, 291/1642), and HBB c.52A>T (0.79%, 13/1642) c.126\_129del (0.67%, 11/1642) were the next most common variant.

#### G6PD deficiency

The carrier rate of G6PD variants was 7.7% (127/1642). The prevalence of heterozygous/carrier G6PD variant and homozygous females in our cohort is 15.28% (127/831) and 0.8% (7/831), respectively (S5 Table). The percentage of symptomatic or biochemical G6PD deficiency in these samples was not known. The prevalence of hemizygous G6PD variant males in our cohort is 8% (65/811). The female carrier rate corresponds with the predicted prevalence calculated with male hemizygous prevalence using the Hardy–Weinberg formula. c.961G>A (p.Val321Met) (G6PD Viantien) and c.577G>A (p.Gly193Ser) (G6PD Mahidol) are the two most common G6PD variants found in our cohort.

#### Other autosomal recessive disorder carrier rate and common variants

The carrier frequency of PV and LPV in GBA causing Gaucher disease is 1.6% (27/1642) with c.605G>A (p.Arg202Gln) (0.54%, 9/1642) and c.1448T>C (p.Leu483Pro) (0.42%, 7/1642) being the two most common variants.

The carrier frequency of PV and LPV in genes associated with hereditary syndromic hearing loss SLC26A2 is 1.1% (18/1642). c.919-2A>G was the most common variant (0.3%, 5/1642). The second most common gene causing syndromic hearing loss carriers was USH2A (Usher syndrome) (0.97%, 16/1642). The carrier frequency of PV and PLV in GJB2, one of the most common genes causing nonsyndromic deafness is 0.6% (9/1642). All cases identified carried c.235delC (p.Leu79CysfsTer3) in GJB2.

Besides Gaucher disease, the most common inherited metabolic disease gene carriers in our cohort are hyperoxaluria (AGXT; 1.1%, 18/1642), Wilson disease (ATP7B; 1%, 17/1642), and phenylketonuria (PAH; 0.79%, 13/1642).

## Discussion

Next Generation Sequencing (NGS) has been increasingly used for expanded genetic carrier screening for multiple genes in a highly efficient manner in clinical laboratories [15]. The carrier rate ranges from as high as 43.6% in Ashkenazi Jewish individuals to 8.5% in East Asians [16], although the number of genes tested and which genes are included are varied among these studies. In our study using 1642 exomes of a Thai population, 39% of Thai individuals carried a PV and an LPV in at least one of 114 genes (113 genes have been recommended by ACMG) [12] plus G6PD, causing G6PD deficiency, which is most common in the Thai population. In this study, 23.8% are carriers of thalassemia and hemoglobinopathies (HBB, HBA1, and HBA2), followed by 7.7% of being a carrier for G6PD deficiency respectively. The carrier rate in our cohort is considered high, similar to the Ashkenazi Jewish group. The most prevalent variants are found in Hb genes and G6PD as aforementioned, are likely the result from the heterozygous advantage in malaria-endemic areas [17] rather than in-group mating as in Ashkenazi Jewish. After excluding thalassemia and hemoglobinopathies, the frequency of the carrier rate in our cohort was still high at 14.7%.

We observed 14 couples (1.7%) who were carriers of PV and LPV in the same gene (S4 Table). Of these couples, seven (0.9%, 7/821) were nonthalassemia carrier couples at risk of having children affected with recessive disorders. Thalassemia carrier screening is already included and available in the screening program of Thai pregnant women; therefore, it is estimated that nearly 1% of the couples could benefit from expanded carrier screening.

The prevalence of common recessive disorders can be estimated in our cohort. In general, this was comparable with the previous estimation of the prevalence rate of common autosomal recessive disorders (thalassemia and G6PD deficiency) [18-24]. This can be implied that the ascertainment and analysis by exclusion of the samples that carry the genetic variants in the disease responsible in their offspring strategy exclude the bias and the frequency rare therefore was comparable to the nondisease or population cohort. This strategy was proven in our previous analysis on the prevalence of the secondary findings. (The majority of the reportable diseases are inherited in an autosomal dominant manner.)

The most common carriers of inherited metabolic disease from our cohort are Gaucher disease caused by mutation in GBA. The result interpretation for this gene is challenging because GBA has a pseudogene. The homology of the GBA pseudogene is highest between exons 8 and 11, which occurs from recombination events [25]. Many clinical exome sequencing laboratories do not analyze or report the GBA variants. The variants commonly identified in our cohort are outside these highly homologous areas and are therefore less likely to be affected in our analysis. Conversely, the prevalence of the GBA variant carrier is likely to be underestimated in our cohort or a similar cohort using exome sequencing.

In our study, carriers of primary hyperoxaluria and Wilson disease PV and LPV were found to be the next most common. Founder variants in East Asian populations (c.2T>C (p.Met1Thr) in AGXT [26-29] and c.2333G>T (p.Arg778Leu) in ATP7B [30-35] were also identified in our cohort. Our study also supported the current newborn screening program in Thailand for Phenylketonuria (PKU) as the estimation of the carrier rates of 1 in 127 individuals comparable with other countries [36-38]; 8–10 children will be born with PKU every year (annual live birth of 600,000).

Hearing loss is one of the major disability problems in childhood, in which hereditary causes both syndromic and nonsyndromic are the main etiologies for these children. We identified that carriers in Pendred syndrome, Usher syndrome, and GJB2 (nonsyndromic hereditary neurosensory hearing loss) were common in Thai cohort similar to previous estimation in other Thai cohorts [39-50].

The carrier of cystic fibrosis in Caucasians are common [51-53], but we identified that only 0.97% of our sample carried PV and LPV in CFTR. As a result of low carrier frequency among Thais, Thais have lower incidence of cystic fibrosis than Caucasians (1/2,000 in Northern European descends vs. 1/400,000). Interestingly, we observed three females (0.18%) bearing 5T (c.1210–12 [5]) of CFTR, which can be found in mild cystic fibrosis but can also result in male infertility due to congenital bilateral absence of the vas deferens [54].

### Limitations

Our study had some limitations. Whole exome sequencing (WES) cannot reliably detect structural variants that are common in  $\alpha$ -thalassemia. Spinal muscular atrophy, a common muscular disorder, most commonly caused by a copy number change in the SMN1 gene and modified by the copy numbers of SMN2, is also missed by WES. The genome analysis with structural variant detection will help estimate these common disorders. Additionally, the frequency of variants based on NGS analysis may be underestimated without validation using Sanger sequencing, particularly for complex recombinant alleles of GBA pseudogenes. The other common PVs that will be missed by WES are trinucleotide repeat expansion. PV in FMR1 causing fragile X syndrome, the most common single-gene disorder cause of intellectual disability especially in males, cannot be detected with exome sequencing data. Our sample size has >80% power to detect alleles with frequency of more than 1%. The detection rate for the uncommon variant will require a larger sample size [55].

### Summary

Of the 114 genes, including the 113 ACMG-recommended genes and G6PD, 39% (640/1642) of Thai individuals were found to be carriers. Our study unveiled the landscape of PVs common in a Thai population based on computational filtering and manual curation which will be fundamental for designing a gene list for carrier screening at an individual level or future public health plan.

### Materials and methods

#### Study samples

Our study sample included the 1642 exomes of unrelated healthy Thai individuals (811 men and 831 women). Written informed consent was obtained from each individual. This study was approved by the Institutional Review Board of the Faculty of Medicine, Chulalongkorn University (IRB No. 264/62). The data are the part, but not all that were deposited, analyzed, curated in the Thai Reference Exome database [56]. The parents of the patient cohort were enrolled as study samples; therefore we excluded the variants in the genes responsible for patient diagnosis.

## Sequencing and bioinformatics analysis

Genomic DNA was extracted from peripheral blood leukocytes. Exomes were captured using either a TruSeq® Exome Kit (Illumina, San Diego, CA) on the NextSeq 500 system or SureSelect Human All Exons (Agilent Inc., Santa Clara, CA) on the HiSeq 4000 and Novaseq 6000 Systems. The reagents used are according to the manufacturer's standard protocol. Sequences were aligned to the human reference genome (GRCh37) using the Burrows–Wheeler Aligner package v 0.7.15 [57]. Variant calling was performed using the Genome Analysis Tool Kit (GATK Best Practice V3.7; Broad Institute, CA, USA), which is called by HaplotypeCaller [58]. ANNOVAR was used to annotate genetic variants. The variant interpretation was limited to the 113 ACMG-recommended genes and G6PD (12). We included variants listed as PV and LPV according to VarSome's ACMG classification [59] with a final manual curation by the authors.

## Statistical analysis

The frequency and 95% confidence interval were calculated using the proportion of carriers in total individuals tested (n/1642) and the Wilson method, respectively.

## Acknowledgments

This study was supported by the Health Systems Research Institute (65-040).

## References

1. Kingsmore S. Comprehensive carrier screening and molecular diagnostic testing for recessive childhood diseases. *PLoS Curr.* 2012:e4f9877ab8ffa9.
2. Xiao Q, Lauschke VM. The prevalence, genetic complexity and population-specific founder effects of human autosomal recessive disorders. *NPJ Genom Med.* 2021;6(1):41.
3. Migeon BR. X-linked diseases: susceptible females. *Genet Med.* 2020;22(7):1156-74.
4. Shotelersuk V, Limwongse C, Mahasirimongkol S. Genetics and genomics in Thailand: challenges and opportunities. *Mol Genet Genomic Med.* 2014;2(3):210-6.
5. Ministry of Public Health Thailand. *Thalassemia in Thailand: 2018* [Available from: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0246736#pone.0246736.ref006>]
6. Cowan RS. Moving up the slippery slope: mandated genetic screening on Cyprus. *Am J Med Genet C Semin Med Genet.* 2009;151C(1):95-103.
7. Bozkurt G. Results from the north cyprus thalassemia prevention program. *Hemoglobin.* 2007;31(2):257-64.

8. Singh K, Bijarnia-Mahay S, Ramprasad VL, Puri RD, Nair S, Sharda S, et al. NGS-based expanded carrier screening for genetic disorders in North Indian population reveals unexpected results - a pilot study. *BMC Med Genet.* 2020;21(1):216.
9. Chan OYM, Leung TY, Cao Y, Shi MM, Kwan AHW, Chung JPW, et al. Expanded carrier screening using next-generation sequencing of 123 Hong Kong Chinese families: a pilot study. *Hong Kong Med J.* 2021;27(1):177-83.
10. Hernandez-Nieto C, Alkon-Meadows T, Lee J, Cacchione T, Iyune-Cojab E, Garza-Galvan M, et al. Expanded carrier screening for preconception reproductive risk assessment: Prevalence of carrier status in a Mexican population. *Prenat Diagn.* 2020;40(5):635-43.
11. Chau JFT, Yu MHC, Chui MMC, Yeung CCW, Kwok AWC, Zhuang X, et al. Comprehensive analysis of recessive carrier status using exome and genome sequencing data in 1543 Southern Chinese. *NPJ Genom Med.* 2022;7(1):23.
12. Gregg AR, Aarabi M, Klugman S, Leach NT, Bashford MT, Goldwasser T, 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.
13. Sathupak S, Leecharoenkiat K, Kampuansai J. Prevalence and molecular characterization of glucose-6-phosphate dehydrogenase deficiency in the Lue ethnic group of northern Thailand. *Sci Rep.* 2021;11(1):2956.
14. Farashi S, Harteveld CL. Molecular basis of alpha-thalassemia. *Blood Cells Mol Dis.* 2018;70:43-53.
15. Rowe CA, Wright CF. Expanded universal carrier screening and its implementation within a publicly funded healthcare service. *J Community Genet.* 2020;11(1):21-38.
16. Lazarin GA, Haque IS, Nazareth S, Iori K, Patterson AS, Jacobson JL, 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.
17. Ha J, Martinson R, Iwamoto SK, Nishi A. Hemoglobin E, malaria and natural selection. *Evol Med Public Health.* 2019;2019(1):232-41.
18. Modell B, Darlison M. Global epidemiology of haemoglobin disorders and derived service indicators. *Bull World Health Organ.* 2008;86(6):480-7.
19. Farashi S, Harteveld CL. Molecular basis of alpha-thalassemia. *Blood Cells Mol Dis.* 2018;70:43-53.
20. Lai K, Huang G, Su L, He Y. The prevalence of thalassemia in mainland China: evidence from epidemiological surveys. *Sci Rep.* 2017;7(1):920.
21. Colah R, Italia K, Gorakshakar A. Burden of thalassemia in India: The road map for control. *Pediatric Hematology Oncology Journal.* 2017;2(4):79-84.
22. Luzzatto L, Nannelli C, Notaro R. Glucose-6-Phosphate Dehydrogenase Deficiency. *Hematol Oncol Clin North Am.* 2016;30(2):373-93.

23. Boonyuen U, Songdej D, Tanyaratsrisakul S, Phuanukoonnon S, Chamchoy K, Praoparotai A, et al. Glucose-6-phosphate dehydrogenase mutations in malaria endemic area of Thailand by multiplexed high-resolution melting curve analysis. *Malar J.* 2021;20(1):194.
24. Phompradit P, Kuesap J, Chaijaroenkul W, Rueangweerayut R, Hongkaew Y, Yamnuan R, et al. Prevalence and distribution of glucose-6-phosphate dehydrogenase (G6PD) variants in Thai and Burmese populations in malaria endemic areas of Thailand. *Malar J.* 2011;10:368.
25. Leija-Salazar M, Sedlazeck FJ, Toffoli M, Mullin S, Mokretar K, Athanasopoulou M, et al. Evaluation of the detection of GBA missense mutations and other variants using the Oxford Nanopore MinION. *Mol Genet Genomic Med.* 2019;7(3):e564.
26. Hulton S-A. Lumasiran: expanding the treatment options for patients with primary hyperoxaluria type 1. *Expert Opinion on Orphan Drugs.* 2021;9(7-10):189-98.
27. Li GM, Xu H, Shen Q, Gong YN, Fang XY, Sun L, et al. Mutational analysis of AGXT in two Chinese families with primary hyperoxaluria type 1. *BMC Nephrol.* 2014;15:92.
28. Yuen YP, Lai CK, Tong GM, Wong PN, Wong FK, Mak SK, et al. Novel mutations of the AGXT gene causing primary hyperoxaluria type 1. *J Nephrol.* 2004;17(3):436-40.
29. Tammachote R, Kingsuwannapong N, Tongkobpetch S, Srichomthong C, Yeetong P, Kingwatanakul P, et al. Primary hyperoxaluria type 1 and brachydactyly mental retardation syndrome caused by a novel mutation in AGXT and a terminal deletion of chromosome 2. *Am J Med Genet A.* 2012;158A(9):2124-30.
30. Kumar M, Gaharwar U, Paul S, Poojary M, Pandhare K, Scaria V, et al. WilsonGen a comprehensive clinically annotated genomic variant resource for Wilson's Disease. *Sci Rep.* 2020;10(1):9037.
31. Jang JH, Lee T, Bang S, Kim YE, Cho EH. Carrier frequency of Wilson's disease in the Korean population: a DNA-based approach. *J Hum Genet.* 2017;62(9):815-8.
32. Qiao L, Ge J, Li C, Liu Y, Hu C, Hu S, et al. Pathogenic gene variation spectrum and carrier screening for Wilson's disease in Qingdao area. *Mol Genet Genomic Med.* 2021;9(8):e1741.
33. Hua R, Hua F, Jiao Y, Pan Y, Yang X, Peng S, et al. Mutational analysis of ATP7B in Chinese Wilson disease patients. *Am J Transl Res.* 2016;8(6):2851-61.
34. Panichareon B, Taweechue K, Thongnoppakhun W, Aksornworanart M, Pithukpakorn M, Yenchitsomanus PT, et al. Six novel ATP7B mutations in Thai patients with Wilson disease. *Eur J Med Genet.* 2011;54(2):103-7.
35. Treepongkaruna S, Pienvichit P, Phuapradit P, Kodcharin P, Wattanasirichaigoon D. Mutations of ATP7B gene in two Thai siblings with Wilson disease. *Asian Biomedicine.* 2010;4(1):163-9.
36. Hillert A, Anikster Y, Belanger-Quintana A, Burlina A, Burton BK, Carducci C, et al. The Genetic Landscape and Epidemiology of Phenylketonuria. *Am J Hum Genet.* 2020;107(2):234-50.



37. Wang T, Okano Y, Eisensmith RC, Lo WH, Huang SZ, Zeng YT, et al. Missense mutations prevalent in Orientals with phenylketonuria: molecular characterization and clinical implications. *Genomics*. 1991;10(2):449-56.
38. Chaiyasap P, Ittiwut C, Srichomthong C, Sangsin A, Suphapeetiporn K, Shotelersuk V. Massive parallel sequencing as a new diagnostic approach for phenylketonuria and tetrahydrobiopterin-deficiency in Thailand. *BMC Med Genet*. 2017;18(1):102.
39. Petit C, Levilliers J, Hardelin JP. Molecular genetics of hearing loss. *Annu Rev Genet*. 2001;35:589-646.
40. Hu H, Zhou P, Wu J, Lei W, Wang Y, Yang Y, et al. Genetic testing involving 100 common mutations for antenatal diagnosis of hereditary hearing loss in Chongqing, China. *Medicine (Baltimore)*. 2021;100(17):e25647.
41. Klumsathian S, Lorlipiwong W, Iemwimangsa N, Sensorn I, Panthan B, Chareonsirisuthigul T, et al. AB034. Carrier frequency of inherited genetic disorders in Thai population: implication for designing expanded carrier screening panel. *Annals of Translational Medicine*. 2017;5(S2):AB034-AB.
42. Snabboon T, Plengpanich W, Saengpanich S, Sirisalipoch S, Keelawat S, Sunthornyothin S, et al. Two common and three novel PDS mutations in Thai patients with Pendred syndrome. *J Endocrinol Invest*. 2007;30(11):907-13.
43. Wattanasirichaigoon D, Limwongse C, Jariengprasert C, Yenchitsomanus PT, Tocharoenthanaphol C, Thongnoppakhun W, et al. High prevalence of V37I genetic variant in the connexin-26 (GJB2) gene among non-syndromic hearing-impaired and control Thai individuals. *Clin Genet*. 2004;66(5):452-60.
44. Choy KW, Cao Y, Lam ST, Lo FM, Morton CC, Leung TY. Target-enriched massively parallel sequencing for genetic diagnosis of hereditary hearing loss in patients with normal array CGH result. *Hong Kong Med J*. 2018;24 Suppl 3(3):11-4.
45. Taniguchi M, Matsuo H, Shimizu S, Nakayama A, Suzuki K, Hamajima N, et al. Carrier frequency of the GJB2 mutations that cause hereditary hearing loss in the Japanese population. *J Hum Genet*. 2015;60(10):613-7.
46. Han SH, Park HJ, Kang EJ, Ryu JS, Lee A, Yang YH, et al. Carrier frequency of GJB2 (connexin-26) mutations causing inherited deafness in the Korean population. *J Hum Genet*. 2008;53(11-12):1022-8.
47. Kim SY, Park G, Han KH, Kim A, Koo JW, Chang SO, et al. Prevalence of p.V37I variant of GJB2 in mild or moderate hearing loss in a pediatric population and the interpretation of its pathogenicity. *PLoS One*. 2013;8(4):e61592.
48. Pollak A, Skorka A, Mueller-Malesinska M, Kostrzewa G, Kisiel B, Waligora J, et al. M34T and V37I mutations in GJB2 associated hearing impairment: evidence for pathogenicity and reduced penetrance. *Am J Med Genet A*. 2007;143A(21):2534-43.
49. Lentz J, Keats B. Usher syndrome type II. *GeneReviews*<sup>®</sup>[Internet]. 2016.

50. Xi Y, Chen G, Lei C, Wu J, Zhang S, Xiao M, et al. Expanded carrier screening in Chinese patients seeking the help of assisted reproductive technology. *Mol Genet Genomic Med*. 2020;8(9):e1340.
51. Antoniou S, Elston C. Cystic fibrosis. *Medicine*. 2016;44(5):321-5.
52. Scotet V, L'Hostis C, Ferec C. The Changing Epidemiology of Cystic Fibrosis: Incidence, Survival and Impact of the CFTR Gene Discovery. *Genes (Basel)*. 2020;11(6).
53. Grody WW, Desnick RJ. Cystic fibrosis population carrier screening: here at last--are we ready? *Genet Med*. 2001;3(2):87-90.
54. Sun W, Anderson B, Redman J, Milunsky A, Buller A, McGinniss MJ, et al. CFTR 5T variant has a low penetrance in females that is partially attributable to its haplotype. *Genet Med*. 2006;8(6):339-45.
55. Fujikura K (2016) Global Carrier Rates of Rare Inherited Disorders Using Population Exome Sequences. *PLoS ONE* 11(5): e0155552. <https://doi.org/10.1371/journal.pone.0155552>
56. Shotelersuk, V, Wichadakul, D, Ngamphiw, C, et al. The Thai reference exome (T-REx) variant database. *Clinical Genetics*. 2021; 100( 6): 703- 712. doi:10.1111/cge.14060
57. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-60.
58. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-303.
59. Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Albarca Aguilera M, Meyer R, et al. VarSome: the human genomic variant search engine. *Bioinformatics*. 2019;35(11):1978-80.

#### Supporting information

S1 Table. Genetic variants of the 18 parents who were responsible for the presenting symptoms in their children

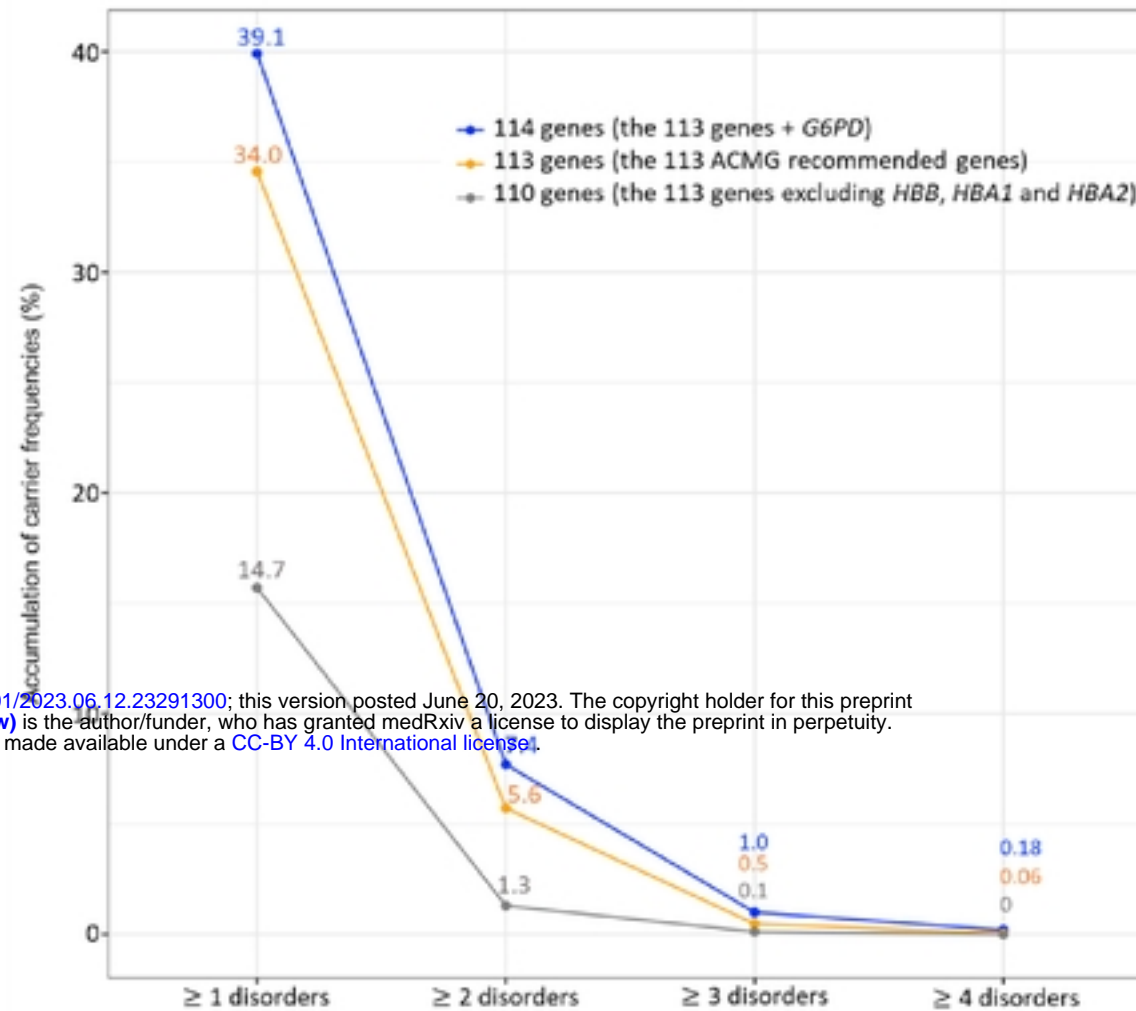
S2 Table. Carrier frequencies of one or more genetic disorders in 114 genes of 1,642 unrelated Thais

S3 Table. The variant carrier rate (VCR) of each variant

S4 Table. Pathogenic or likely pathogenic variants in the same genes that were harbored by the couples

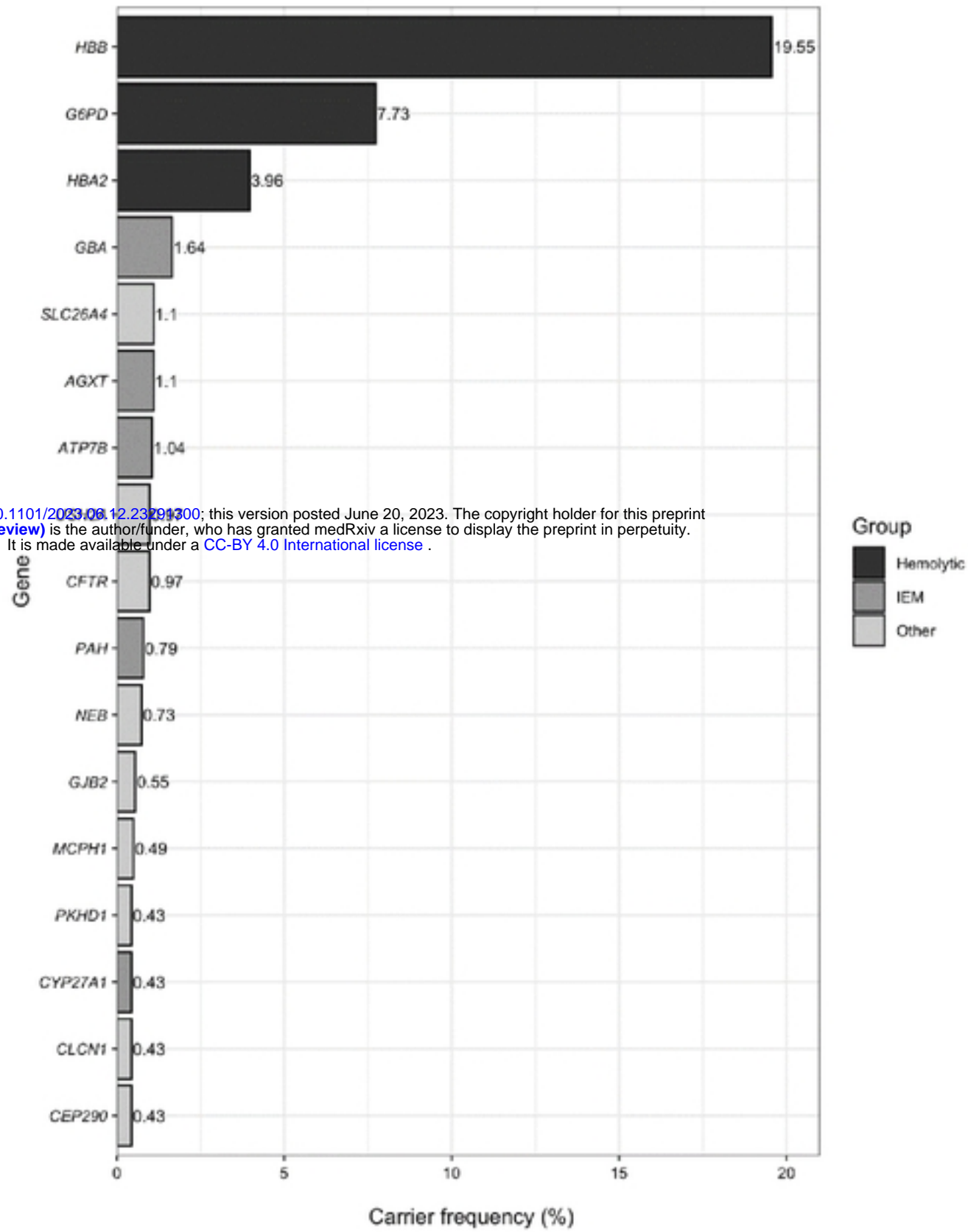
S5 Table. Characteristics of glucose-6-phosphate dehydrogenase (G6PD) phenotypes in 1,642 unrelated Thais

medRxiv preprint doi: <https://doi.org/10.1101/2023.06.12.23291300>; this version posted June 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).



**Fig 1.** Carrier frequencies of one or more genetic disorders in 114 genes of 1,642 unrelated Thais. Orange color represents carrier frequencies in 113 ACMG recommended genes. Blue color represents those in 114 genes, which are the 113 genes and G6PD. Gray color represents those in 110 genes, which are the 113 genes excluding three genes underlying thalassemia and hemoglobinopathies (*HBB*, *HBA1*, and *HBA2*).

medRxiv preprint doi: <https://doi.org/10.1101/2023.06.12.23291300>; this version posted June 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).



**Fig 2.** The 17 genes underlying autosomal recessive disorders with carrier frequencies of  $\geq 1/250$  (0.4%) in the 1,642 unrelated Thai individuals.