Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Investigation of the Association Between G6PD, UGT1A1, and SLCO1B1 Gene Polymorphisms and Neonatal Hyperbilirubinemia Using MALDI-TOF MS Technology

Wen-Fang Wang
Haikou Maternal and Child Health Hospital
Li-Sen Shi
Haikou Maternal and Child Health Hospital
Zeng-Wen Xing
Haikou Maternal and Child Health Hospital
Yan-Mei Han
Haikou Maternal and Child Health Hospital
Wei-Qiang Liu
liuwq06@126.com

Longgang District Maternity & Child Healthcare Hospital of Shenzhen City

Research Article

Keywords: Glucose-6-phosphate dehydrogenase (G6PD) activity, Uric acid-glucuronide transferase 1A1 (UGT1A1), Solute carrier organic anion transporter family member 1B1 (SLCO1B1), Neonatal hyperbilirubinemia (NHB), Single nucleotide polymorphism

Posted Date: June 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9967522/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Investigation of the Association Between *G6PD*, *UGT1A1,* and *SLCO1B1* Gene Polymorphisms and Neonatal Hyperbilirubinemia Using MALDI-TOF MS Technology

Wen-Fang Wang1,2,#, Li-Sen Shi2,#, Zeng-Wen Xing2, Yan-Mei Han2, Wei-Qiang Liu3,*

1Clinical Laboratory Diagnostics, Shantou University Medical College, Shantou, Guangdong Province, China

2Medical Genetics and Prenatal Diagnosis, Haikou Maternal and Child Health Hospital, Haikou, Hainan Province, China

3Central Laboratory (Genetics Lab), Longgang District Maternity & Child Healthcare Hospital of Shenzhen City (Affiliated Shenzhen Women and Children's Hospital(Longgang) of Shantou University Medical College) Shenzhen City， China

#These authors have contributed equally to this work. *Correspondence: liuwq06@126.com

Objective: To investigate the relationship between *neonatal G6PD deficiency and G6PD, UGT*1A*1,* and *SLCO1B1* gene polymorphisms and phenotypes, as well as their association with neonatal hyperbilirubinemia (NHB), using MALDI-TOF MS technology.

Methods: A total of 210 cases of neonatal hyperbilirubinemia were selected as the study group, and 195 healthy subjects were recruited as the control group. G6PD enzyme activity was quantitatively measured in peripheral blood samples.The MassARRAY® MALDI-TOF system (Agena Bioscience®, California, USA) was used to analyze *pathogenic variants of G6PD* and single-nucleotide polymorphisms (SNPs) at *UGT1A1* and *SLCO1B1* gene loci in the samples. The relationships between these variants and quantified G6PD activity and bilirubin levels were evaluated.

Results: In the correlation analysis of 10 SNP loci with serum total bilirubin concentration, G6PD 95A>G (rs137852340),1388G>A (rs72554664) in G6PD, 211G>A (rs4148323) in UGT1A1, and -11187G>A (rs4149015) in *SLCO1B1* had a significant effect on serum total bilirubin levels.

Conclusion: G6PD 95A>G (rs137852340), 1388G>A (rs72554664),UGT1A1 211 G>A (rs4148323), *and SLCO1B1* -11187 G>A (rs4149015) are important risk factors for the

development of neonatal hyperbilirubinemia, are associated with the risk of its occurrence and progression, and serve as potential biomarkers for predicting this condition.

Keywords: Glucose-6-phosphate dehydrogenase (G6PD) activity; Uric acid-glucuronide transferase 1A1 (UGT1A1); Solute carrier organic anion transporter family member 1B1 (SLCO1B1); Neonatal hyperbilirubinemia (NHB); Single nucleotide polymorphism

Neonatal hyperbilirubinemia (NHB) is one of the most common clinical conditions in newborns; 60%–80% of newborns develop jaundice of varying severity after birth, requiring careful assessment and repeated management[1-2] .In many newborns, low concentrations of bilirubin are generally harmless and even offer certain antioxidant benefits; therefore, some cases of NHB are considered a benign, transient physiological phenomenon:However, in a small number of infants, serum total bilirubin (TSB) may rise to dangerous levels, posing a direct threat of acute bilirubin encephalopathy, which can lead to brain damage. Even if the infant survives, chronic or permanent damage and sequelae may occur in the central nervous system [3] .In addition to established risk factors for neonatal hyperbilirubinemia such as preeclampsia during pregnancy, gestational age, and preterm birth[4-5] , the occurrence of NHB is associated with genetic and environmental factors[6] .

In recent years, the importance of genetic factors has been increasingly recognized[7] . Multiple studies have shown that polymorphisms in three genes are associated with an increased risk of NHB: (1) the red blood cell enzyme glucose-6-phosphate dehydrogenase (*G6PD*) ;(2) the hepatic bilirubin-conjugating enzyme UGT1A1 (*UGT1A1*)[9]-11] ;(3) Solute-Linking Carrier Organic Ion Transporter Polypeptide 1B1 (*SLCO1B1*)[12] . Pathogenic variants in these genes lead to dysfunction of the corresponding enzymes or transporters, severely impairing bilirubin metabolism. These genetic variants may interact with one another or with environmental factors to cause NHB[13] .

Glucose-6-phosphate dehydrogenase (*G6PD*) is the enzyme that catalyzes the first step of the pentose phosphate pathway.It is used to produce reduced nicotinamide adenine dinucleotide phosphate, which is essential for protecting red blood cells from oxidative stress. *G6PD* deficiency is globally recognized as a risk factor for severe hyperbilirubinemia, and this deficiency is caused by pathogenic variants in the G6PD gene[14] .Pathogenic G6PD

variants are a significant risk factor for severe NHB and may even lead to kernicterus. *G6PD* gene variants may predispose individuals to NHB by triggering acute hemolytic episodes.Additionally, pathogenic G6PD variants can lead to mild hemolysis, and when combined with UGT1A1 gene polymorphisms, may result in severe NHB[8]. To date, the 1388 G>A (rs72554664) variant in the *G6PD* SNP is common in Asians but has been rarely reported in other ethnic groups.

Unconjugated bilirubin is conjugated with glucuronic acid in the liver by uridine diphosphate glucuronosyltransferase (UGT), encoded by the UGT1A1 gene, to form conjugated bilirubin, which is then excreted from the body through a series of complex processes.Pathogenic variants in the UGT1A1 gene cause structural or functional defects in the enzyme, which can reduce UGT activity and lead to elevated serum bilirubin levels[15] . Among more than 100 types of pathogenic variants in UGT1A1, the 211G>A (rs4148323) allele is the most common pathogenic variant in Asian newborn cohorts,A multicenter research group at Shanghai Children's Hospital found that among 240 children with severe hyperbilirubinemia, the c.211G>A mutation rate was 65.0% and the A allele frequency was 43.8%, significantly higher than the 19.0% observed in the control group;[16] .

The organic anion transport polypeptide encoded by the SLCO1B1 gene is an important hepatic transporter that, together with a series of specialized transporters located on the membranes of hepatocytes and intestinal cells, facilitates the transmembrane transport of bilirubin. Cellular studies and related clinical research have shown that pathogenic variants at polymorphic sites in the SLCO1B1 gene reduce the transport activity of the SLCO1B1 gene to varying degrees[17] .The 388A>G (rs2306283), 521T>C (rs4149056), and -11187 G>A (rs4149015) polymorphisms have been reported to affect the function of their corresponding protein, OATP-C [18] .Several studies have shown that SLCO1B1 polymorphisms can influence serum bilirubin levels in adults with hyperbilirubinemia or in healthy adults[19] . However, there are currently few reports on whether SLCO1B1 polymorphisms are associated with bilirubin levels in pathological conditions, particularly in the special population of newborns.The objective of this study is to determine G6PD activity and identify specific genotypes associated with it in the local neonatal population, as well as to assess the impact of *UGT1A1*

and *SLCO1B1* gene polymorphisms on the risk of hyperbilirubinemia, thereby providing a theoretical basis for disease prevention and treatment.

1. Materials and Methods

o 1.1. Study Population

From January 2024 to January 2025, 210 newborns with hyperbilirubinemia were recruited from the neonatal ward of Haikou Maternal and Child Health Hospital to form the case group, along with 195 healthy control subjects. All individuals included in this study were free of congenital biliary atresia, liver dysfunction, polycythemia, sepsis, and other related conditions.To rule out the possibility of Crigler-Najjar syndrome, all participants had serum total bilirubin levels below 425 μmol/L, and the proportion of unconjugated bilirubin to total bilirubin was ≥80%. Collected data included gender, gestational age, birth weight, feeding method, mode of delivery, total and unconjugated bilirubin levels, and hematocrit.Informed consent was obtained from all guardians of the participants. The study protocol strictly adhered to the ethical guidelines of the Declaration of Helsinki and was approved by the Ethics Committee of Haikou Maternal and Child Health Hospital (Ethics Approval No.: [2024]01045). Clinical trial number: not applicable.

1.2 Study Methods

1.2.1 Sample Collection On the day of admission, venous blood (2–3 mL) was collected from each newborn for complete blood count and serum biochemical analysis.A fully automated biochemical analyzer (Hitachi 7080, Japan) was used to quantitatively measure G6PD enzyme activity (rate method, 37°C, NADP+ as coenzyme, expressed in U/g Hb), as well as serum bilirubin (total and direct bilirubin) and alkaline phosphatase (ALP) levels.G6PD activity <2.0 U/g Hb was defined as G6PD deficiency. Additionally, the remaining anticoagulated peripheral blood after routine blood tests was stored at -80°C for DNA extraction and genotyping of SNP loci.

1.2.2 Diagnosis of Pathological Jaundice

Based on the “Guidelines for the Diagnosis and Treatment of Neonatal Hyperbilirubinemia” (2025) published in the *Chinese Journal of Pediatrics*. A peak serum total bilirubin level ≥221 μmol/L was classified as pathological jaundice, while a peak serum

total bilirubin level <221 μmol/L was classified as normal bilirubin levels.All newborns in the case group were full-term or near-term (gestational age ≥ 35 weeks) and developed jaundice within 72 hours of birth or had serum total bilirubin levels exceeding the 95th percentile curve for their corresponding gestational age and postnatal age.

1.2.3 Selection of SNP Locations

The *G6PD*, UGT1A1, and SLCO1B1 genes were selected due to their critical roles in bilirubin metabolism. The tag SNPs selected for validation adhered to the following criteria: (1) SNPs located within the gene coding region, promoter region, and the 2-kb regulatory region upstream and downstream;(2) SNPs were screened using the NCBI Human Genome SNP Database and the gnomAD database, with priority given to hotspots with known biological functions or those reported in the literature (see Table S1 in the Supplementary Materials for details). Based on these parameters, a total of 12 target SNPs were selected, including 5 from the G6PD gene, 4 from the UGT1A1 gene, and 3 from the SLCO1B1 gene.It is worth noting that this study did not include testing for the UGT1A1 gene promoter region (TA) 7-repeat variant (UGT1A1*28, rs3064744) for the following reasons: (1) Due to the sequence repeat characteristics of this locus, it is difficult to design stable extension primers on the MassARRAY platform used in this study;(2) Existing literature has sufficiently demonstrated that the frequency of this variant in Han Chinese newborns is extremely low (<1%); given the sample size of this study, the expected number of detected cases would be fewer than 5, resulting in limited statistical power; (3) This study focuses on evaluating the combined effects of functional SNPs in the coding region on G6PD enzyme activity. Detection of this variant will be supplemented in subsequent multicenter studies with expanded sample sizes using pyrosequencing or capillary electrophoresis.

1.2.4 DNA Extraction, Amplification, and Genotyping

Genomic DNA was isolated and extracted from peripheral blood samples using a magnetic bead method with a kit provided by Guangzhou Darui Biotechnology Co., Ltd., following the manufacturer’s instructions.The extracted DNA was diluted with 200 μL of nuclease-free water and stored at -80°C for later use. Genotyping was performed using

matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) with the MassARRAY system (Agena Bioscience®, California, USA).Twelve SNP loci were analyzed, specifically: G6PD gene 95A>G (rs137852340), 1388G>A (rs72554664), 1376G>T (rs72554665), 871G>A (rs137852327), 1024C>T(rs137852342); UGT1A1 gene 865-2441G>A (rs3771341), 1091C>T (rs34946978), 686C>A (rs35350960), and 211G>A (rs4148323);SLCO1B1 gene -11187G>A (rs4149015), 388A>G (rs2306283), and 521T>C (rs4149056). For the above SNPs, amplification primers and single-base extension primers for the target SNP sites were designed using Genotyping Tools and MassARRAY Assay Design software;PCR amplification and single-base extension were performed to obtain PCR products; after purifying the extension products using resin, the samples were spotted onto SpectroCHIP chips and analyzed using the MassARRAY® MALDI-TOF system (Agena Bioscience®, California, USA).Finally, the detection results were analyzed for genotyping using MassARRAY Typer 4.0 software (Agena Bioscience®, California, USA), and the final results were generated.

1.2.5 Primer Design and Quality Control

PCR amplification primers and single-base extension primers for each SNP locus were designed using MassARRAY Assay Design software. The specificity of all primers was verified via BLAST alignment. To validate the accuracy of genotyping results, each batch of testing included positive controls with known genotypes, negative controls (no-template controls, NTCs), and duplicate samples (5% of samples were randomly selected for duplicate testing).In addition, Sanger sequencing was used to validate 10% of the random samples, with 100% consistency in genotyping results. [20] . The principle involves irradiating a co-crystallized film formed by the sample and the matrix with a laser; the matrix absorbs the laser energy and transfers it to the biomolecules, transferring protons to the biomolecules during the ionization process, thereby ionizing them.Under the influence of an electric field, ions accelerate through the flight tube. Based on the proportional relationship between the mass-to-charge ratio (M/Z) of ions and their flight time, the time it takes for ions to reach the detector is measured, thereby distinguishing different ions and enabling qualitative identification. Quantitative analysis is then performed by detecting the mass spectrometry peak height, which is proportional to the molecular weight of each DNA molecule, thereby

enabling SNP genotyping, detection of insertions or deletions, and quantitative expression analysis [21].The detection process is as follows: First, the target DNA sequence is amplified using PCR technology; subsequently, shrimp alkaline phosphatase is used to remove excess deoxyribonucleotides from the products. Specific extension primers bind to the products and extend a modified deoxy-nucleoside triphosphate. Finally, the extension products are analyzed by time-of-flight mass spectrometry, and genotyping is performed based on differences in molecular weight.MALDI-TOF MS technology can detect low-frequency alleles or rare mutations as low as 0.1%, offering extremely high accuracy and sensitivity [22].

1.3 Statistical Analysis

All SNPs were evaluated for Hardy-Weinberg equilibrium using the chi-square test. Demographic variables between the patient and control groups were compared using the chi-square test and t-tests. Allele frequencies were determined by direct counting (i.e., counting the number of times each allele appeared in the sample); comparisons of genotype and allele frequencies between the case and control groups were assessed using the chi-square test.Logistic regression was used to calculate odds ratios (OR) and 95% confidence intervals (CI) to assess the association between three different genetic models (dominant, recessive, and additive) and neonatal hyperbilirubinemia.Correlation analysis between different genotypes and bilirubin levels was performed using the nonparametric Kruskal-Wallis rank-sum test; data are presented as mean ± standard deviation (SD) or median and interquartile range (IQR). We also performed logistic regression adjusted for sex, gestational age, birth weight, feeding method, mode of delivery, and hematocrit to explore the association between SNPs and the risk of hyperbilirubinemia.Significance between groups was assessed using two-sided tests, with a significance level of $P < 0.05$. All statistical graphs and data were generated and analyzed using R version 3.4.1. All P-values are two-sided; P-values less than 0.05 were considered statistically significant.

2. Results

2.1 Results of the Hardy-Weinberg Equilibrium Test

To ensure the representativeness of the population, Hardy-Weinberg (HWE) tests were

performed on 12 loci across three genes.Among these, the two loci rs72554665 and rs137852327 are both located on the G6PD gene and are on the X chromosome, exhibiting X-linked inheritance. Due to the influence of sex ratio, low mutation frequency, sample size limitations, and potential case enrichment, the genotype distribution deviated from Hardy-Weinberg equilibrium ($P<0.05$).The remaining 10 SNP loci include 95A>G (rs137852340), 1388G>A (rs72554664), and 1024C>T (rs137852342);UGT1A1 gene 865-2441G>A (rs3771341), 1091C>T (rs34946978), 686C>A (rs35350960), and 211G>A (rs4148323);The SLCO1B1 gene variants -11187G>A (rs4149015), 388A>G (rs2306283), and 521T>C (rs4149056) were all in Hardy-Weinberg equilibrium. Therefore, the 10 SNPs that met the Hardy-Weinberg equilibrium criteria were subjected to further analysis.See Table S1

2.2 Comparison of Clinical Characteristics Between the Case Group and the Control Group

The results of the comparison of baseline clinical data between the control group and the NHB group are shown in Table 1. There was a significant difference in gestational age between the control group and the NHB group, with the NHB group having a lower gestational age than the control group ($P < 0.05$).Compared with the control group, the hematocrit (HCT) level was significantly higher in the NHB group ($P < 0.001$), while there were no statistically significant differences in gender, birth weight, or alkaline phosphatase (ALP) levels ($P > 0.05$).

2.3 Differences in the Distribution of Genotypes and Allele Frequencies at Variant Sites of the G6PD, UGT1A1, and SLCO1B1 Genes Between the Two Groups of Patients

Genotyping was performed for a total of 10 SNP loci across the G6PD, UGT1A1, and SLCO1B1 genes (see Table 2 and Figure 1). Analysis of genotype and allele frequencies revealed that the genotype and allele frequencies of G6PD 95A>G (rs137852340) andand 1388G>A (rs72554664) showed significant differences in genotype and allele frequencies between the control and NHB groups, with these differences being statistically significant. The detection rate of the T allele at 1024C>T (rs137852342) was higher in the case group (2.38%) than in the control group (0.26%),and this difference was statistically significant ($P = 0.012$). The frequencies of the GG, AG, and AA genotypes for UGT1A1 211 G>A

(rs4148323) were 128 (61.0%), 71 (33.8%), and 11 (5.2%), respectively;In this study, only the heterozygous CT genotype was detected for the 1091C>T (rs34946978) genotype. The genotypes (TT, TC, and CC) at the SLCO1B1 521T>C (rs4149056) locus and the genotypes (GG, GA, and AA) at the SLCO1B1-11187G>A (rs4149015) genotypes (GG, GA, and AA), as well as the allele A, were detected at a higher rate in the case group (21.43%) than in the control group (11.79%), with a statistically significant difference ($P < 0.001$) (Table 2).

2.4 Correlation Analysis Between G6PD, UGT1A1, and SLCO1B1 Genotypes and Bilirubin Levels

Analysis of the mean bilirubin levels measured across the three genotypes (G6PD, UGT1A1, and SLCO1B1) revealed several noteworthy characteristics (Table 3):Since G6PD deficiency is an X-linked incompletely dominant trait, when male heterozygotes (G) and female homozygotes (GG) were combined into the mutant group, their median serum total bilirubin reached 276.5 μmol/L, which was significantly higher than the 221.5 μmol/L observed in the wild-type group comprising male heterozygotes (A) and female homozygotes (AA) ($P = 0.003$);Similarly, the median bilirubin level in the UGT1A1 211G>A (rs4148323) homozygous mutant (AA) group was 262.3 μmol/L, which was also significantly higher than the 194.8 μmol/L observed in the wild-type (GG) group ($P < 0.001$).This "dose-response" pattern of increasing levels is corroborated by the logistic regression results in Table 5, which were adjusted for covariates such as gestational age, HCT, and sex—the additive models for the aforementioned loci remained independently associated with NHB risk, suggesting that their pathogenic effects do not depend on traditional risk factors such as gestational age or hematocrit.It is worth noting that although SLCO1B1 -11187G>A (rs4149015) was significant under the dominant model, the bilirubin level in the homozygous mutant (AA) group (234.3 μmol/L) was only slightly higher than that in the heterozygous group (229.2 μmol/L), suggesting that this locus may primarily exert its effect through a dominant negative mechanism.

2.5 Logistic regression analysis of the association between genotypes of 10 SNPs and neonatal hyperbilirubinemia using dominant, recessive, and additive genetic models

The study showed that G6PD 95A>G (rs137852340) and 1388G>A (rs72554664) increased the risk of elevated neonatal peripheral blood bilirubin levels under both dominant

and additive inheritance models (P<0.05);Under the dominant inheritance model, the UGT1A1 211G>A (rs4148323) AA and GA genotypes were associated with an increased risk of hyperbilirubinemia (OR 2.406, p < 0.001), and they also significantly influenced neonatal peripheral blood bilirubin levels under the additive inheritance model (P < 0.001);SLCO1B1 -11187G>A (rs4149015) and 521T>C (rs4149056) significantly influenced neonatal peripheral blood bilirubin levels under both dominant and additive inheritance models (P < 0.05). Additionally, to ensure data stability, this study did not analyze group models where the number of individuals with any given variant was less than 10. See Table 4.

2.6 Association Analysis of Alleles and Genotypes of G6PD, UGT1A1, and SLCO1B1 Variants with Neonatal Hyperbilirubinemia (Logistic Regression Model)

Logistic regression analysis of the association between neonatal hyperbilirubinemia (NHB) and genotype frequencies, as shown in Table 5, identified several significant genetic markers: G6PD (95A>G, 1388G>A, and 1024C>T), UGT1A1 (211G>A and 1091C>T), and SLCO1B1 -11187G>A(rs4149015). These loci provide key insights into the genetic basis of NHB. Specifically, the G6PD 95A>G locus demonstrated robust significance: the unadjusted odds ratio (OR) for the AG genotype was 10.89 (P = 0.02), and it remained significant after adjustment (OR = 9.67, P = 0.04).Similarly, the A allele showed strong association (uncorrected OR = 11.45, P = 0.02; corrected OR = 10.37, P = 0.02), indicating that it is a major genetic risk factor.Further supporting the involvement of G6PD, the 1388G>A locus was also significant: the unadjusted OR for the GA genotype was 2.73 (P = 0.006), which decreased slightly after adjustment but remained significant (OR = 2.32, P = 0.04).The A allele showed a similar trend (unadjusted OR = 2.23, P = 0.007; adjusted OR = 2.22, P = 0.01). The UGT1A1 211G>A polymorphism showed a strong association.The unadjusted OR for the GA genotype was 2.37 (P < 0.001), and this association persisted in the adjusted analysis (OR = 2.17, P = 0.002). The A allele also demonstrated robust significance (unadjusted OR = 2.05, P < 0.001; adjusted OR = 1.79, P = 0.005).Overall, these findings confirm that specific genetic variants in G6PD, UGT1A1, and SLCO1B1 are significantly associated with the risk of NHB. This evidence lays the foundation for further research into the roles of these genes in bilirubin metabolism and may contribute to the development of genetic screening and prevention strategies. Future studies should explore gene-environment interactions and

variations in prevalence across different populations.

3. Discussion

The development of neonatal hyperbilirubinemia is influenced by a variety of complex and interrelated factors. Although the exact etiology remains unclear, there has been growing attention on the role of genetic factors in jaundice[23] . Studies have confirmed that genetic variations are the underlying cause of clinical pathogenic factors[24] , with genetic polymorphisms in bilirubin metabolism genes likely playing a significant role in the development of hyperbilirubinemia[25] . Our study indicates a significant association between susceptibility to neonatal hyperbilirubinemia and polymorphisms in the G6PD, UGT1A1, and SLCO1B1 genes.

This study suggests that the G6PD 95A>G (rs137852340) and 1388 G>A (rs72554664) SNP loci are associated with the risk of hyperbilirubinemia; individuals carrying the pathogenic variant allele exhibit significantly elevated serum bilirubin levels and a significantly increased risk of hyperbilirubinemia.A previous study consistent with our findings reported that the pathogenic variant 1388 G>A (rs72554664) in the G6PD gene affects enzyme function, leading to reduced enzyme activity or a decrease in enzyme levels within cells[26] .The authors of another study screened for pathogenic variants associated with neonatal G6PD deficiency and identified six pathogenic variants in the G6PD gene, including the 1388 G>A (rs72554664) variant, which accounted for 95.5% of G6PD deficiency cases,suggesting that this variant can *affect* G6PD enzyme activity and increase the risk of hyperbilirubinemia[27] .The 1388 G>A (rs72554664) SNP is located in exon 12 and has been found to cause the substitution of arginine for histidine, which may impair G6PD enzyme activity, thereby predisposing individuals to hyperbilirubinemia[14] .Furthermore, whether different G6PD genotypes correlate with the degree of elevated peak bilirubin levels requires further investigation in future studies with a larger sample size.

Variations in genes related to the bilirubin metabolic pathway may affect bilirubin

binding and excretion, thereby being closely associated with neonatal hyperbilirubinemia. Studies on single-gene polymorphisms have been reported both domestically and internationally[28-29] .Studies *have shown* that genetic variants in the promoter and coding regions of *UGT1A1 and SLCO1B1* may lead to reduced enzyme levels or loss of function, thereby increasing the risk of neonatal hyperbilirubinemia. In particular, the 211G>A (rs4148323) variant in the *UGT1A1* coding region is a risk factor for neonatal hyperbilirubinemia[30,31]. In this study, the UGT1A1 211G>A (rs4148323) variant significantly influenced peak neonatal peripheral blood bilirubin levels under both dominant and additive genetic models, consistent with findings from other studies, further confirming this locus as a risk locus for neonatal hyperbilirubinemia[18] .However, in this study, 865-2441G>A (rs3771341) and 1091C>T (rs34946978) showed no significant association with neonatal hyperbilirubinemia in logistic regression analyses adjusted for relevant parameters or under any of the three genetic models. Whether these two loci are risk loci for neonatal hyperbilirubinemia remains to be further confirmed.

OATP2 is a class of transport proteins highly expressed on the basolateral membrane of hepatocytes and encoded by the SLCO1B1 gene. It mediates the uptake of various organic anions—including conjugated and unconjugated bilirubin and unconjugated bile salts—from the blood into the liver[32] .Pathogenic variants in the SLCO1B1 gene can impair OATP2 transport activity, affecting bilirubin uptake and leading to significantly elevated plasma bilirubin concentrations, thereby increasing the risk of neonatal hyperbilirubinemia. Additionally, in the general population, these variants are associated with elevated serum bilirubin levels[33-34] .In this study, the SLCO1B1 -11187G>A (rs4149015) and c.521T>C (rs4149056) variants significantly influenced peak neonatal peripheral blood bilirubin concentrations under both dominant and additive genetic models ($P < 0.001$).however, reports on the association of the SLCO1B1 388A>G (rs2306283) locus are contradictory, showing inconsistent patterns of results; nevertheless, variations at these specific loci do not systematically negate the overall haplotype study findings.For example, Weng et al. and Zhou et al.[35-36] analyzed newborns from Taipei and Fujian, respectively, and found no significant association between this locus and peak bilirubin levels under additive, dominant, or recessive

genetic models (all P > 0.05).In contrast, a cohort study by Boo et al.[37] involving 580 Malaysian newborns showed that the frequency of the 388G allele reached 26.3%, and carriers had a 1.6-fold increased risk of hyperbilirubinemia (P = 0.01).These differing results suggest that the 388A>G locus may exhibit population-specific effects; however, the scope of this variation is limited and does not constitute a complete contradiction in conclusions. A local study in Hainan (conducted on Han Chinese health examination participants) found[38] that the frequency of the SLCO1B1 *15 risk allele was approximately 6.5%, which is similar to that of Han Chinese in South and Central China but significantly lower than that of the Bai ethnic group in Yunnan (4.8% vs. approximately 13%). This indicates that Han Chinese on the island are consistent with mainland Han Chinese but exhibit regional differences compared to ethnic minorities in Southwest China. Therefore, future studies could further investigate the relationship between SLCO1B1 gene polymorphisms and the incidence of jaundice in specific local populations.

The matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) platform used in this study enables rapid and accurate detection of common genetic variants associated with neonatal hyperbilirubinemia, demonstrating high sensitivity and specificity.This method can effectively distinguish between G6PD wild-type, hemizygous, homozygous, and heterozygous genotypes. Its inherent high-throughput capability allows for the evaluation of up to 192 samples per chip, thereby reducing the estimated turnaround time to 8 hours—a significant improvement over traditional DNA-based screening methods.Furthermore, its high analytical sensitivity enables direct application to dried blood spots (DBS), eliminating the need to recall patients or parents for venous blood collection. Mechanistically, MALDI-TOF MS products terminate after a single nucleotide extension, reducing the probability of mismatches and improving amplification efficiency compared to standard PCR products[39]. Despite these clinical advantages,

this study still has inherent limitations: the single-center, retrospective sampling approach limits the generalizability of the results and the sample size, potentially masking pathogenic variants with low population prevalence or weak effects. Furthermore, since multiple SNPs were identified within a single gene, the lack of haplotype analysis is another

limitation of this study; future research should investigate the relationship between haplotypes and bilirubin levels.This study also has several limitations: (1) The lack of gene-gene interaction assessment precludes analysis of the synergistic effects of combined pathogenic variants in G6PD, UGT1A1, and SLCO1B1 on hyperbilirubinemia. (2) The cross-sectional study design did not track bilirubin peaks over time, which may underestimate the true effects of genetic loci.(3) Environmental factors (such as infection and inadequate feeding) were not collected, limiting the analysis of gene-environment interactions. To address these limitations, future studies should focus on five key areas: First, validate the reproducibility of the current findings by expanding the sample size and conducting multicenter prospective cohort studies. Second, precisely define the bilirubin peak phenotype using time-series data collection protocols combined with nonlinear mixed-effects models.Third, systematically investigate the epistatic effects of single-nucleotide polymorphisms (SNPs) and structural variants across the entire bilirubin metabolism pathway using polygenic risk scores (PRS) and machine learning methods.Fourth, elucidate gene-environment interactions by simultaneously recording variables (such as perinatal infections, breastfeeding rates, and birth weight) and analyzing them via additive interactions and exposure-weighted regression. Finally, combine high-throughput sequencing with functional validation experiments, prioritizing low-frequency or rare variants with regulatory potential, to provide a basis for personalized neonatal jaundice interventions.

Conclusion

This study analyzed the relationship between 10 potential risk factors and neonatal hyperbilirubinemia, identifying the G6PD gene 95A>G and 1388G>A, the UGT1A1 gene 211G>A, and the SLCO1B1 gene -11187G>A loci as major risk factors for neonatal hyperbilirubinemia.Pathological neonatal hyperbilirubinemia is a complex pediatric condition resulting from the combined effects of one or more factors. This study has deepened our understanding of jaundice, and future research will further elucidate the relationship between gene-gene interactions and gene-environment interactions and neonatal hyperbilirubinemia.

References

[1] Sarici SU. Incidence and etiology of neonatal hyperbilirubinemia. J Trop Pediatr. 2010;56(2):128-129. doi:10.1093/tropej/fmp041

[2] Huo Yixuan, Peng Cheng, Hou Xinlin, et al. Revision of the American Academy of Pediatrics Clinical Practice Guidelines on Neonatal Hyperbilirubinemia: Management of Hyperbilirubinemia in Neonates at 35 Weeks' Gestational Age and Older [J]. Chinese Journal of Neonatology, 2023, 38(9): 513-524. DOI: 10.3760/cma.j.issn. 2096-2932.2023.09.001.

[3] Atasilp C, Kanjanapipak J, Vichayaprasertkul J, et al. Associations between UGT1A1 and SLCO1B1 polymorphisms and susceptibility to neonatal hyperbilirubinemia in the Thai population. BMC Pediatrics. 2022;22(1):243. Published 2022 May 2. doi:10.1186/s12887-022-03311-4.

[4] Scrafford CG, Mullany LC, Katz J, et al. Incidence of and risk factors for neonatal jaundice among newborns in southern Nepal. Trop Med Int Health. 2013;18(11):1317-1328. doi:10.1111/tmi.12189

[5] Muchowski KE. Evaluation and treatment of neonatal hyperbilirubinemia. Am Fam Physician. 2014;89(11):873-878.

[6] Kemper AR, Newman TB, Slaughter JL, et al. Clinical Practice Guideline Revision: Management of Hyperbilirubinemia in the Newborn Infant 35 or More Weeks of Gestation. Pediatrics. 2022;150(3):e2022058859. doi:10.1542/peds.2022-058859

[7] Bancone G, Gornsawun G, Peerawaranun P, et al. Contribution of genetic factors to high rates of neonatal hyperbilirubinemia on the Thailand-Myanmar border. PLOS Global Public Health. 2022;2(6):e0000475. Published 2022 Jun 17. doi:10.1371/journal.pgph.00004757.

[8] Chávez-Peña T, Martínez-Camberos A, Cossio-Gurrola G, et al. Prevalence of UGT1A1 (TA)n promoter polymorphism in Panamanian neonates with G6PD deficiency. J Genet. 2020;99:63.

[9] Vukovic M, Radlovic N, Lekovic Z, et al. UGT1A1 (TA)n Promoter Genotype: Diagnostic and Population Pharmacogenetic Marker in Serbia. Balkan J Med Genet. 2018;21(1):59-68. Published 2018 Oct 29. doi:10.2478/bjmg-2018-0012

[10] Pasha YZ, Kacho MA, Niaki HA, Tarighati M, Alaee E. The Association between

Prolonged Jaundice and TATA Box Dinucleotide Repeats in Gilbert's Syndrome. J Clin Diagn Res. 2017;11(9):GC05-GC07. doi:10.7860/JCDR/2017/19376.10597

[11] Amandito R, Rohsiswatmo R, Carolina E, Maulida R, Kresnawati W, Malik A. Profiling of UGT1A1*6, UGT1A1*60, UGT1A1*93, and UGT1A1*28 Polymorphisms in Indonesian Neonates With Hyperbilirubinemia Using Multiplex PCR Sequencing. Front Pediatr. 2019;7:328. Published 2019 Aug 7. doi:10.3389/fped.2019.00328

[12] Fan J, He HY, Li HH, et al. Associations between UGT1A1, SLCO1B1, SLCO1B3, BLVRA, and HMOX1 polymorphisms and susceptibility to severe neonatal hyperbilirubinemia in the Chinese Han population. BMC Pediatr. 2024;24(1):82. Published 2024 Jan 26. doi:10.1186/s12887-024-04537-0

[13] Watchko JF, Tiribelli C. Bilirubin-induced neurologic damage—mechanisms and management approaches. N Engl J Med. 2013;369(21):2021-2030. doi:10.1056/NEJMra1308124

[14] Li Y, Wu T, Chen L, Zhu Y. Associations between G6PD, OATP1B1, and BLVRA variants and susceptibility to neonatal hyperbilirubinemia in a Chinese Han population. J Paediatr Child Health. 2019;55(9):1077-1083. doi:10.1111/jpc.14346

[15] Xu JX, Lin F, Chen ZK, et al. Co-inheritance of G6PD deficiency and the 211G variant of UGT1A1 in neonates with hyperbilirubinemia in eastern Guangdong. BMC Pediatrics. 2021;21(1):564. Published 2021 Dec 11. doi:10.1186/s12887-021-03010-6

[16] Shanghai Children’s Hospital Multicenter Research Group. Association between UGT1A1 gene polymorphisms and unexplained severe neonatal hyperbilirubinemia [J]. Chinese Journal of Neonatology, 2025, 40(7): 481-486.

[17] Zhang Haixia, Zhao Xin, Yang Zhi, et al. A study on the association between neonatal hyperbilirubinemia and the organic anion transporter 1B1 T521C/A388G polymorphism. Chinese Journal of Pediatrics, 2010, 48(09):650-655. DOI:10.3760/cma.j.issn.0578-1310.2010.09.004

[18] Xiao Zhiqi, Guo Hongchuang, Li Lianxiang, et al. Study on the Relationship Between G6PD Activity and Polymorphisms in the UGT1A1, SLCO1B1, and ABCC2 Genes and Neonatal Hyperbilirubinemia [J]. Journal of Prenatal Diagnosis and Treatment, 2018, 10(5):163-168.

[19] Zhang W, He YJ, Gan Z, et al. OATP1B1 polymorphism is a major determinant of serum bilirubin level but not associated with rifampicin-mediated bilirubin elevation. Clin Exp Pharmacol Physiol. 2007;34(12):1240-1244. doi:10.1111/j.1440-1681.2007.04798.x

[20] Rodrigo MA, Zitka O, Krizkova S, Moulick A, Adam V, Kizek R. MALDI-TOF MS as an evolving cancer diagnostic tool: a review. J Pharm Biomed Anal. 2014;95:245-255.

[21] Jang KS, Kim YH. Rapid and robust MALDI-TOF MS techniques for microbial identification: a brief overview of their diverse applications. J Microbiol. 2018;56(4):209-216.

[22] Zhong XY, Holzgreve W. MALDI-TOF MS in Prenatal Genomics. Transfus Med Hemother. 2009;36(4):263-272.

[23] Jon F, Watchko, Zhili, Lin. Exploring the genetic architecture of neonatal hyperbilirubinemia. [J]. Semin Fetal Neonatal Med, 2009, 15: 0.

[24] Zhou J, Yang C, Zhu W, et al. Identification of Genetic Risk Factors for Neonatal Hyperbilirubinemia in Fujian Province, Southeastern China: A Case-Control Study. [J]. Biomed Res Int, 2018,:7803175.

[25] Fan J, He HY, Li HH, et al. Associations between UGT1A1, SLCO1B1, SLCO1B3, BLVRA, and HMOX1 polymorphisms and susceptibility to severe neonatal hyperbilirubinemia in the Chinese Han population. BMC Pediatrics. 2024;24(1):82. Published 2024 Jan 26. doi:10.1186/s12887-024-04537-0

[26] Wong FL, Ithnin A, Othman A, Cheah FC. Glucose-6-phosphate dehydrogenase (G6PD)-deficient infants: Enzyme activity and gene variants as risk factors for phototherapy in the first week of life. J Paediatr Child Health. 2017;53(7):705-710. doi:10.1111/jpc.13509

[27] Tang F, Huang YL, Jiang X, et al. Chinese Journal of Pediatrics. 2018;56(5):359-363. doi:10.3760/cma.j.issn.0578-1310.2018.05.010

[28] Mu SC, Chen YL, Tsai LY, Shih YL, Chen ES, Huang CS. Prolonged Hyperbilirubinemia in a Neonate with a Novel Mutation in the UDP-glucuronosyltransferase 1A1 Gene. Neonatology. 2016;109(3):235-238. doi:10.1159/000443365

[29] Pasha YZ, Kacho MA, Niaki HA, Tarighati M, Alaee E. The Association between Prolonged Jaundice and TATA Box Dinucleotide Repeats in Gilbert's Syndrome. J Clin Diagn Res. 2017;11(9):GC05-GC07. doi:10.7860/JCDR/2017/19376.10597

[30] Dai X, Wu C, He Y, et al. A genome-wide association study for serum bilirubin levels and gene-environment interaction in a Chinese population. Genet Epidemiol. 2013;37(3):293-300. doi:10.1002/gepi.21711

[31] Yang H, Wang Q, Zheng L, et al. Clinical Significance of UGT1A1 Genetic Analysis in Chinese Neonates with Severe Hyperbilirubinemia. Pediatr Neonatol. 2016;57(4):310-317. doi:10.1016/j.pedneo.2015.08.008

[32] He Cuihong. Advances in Research on Neonatal Hyperbilirubinemia and Genetic Polymorphisms [J]. Chinese Journal of Contemporary Pediatrics, 2020, 22(3):280-284. DOI:10.7499/j.issn.1008-8830.2020.03.018.

[33] Zhang X, Pu Z, Ge J, Shen J, Yuan X, Xie H. Association of CYP2D6*10, OATP1B1 A388G, and OATP1B1 T521C polymorphisms with overall survival in breast cancer patients after tamoxifen therapy. Med Sci Monit. 2015;21:563-569. Published 2015 Feb 21. doi:10.12659/MSM.893473

[34] Niemi M, Pasanen MK, Neuvonen PJ. Organic anion transporting polypeptide 1B1: a genetically polymorphic transporter of major importance for hepatic drug uptake. Pharmacol Rev. 2011;63(1):157-181. doi:10.1124/pr.110.002857

[35] Weng YH, Chiu YW, Cheng SW, et al. Risk assessment of gene variants for neonatal hyperbilirubinemia in Taiwan[J]. BMC Pediatr, 2016, 16(1): 144

[36] Zhou J, Yang C, Zhu W, et al. Identification of genetic risk factors for neonatal hyperbilirubinemia in Fujian Province, Southeastern China: a case-control study[J]. Biomed Res Int, 2018, 2018: 7803175.

[37] Boo NY, Sin S, Chee SC, et al. Genetic Factors and Delayed TSB Monitoring and Treatment as Risk Factors Associated with Severe Hyperbilirubinemia in Term Neonates Admitted for Phototherapy. J Trop Pediatr. 2020;66(6):569-582. doi:10.1093/tropej/fmaa016.

[38] Ruan Heqiu, Chen Jun, Xia Mengjuan, et al. Analysis of Statin-Related Gene Polymorphisms and Their Clinical Significance in the Han Chinese Population of Hainan

[J]. International Journal of Geriatrics, 2021, 42(3):143-147,176. DOI:10.3969/j.issn.1674-7593.2021.03.004

[39] Xu Mingli. Detection of SMN1 and SMN2 Gene Copy Numbers in Spinal Muscular Atrophy and Identification of a New Deletion Mutation in the α0-Thalassemia Gene Using MALDI-TOF Mass Spectrometry [D]. Southern Medical University, 2020

Table 1

Table 1 Comparison of Demographic and Clinical Characteristics Between the Hyperbilirubinemia Group and the Control Group			
	Hyperbilirubinemia group	Control group	P value
	(N=210)	(N=195)	
	No. (%)	No. (%)	
Sex			0.09
Male	101 (48.09)	111 (56.92)	
Female	109 (51.91)	84 (43.08)	
Ethnicity			1.00
Han	208 (99.05)	193 (98.97)	
Others	2 (0.95)	2 (1.03)	
Gestational age (weeks)			0.04
<37	20 (9.52)	33 (16.92)	
≥37	190 (90.48)	162 (83.08)	
Birth weight (kg)			0.37
<2.5	21 (10.00)	26 (13.33)	
≥2.5	189 (90.00)	169 (86.67)	
Feeding patterns			0.39
Breastfeeding	123 (58.57)	103 (52.82)	
Bottle feeding	67 (31.90)	75 (38.46)	
Mixed feeding	20 (9.53)	17 (8.72)	
ALP (U/L)			0.52
<150	53 (25.24)	45 (23.07)	
150–420	157 (74.76)	149 (76.41)	
>420	0 (0.00)	1 (0.42)	
HCT (%)			<0.001
<35	55 (26.19)	83 (42.56)	
35–45	133 (63.33)	57 (29.13)	
>45	22 (10.48)	55 (28.21)	
Clinical characteristics of the control group and the NHB group. ALP: alkaline phosphatase, HCT: hematocrit. Group comparisons were performed using the χ^2 test; $P < 0.05$ was considered statistically significant.			

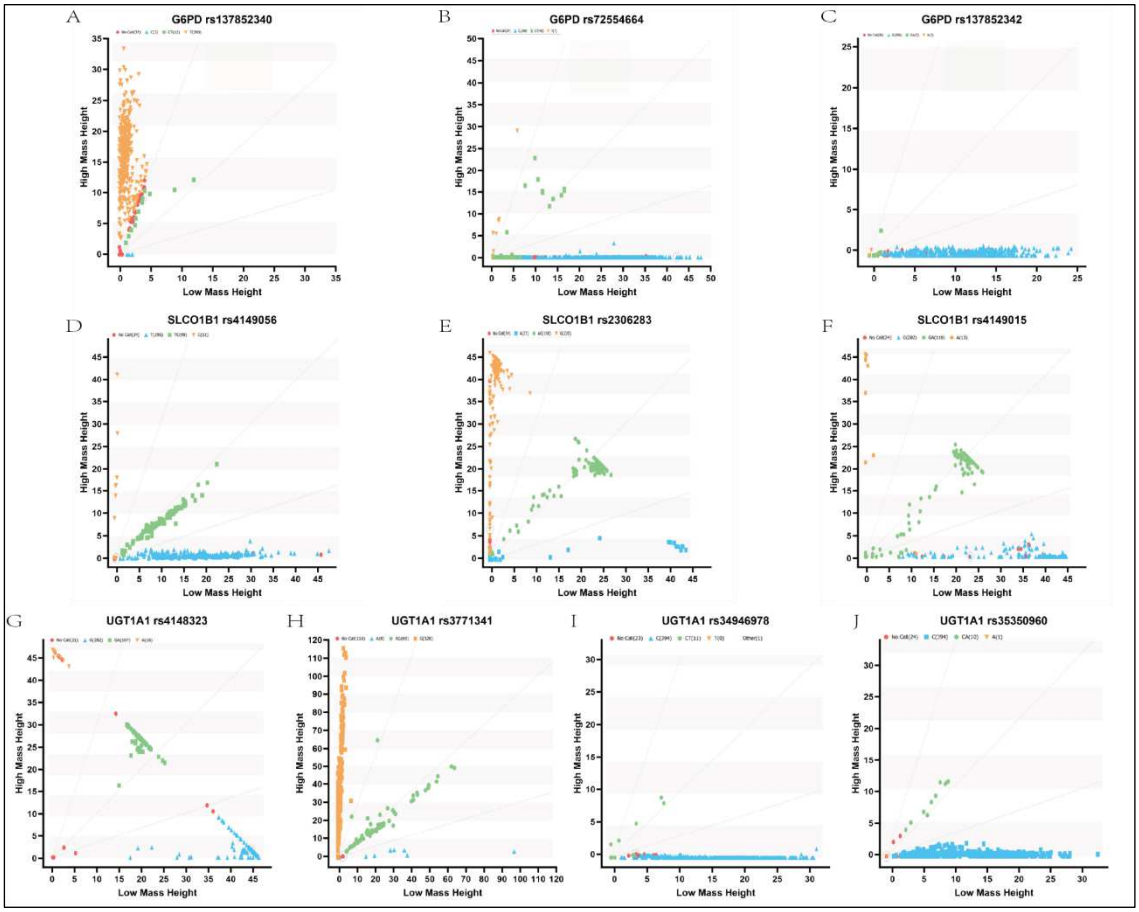


Figure 1 MassARRAY mass spectrometry scatter plots for the G6PD gene (rs137852340, rs72554664, rs137852342), SLCO1B1 gene (rs4149056, rs2306283, rs4149015), and UGT1A1 gene(rs4148323, rs3771341, rs34946978, rs35350960) MassARRAY mass spectrometry scatter plots

Figure 1 shows the MassARRAY mass spectrometry genotyping scatter plots for 10 SNP loci. The descriptions of each sub-figure are as follows:

Figure 1A (G6PD rs137852340, 95A>G): In the case group, there were 11 cases with the AG genotype (5.2%) and 3 cases with the GG genotype (1.4%); in the control group, there was only 1 case with the AG genotype (0.5%).The frequency of the G allele was significantly higher in the case group (0.0405) than in the control group (0.0026), with a statistically significant difference ($P < 0.001$).

Figure 1B (G6PD rs72554664, 1388G>A): In the case group, there were 29 cases with the GA genotype (13.8%) and 5 cases with the AA genotype (2.4%); in the control group, there were 11 cases with the GA genotype (5.6%) and 2 cases with the AA genotype (1.0%).The frequency of the A allele was higher in the case group (0.0929) than in the control group (0.0385), and the difference was statistically significant ($P = 0.01$).

Figure 1C (G6PD rs137852342, 1024C>T): In the case group, there were 6 cases with the CT genotype (2.9%) and 2 cases with the TT genotype (1.0%), while in the control group, there was only 1 case with the CT genotype (0.5%).The T allele frequency was higher in the case group

(0.0238) than in the control group (0.0026), and the difference was statistically significant (P = 0.012).

Figure 1D (SLCO1B1 rs4149056, 521T>C): In the case group, there were 58 cases with the TC genotype (27.6%) and 11 cases with the CC genotype (5.2%); in the control group, there were 40 cases with the TC genotype (20.5%) and no cases with the CC genotype.The frequency of the C allele was significantly higher in the case group (0.1905) than in the control group (0.1026), with a statistically significant difference (P < 0.001).

Figure 1E (SLCO1B1 rs2306283, 388A>G): The genotype distributions in the case and control groups were similar, with A allele frequencies of 0.2524 and 0.2718, respectively; the difference was not statistically significant (P = 0.58).

Figure 1F (SLCO1B1 rs4149015, -11187G>A): In the case group, 72 cases (34.3%) had the GA genotype and 9 cases (4.3%) had the AA genotype; in the control group, 38 cases (19.5%) had the GA genotype and 4 cases (2.1%) had the AA genotype.The frequency of the A allele was significantly higher in the case group (0.2143) than in the control group (0.1179), with a statistically significant difference (P < 0.001).

Figure 1G (UGT1A1 rs4148323, 211G>A): In the case group, there were 71 cases with the GA genotype (33.8%) and 11 cases with the AA genotype (5.2%); in the control group, there were 36 cases with the GA genotype (18.5%) and 5 cases with the AA genotype (2.6%).The frequency of the A allele was significantly higher in the case group (0.2214) than in the control group (0.1179), with a statistically significant difference (P < 0.001).

Figure 1H (UGT1A1 rs3771341, 865-2441G>A): The genotype distributions in the case and control groups were similar, with A allele frequencies of 0.1000 and 0.1103, respectively; the difference was not statistically significant (P = 0.03, but not significant after correction for multiple comparisons).

Figure 1I (UGT1A1 rs34946978, 1091C>T): The CT genotype was observed in 9 cases (4.3%) and 2 controls (1.0%).The frequency of the T allele was higher in the case group (0.0214) than in the control group (0.0051), but the difference was not statistically significant (P = 0.07).

Figure 1J (UGT1A1 rs35350960, 686C>A): The genotype distributions were similar between the case and control groups, with A allele frequencies of 0.0167 and 0.0128, respectively; the difference was not statistically significant (P = 0.78).

Table S1. Detailed information of the 12 selected SNP loci in G6PD, UGT1A1, and SLCO1B1 genes.

Gene	SNP	rs ID	Chromosome	Position (GRCh37)	Region	Amino Acid Change
G6PD	95A>G	rs137852340	Xq28	153764636	Exon 2	His32Arg
G6PD	1388G>A	rs72554664	Xq28	153764929	Exon 12	Arg459His
G6PD	1376G>T	rs72554665	Xq28	153764917	Exon 12	Arg459Leu
G6PD	871G>A	rs137852327	Xq28	153764412	Exon 8	Val291Met

G6PD	1024C>T	rs137852342	Xq28	153764565	Exon 10	Leu342Phe
UGT1A1	865-2441G>A	rs3771341	2q37	234668879	Intron 1	—
UGT1A1	1091C>T	rs34946978	2q37	234669353	Exon 3	Pro364Leu
UGT1A1	686C>A	rs35350960	2q37	234669118	Exon 2	Pro229Gln
UGT1A1	211G>A	rs4148323	2q37	234668879	Exon 1	Gly71Arg
SLCO1B1	-11187G>A	rs4149015	12p12	21353647	Promoter	—
SLCO1B1	388A>G	rs2306283	12p12	21341303	Exon 4	Asn130Asp
SLCO1B1	521T>C	rs4149056	12p12	21341946	Exon 6	Val174Ala

Note: SNPs were selected from the coding region, promoter region, and 2-kb regulatory region upstream and downstream of each gene based on the NCBI dbSNP and gnomAD databases. All positions are reported according to the GRCh37/hg19 human genome assembly.

Table 2 Comparison of genotype and allele frequency distributions for G6PD, UGT1A1, and SLCO1B1 between the case group and the control group [n(%)]

Gene	SNPs	Genotypes/ Alleles	Hyperbilirubinemia group (N=210)	Control group (N=195)	P-value
G6PD	95A>G (rs137852340)				0.01
		AA	196 (0.9330)	194 (0.9950)	
		AG	11 (0.0520)	1 (0.0050)	
		GG	3 (0.0140)	0 (0.0000)	
					<0.001
		A allele	403 (0.9595)	389 (0.9974)	
		G allele	17 (0.0405)	1 (0.0026)	
	1388G>A (rs72554664)				0.01
		GG	176 (0.8380)	182 (0.9330)	
		GA	29 (0.1380)	11 (0.0560)	
		AA	5 (0.0240)	2 (0.0100)	
					0.01
		G allele	381 (0.9071)	375 (0.9615)	
		A allele	39 (0.0929)	15 (0.0315)	
	1024C>T (rs137852342)				0.08

		CC	202 (0.9620)	194 (0.9950)	
		CT	6 (0.0290)	1 (0.0050)	
		TT	2 (0.0100)	0 (0.0000)	
					0.01
		C allele	410 (0.9762)	389 (0.9974)	
		T allele	10 (0.0238)	1 (0.0026)	
	211G>A (rs4148323)				<0.001
		GG	128 (0.6100)	154 (0.7900)	
		GA	71 (0.3380)	36 (0.1850)	
		AA	11 (0.0520)	5 (0.0260)	
					<0.001
		G allele	327 (0.7786)	344 (0.8821)	
		A allele	93 (0.2214)	46 (0.1179)	
	865-2441G>A (rs3771341)				0.20
		GG	123 (0.5860)	99 (0.5080)	
		GA	41 (0.1950)	39 (0.2000)	
		AA	46 (0.2190)	57 (0.2920)	
		G allele	287 (0.6833)	237 (0.6077)	0.03
UGT1A1		A allele	133 (0.3167)	153 (0.3923)	
	1091C>T (rs34946978)				0.09
		CC	201 (0.9570)	193 (0.9900)	
		CT	9 (0.0430)	2 (0.0100)	
		TT	0 (0.00)	0 (0.00)	
		C allele	411 (0.9786)	388 (0.9949)	0.07
		T allele	9 (0.0211)	2 (0.0051)	
	686C>A (rs35350960)				0.62
		CC	204 (0.9710)	190 (0.9740)	
		CA	5 (0.0240)	5 (0.0260)	
		AA	1 (0.0050)	0 (0.0000)	
					0.78
		C allele	413 (0.9833)	385 (0.9872)	
		A allele	7 (0.0214)	5 (0.0218)	
SLCO1B1	521T>C (rs4149056)				<0.001

		TT	141 (0.6710)	155 (0.7950)	
		TC	58 (0.2760)	40 (0.2050)	
		CC	11 (0.0520)	0 (0.0000)	
					<0.001
		T allele	340 (0.8095)	350 (0.8974)	
		C allele	80 (0.1905)	40 (0.1026)	
	388A>G (rs2306283)				0.72
		AA	14 (0.0670)	13 (0.0670)	
		AG	78 (0.3710)	80 (0.4100)	
		GG	118 (0.5620)	102 (0.5230)	
					0.58
		A allele	106 (0.2524)	106 (0.2718)	
		G allele	314 (0.7476)	284 (0.7282)	
	-11187G>A (rs4149015)				<0.001
		GG	129 (0.6140)	153 (0.7850)	
		GA	72 (0.3430)	38 (0.1950)	
		AA	9 (0.0430)	4 (0.0210)	
					<0.001
		G allele	330 (0.7857)	344 (0.8821)	
		A allele	90 (0.2143)	46 (0.1179)	

Intergroup comparisons of genotype and allele frequencies were performed using the χ^2 test; $P < 0.05$ was considered statistically significant. Since G6PD is an X-linked gene, statistical analysis was stratified by sex: males were counted as having 1 allele, and females as having 2 alleles, when calculating allele frequencies.In genotype statistics, male heterozygotes (A/G) are uniformly denoted as homozygotes (AA/GG) in the table format. Note: Allele frequencies are reported as decimals (e.g., 0.9595 represents 95.95%), consistent with the gnomAD database format; genotype frequencies are expressed as percentages (e.g., 0.0520 represents 5.20%). Note: Allele frequencies are reported as decimals (e.g., 0.9595 represents 95.95%), consistent with the gnomAD database format; genotype frequencies are expressed as percentages (e.g., 0.0520 represents 5.20%). Note: Allele frequencies are reported as decimals (e.g., 0.9595 represents 95.95%), consistent with the gnomAD database format;genotype frequencies are expressed as ratios (e.g., 0.0520 represents 5.20%). Note: Allele frequencies are reported as decimals (e.g., 0.9595 represents 95.95%), consistent with the gnomAD database format; genotype frequencies are expressed as ratios (e.g., 0.0520 represents 5.20%). Note: Allele frequencies are reported as decimals (e.g., 0.9595 represents 95.95%), consistent with the gnomAD database format; genotype frequencies are expressed as ratios (e.g., 0.0520 represents 5.20%). Note: Allele frequencies are reported as decimals (e.g., 0.9595 represents 95.95%), consistent with the gnomAD database format; genotype frequencies are expressed as ratios (e.g., 0.0520 represents 5.20%).

Table 3 Correlation analysis between G6PD, UGT1A1, and SLCO1B1 genotypes and bilirubin levels

Gene	SNPs	Genotypes	Bilirubin level (μmol/L)			
			Median	IQR	Maximum	P-value
G6PD	95A>G (rs137852340)					0.003
		AA	221.46	[47.61, 25.00]	384.86	
		AG	265.99	[249.36, 286.55]	325.53	
		GG	276.51	[253.55, 289.20]	301.89	
	1388G>A (rs72554664)					0.06
		GG	210.59	[48.58, 250.49]	371.8	
		GA	241.44	[157.78, 281.38]	384.86	
		AA	227.12	[124.64, 245.39]	296.48	
	1024C>T (rs137852342)					0.30
		CC	221.73	[48.32, 255.18]	384.86	
		CT	233.91	[226.24, 279.74]	313.73	
		TT	236.76	[232.64, 240.87]	244.98	
UGT1A1	211G>A (rs4148323)					<0.001
		GG	194.8	[39.96, 240.35]	384.86	
		GA	232.9	[169.94, 273.43]	366.99	
		AA	262.3	[203.99, 312.68]	377.46	
	865-2441G>A (rs3771341)					0.36
		GG	222.78	[52.81, 259.22]	384.86	
		GA	221.63	[37.2, 240.11]	337.77	
		AA	197.70	[167.86, 225.04]	256.79	
	1091C>T (rs34946978)					0.04
		CC	221.73	[47.85, 253.39]	384.86	
		CT	255.59	[229.54, 272.75]	316.30	
	686C>A (rs35350960)					0.23
		CC	222.18	[54.72, 255.30]	384.86	
		CA	142.26	[26.93, 245.37]	269.34	
		AA	313.73	[313.73, 313.73]	313.73	
	521T>C (rs4149056)					<0.001
		TT	202.45	[40.71, 244.24]	384.86	
		TC	228.73	[104.29, 267.70]	366.99	

SLCO1B1		CC	279.23	[230.38, 302.49]	325.53	
	388A>G (rs2306283)					0.74
		AA	223.83	[83.87, 247.67]	369.78	
		AG	208.25	[37.13, 252.89]	371.80	
		GG	222.78	[80.42, 255.89]	384.86	
	-11187G>A (rs4149015)					0.001
		GG	198.25	[38.39, 244.91]	384.86	
		GA	229.23	[175.57, 266.72]	366.84	
		AA	234.31	[154.45, 278.29]	334.05	
The Kruskal-Wallis test was used to analyze differences in bilirubin levels among different genotypes; $P < 0.05$ indicates a statistically significant difference;Since G6PD is an X-linked gene, the statistical analysis accounted for sex: when calculating allele frequencies, males were counted as having 1 allele and females as having 2 alleles; in genotype statistics, male heterozygotes (A/G) were uniformly denoted as homozygotes (AA/GG) in the table format. IQR: interquartile range (25–75th percentile).						

Table 4 Results of logistic regression analysis on the association between three genetic models and neonatal hyperbilirubinemia

Gene	SNPs	Models	OR (95% CI)	P-value
G6PD	95A>G (rs137852340)	Dominant model	13.86 (2.74–252.34)	0.01
		Recessive model	-	-
		Additive model	11.44 (2.46–205.01)	0.02
	1388G>A (rs72554664)			
		Dominant model	2.71 (1.41–5.47)	<0.001
		Recessive model	2.35 (0.50–16.05)	0.31
		Additive model	2.23 (1.28–4.13)	0.01
	1024C>T (rs137852342)			
		Dominant model	7.68 (1.39–143.20)	0.06
		Recessive model	-	-
		Additive model	6.44 (1.38–115.47)	0.07
UGT1A1	211G>A (rs4148323)			
		Dominant model	2.41 (1.55–3.77)	<0.001

		Recessive model	2.10 (0.75–6.77)	0.18
		Additive model	2.05 (1.41–3.04)	<0.001
	865-2441G>A (rs3771341)			
		Dominant model	0.94 (0.57–1.55)	0.44
		Recessive model	0.55 (0.11–2.27)	0.42
		Additive model	0.90 (0.64–1.02)	0.07
	1091C>T (rs34946978)			
		Dominant model	4.32 (1.09–28.58)	0.06
		Recessive model	-	-
		Additive model	4.32 (1.09–28.58)	0.06
	686C>A (rs35350960)			
		Dominant model	1.12 (0.33–3.93)	0.86
		Recessive model	-	-
		Additive model	1.26 (0.43–4.05)	0.67
SLCO1B1	521T>C (rs4149056)			
		Dominant model	1.89 (1.21–2.99)	0.01
		Recessive model	-	-
		Additive model	2.02 (1.35–3.08)	<0.001
	388A>G (rs2306283)			
		Dominant model	1.00 (0.45–2.19)	0.99
		Recessive model	1.69 (0.79–1.73)	0.43
		Additive model	1.11 (0.81–1.52)	0.53
	-11187G>A (rs4149015)			
		Dominant model	2.29 (1.48–3.57)	<0.001
		Recessive model	2.14 (0.68–7.99)	0.21
		Additive model	2.02 (1.38–3.02)	<0.001

Table note: To ensure data stability, models for groups with fewer than 10 individuals in any variant category were not analyzed; therefore, some models are marked as "-". Dominant model: mutation homozygotes + heterozygotes vs. wild-type homozygotes; recessive model: mutation homozygotes vs. heterozygotes + wild-type homozygotes; additive model: increasing by allelic dose. The logistic regression in this table was not adjusted for confounding factors.OR: odds ratio; CI: confidence interval;

Table 5 Results of logistic regression analysis on the association between genotype frequencies and neonatal hyperbilirubinemia using crude or adjusted models

Table 5 Results of logistic regression analysis on the association between genotype frequencies and neonatal hyperbilirubinemia using crude or adjusted models

Gene	SNPs	Genotypes/ Alleles	Crude OR (95% CI)	P-value	Adjusted OR (95% CI)	P-value
G6PD	95A>G (rs137852340)					
		AA	Ref		Ref	
		AG	10.89 (2.09–199.98)	0.02	9.67 (1.70–182.99)	0.04
		A	Ref		Ref	
		G	11.45 (2.46–205.10)	0.02	10.37 (2.15–189.26)	0.02
	1388G>A (rs72554664)					
		GG	Ref		Ref	
		GA	2.73 (1.36–5.86)	0.006	2.32 (1.05–5.42)	0.04
		AA	2.59 (0.55–18.22)	0.26	4.46 (0.80–35.41)	0.11
		G	Ref		Ref	
		A	2.23 (1.28–4.13)	0.007	2.22 (1.23–4.25)	0.01
	1024C>T (rs137852342)					
		CC	Ref		Ref	
		CT	5.76 (0.97–109.36)	0.11	4.21 (0.64–83.63)	0.20
		C	Ref		Ref	
		T	6.44 (1.38–115.47)	0.07	5.04 (1.09–91.04)	0.11
UGT1A1	211G>A (rs4148323)					
		GG	Ref		Ref	
		GA	2.37 (1.50–3.81)	<0.001	2.17 (1.33–3.59)	0.002
		AA	2.65 (0.94–8.58)	0.08	1.72 (0.57–5.89)	0.36
		G	Ref		Ref	
		A	2.05 (1.41–3.04)	<0.001	1.79 (1.20–2.72)	0.005

	865-2441G>A (rs3771341)					
		GG	Ref		Ref	
		GA	1.00 (0.60–1.69)	1	0.86 (0.49–1.51)	0.60
		AA	0.55 (0.11–2.29)	0.42	0.54 (0.10–2.39)	0.43
		G	Ref		Ref	
		A	0.90 (0.59–1.39)	0.65	0.81 (0.51–1.38)	0.38
	1091C>T (rs34946978)					
		CC	Ref		Ref	
		CT	4.32 (1.10–28.58)	0.06	5.2 (1.17–38.13)	0.05
		C	Ref		Ref	
		T	4.32 (1.10–28.58)	0.06	5.2 (1.17–38.13)	0.05
	686C>A (rs35350960)					
		CC	Ref		Ref	
		CA	0.93 (0.25–3.39)	0.91	0.71 (0.18–2.83)	0.62
		C	Ref		Ref	
		A	1.26 (0.43–4.05)	0.68	0.92 (0.29–3.23)	0.89
SLCO 1B1	521T>C (rs4149056)					
		TT	Ref		Ref	
		TC	1.59 (1.01–2.55)	0.048	1.39 (0.85–2.31)	0.18
		T	Ref		Ref	
		C	2.02 (1.36–3.08)	<0.001	1.847 (1.20–2.89)	0.006
	388A>G (rs2306283)					
		AA	Ref		Ref	
		AG	0.91 (0.39–2.06)	0.81	0.83 (0.34–2.06)	0.692
		GG	1.07 (0.48–2.40)	0.86	0.89 (0.36–2.16)	0.80
		A	Ref		Ref	

		G	1.11 (0.81–1.52)	0.528	1.01 (0.71–1.42)	0.98
	-11187G>A (rs4149015)					
		GG	Ref		Ref	
		GA	2.25 (1.43–3.58)	<0.001	2.39 (1.47–3.97)	<0.001
		AA	2.67 (0.85–10.03)	0.109	2.36 (0.69–9.48)	0.19
		G	Ref		Ref	
		A	2.02 (1.38–3.02)	<0.001	2.01 (1.37–3.19)	<0.001

Table Notes: Statistical methods: Binary logistic regression analysis was used to calculate crude odds ratios and adjusted odds ratios, with 95% confidence intervals provided; Adjustment factors: gestational age, hematocrit; Reference group: The wild-type homozygous genotype/wild-type allele of each gene was used as the reference (labeled as Ref); OR: odds ratio; CI: confidence interval; $P < 0.05$ indicates statistical significance;The wide range of 95% CIs for some loci is related to the small sample size for the corresponding genotypes.

Statements

Data availability statement

All data generated or analyzed during this study are included in this article and its supplementary material. No additional raw data repository deposition is required or applicable, as all relevant genotyping results, statistical analyses, and summary tables are fully presented within the published article and its supplementary files.

Ethics statement

The studies involving humans were approved by the Ethics Committee of Haikou Maternal and Child Health Hospital (IRB approval No. [2024]01045). Written informed consent to participate was obtained from the participants' legal guardians. Human Ethics and Consent to Participate declarations: This study was conducted in accordance with the Declaration of

Helsinki. All participants provided written informed consent to participate. Consent to Participate: Written informed consent was obtained from the legal guardians of all participants. Ethics Approval: The Ethics Committee of Haikou Maternal and Child Health Hospital (Internal Review Board). Clinical trial number: not applicable.

Author contributions

WW: Writing - original draft, Writing - review & editing. LS: Formal analysis, Investigation, Writing - review & editing. ZX: Data curation, Writing - original draft. YH: Supervision, Writing - review & editing. WL: Data curation, Project administration, Writing - review & editing.

Funding Declaration

Funding Declaration: This study was supported by the Fostering Talents Project of Haikou Maternal and Child Health Hospital (RC24Y01).

Conflict of interest

The authors declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.