# Genetic study of intrahepatic cholestasis of pregnancy in 101,023 Chinese women unveils East Asian-specific etiology linked to historic HBV infection

Yanhong Liu[1*], Yuandan Wei[1,2*], Xiaohang Chen[3*], Shujia Huang[4*], Yuqin Gu[1], Zijing Yang[1,3], Liang Hu[3], Xinxin Guo[1], Hao Zheng[1], Mingxi Huang[4], Shangliang Chen[5], Tiantian Xiao[6], Yang Zhang[1], Guo-Bo Chen[7], Likuan Xiong[2,8], Xiu Qiu[4#], Fengxiang Wei[3,9#], Jianxin Zhen[2#], Siyang Liu[1#]

1. School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, Guangdong 518107, China
2. Central Laboratory, Shenzhen Baoan Women's and Children's Hospital, Shenzhen, Guangdong 518102, China
3. The Genetics Laboratory, Longgang District Maternity & Child Healthcare Hospital of Shenzhen City, Shenzhen, Guangdong, 518172, China
4. Division of Birth Cohort Study, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, 510623, China
5. Department of transfusion, Shenzhen Baoan Women's and Children's Hospital, Shenzhen, Guangdong 518102, China
6. Xiangya School of Medicine, Central South University, Changsha, Hunan 410078, China

7. Center for Productive Medicine, Department of Genetic and Genomic Medicine, Clinical Research Institute, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou 310014, Zhejiang, China
8. Shenzhen Key Laboratory of Birth Defects Research, Shenzhen, Guangdong 518102, China
9. Longgang Maternity and Child Institute of Shantou University Medical College, Shenzhen, Guangdong, 518172, China

*: Those authors contribute equally as co-first authors

#: Correspondence can be addressed to

Siyang Liu liusy99@mail.sysu.edu.cn

Jianxin Zhen jxzhen@qq.com

Fengxiang Wei haowei727499@163.com

Xiu Qiu xiu.qiu@bigcs.org

## Data availability

GWAS summary statistics for TBA and ICP phenotypes will be made publicly available in the GWAS catalog (https://www.ebi.ac.uk/gwas/) upon publication.

## Author contribution

SL, FW, JZ, and XQ conceived the study. JZ, XC, LH, YW, ZY, HZ, SC, and TX collected and organized the data from the maternity testing system. YL, YG, XC, LH, and XG conducted data pre-processing and preliminary analyses. YL performed all statistics, GWAS, and evolutionary analysis. YL and YW performed the visualization of all results. JZ, SH, MH, and XQ provided the validation data. SL, FW, JZ, XQ, YZ, GBC, and LX provided professional guidance and interpretation of data. YL & SL wrote the manuscript with input from all authors. All authors contributed to manuscript revisions and approved the final version of the article. SL is responsible for the integrity of the work as a whole.

## Competing interests

All authors declare no competing financial interests.

## Financial support statement

The study was supported by Shenzhen Basic Research Foundation (20220818100717002), Guangdong Basic and Applied Basic Research Foundation (2022B1515120080, 2020A1515110859), National Natural Science Foundation of China (31900487, 82203291), and the Shenzhen Health Elite Talent Training Project.

## Abstract

### Background & Aims

Intrahepatic cholestasis of pregnancy (ICP) is the most common and high-risk liver disorder during the critical period of human reproduction. Despite varying prevalence across populations, a mechanistic understanding of this phenomenon is lacking. This study delves into the genetic etiology of ICP in East Asians, drawing comparisons with Europeans to comprehend ICP etiology in the context of genetic background and evolution.

### Methods

We conducted the hitherto largest-scale genome-wide association studies (GWAS) on total bile acid concentration (TBA) and ICP among 101,023 Chinese pregnancies. The findings were subsequently replicated in two cohorts and compared with European populations. Additionally, phenome-wide association and spatio-temporal evolution analyses were employed to understand the function and explore evolutionary pattern of sites associated with ICP.

### Results

We identified eight TBA and five ICP loci, including ten novel loci. Notably, we found an East-Asian-specific genetic locus at 14q24.1, contributing to a 6.41 μmol/L increase in TBA and a 15.23-fold higher risk of ICP per risk allele (95% *CI*: 15.10 to

15.36, $P = 9.23\times10^{-375}$). Phenome-wide association studies and spatial-temporal evolution analyses revealed that the 14q24.1 ICP risk locus exhibits resistance to hepatitis B infection and has become prevalent only within the last 3,000 years in East and Southeast Asia.

## Conclusions

Our investigations have unraveled a distinct etiology of ICP between Europeans and East Asians, and has linked ICP etiology in East Asians to a historical HBV epidemic in East and Southeast Asia within the last 3,000 years. These findings lay the groundwork for an improved biological understanding of ICP pathophysiology. Further exploration and utilization of these variations hold the potential for more precise detection, assessment, and treatment of ICP.

## Lay summary

Intrahepatic cholestasis of pregnancy (ICP) is a prevalent and high-risk liver disorder that occurs during pregnancy, a critical period in human reproduction. It affects approximately 1% to 6.06% pregnancies and has been associated with severe adverse outcomes such as preterm birth and stillbirth. While rare and common variants associated with ICP have been identified in the European population, the genetic basis of ICP in East Asian population remains uncharacterized. Here, we conducted the largest-scale genome-wide association studies to date for TBA and ICP among

101,023 Chinese pregnant women, including 4,703 cases and 96,320 controls from two hospitals in Shenzhen, China. We replicated our findings in two independent Chinese cohorts and compared them with ICP genetic studies in the European population. We identified eight and five genome-wide significant loci for TBA and ICP, respectively, including ten novel loci. Notably, we identified an East-Asian-specific genetic locus contributing to a 6.41 μmol/L increase in TBA per risk allele and a 15.23-fold higher risk of ICP. Further exploration through phenome-wide association studies and spatial-temporal evolution analyses revealed that the 14q24.1 ICP risk locus exhibits resistance to hepatitis B infection and has become prevalent only within the last 3,000 years in East and Southeast Asia. These findings suggest a historical HBV epidemic in East and Southeast Asia within 3,000 years may have contributed to the increased prevalence of ICP and TBA risk alleles among East Asians. Our study unravels a distinct genetic etiology of ICP between Europeans and East Asians. These findings lay the foundation for an improved understanding of ICP pathophysiology and emphasize the need for integrating population evolution into genetic medicine for personalized genomics and clinical guidance.

## Highlights

(1) In the most powerful genome-wide association studies on TBA and ICP in East Asians to date, we identified eight and five genetic loci, respectively, of which, 7 and 3 were novel discoveries.

(2) One of the novel loci, the 14q24.1 locus, stands out as it contains unique causal genetic variants specific to East-Asians. These variants demonstrate large effects, contributing to an average increase of 6.41 μmol/L in TBA per risk allele and a 15.23-fold higher risk of ICP.

(3) The risk mutations associated with ICP at the 14q24.1 exhibit resistance to hepatitis B infection and has only become prevalent within the last 3000 years in East and Southeast Asia.

**Key words**: Intrahepatic cholestasis of pregnancy; Total bile acid; Genetic etiology; Evolutionary history; East Asians; 14q24.1

## Introduction

Intrahepatic cholestasis of pregnancy (ICP) is the most prevalent and high-risk liver disorder occurring during pregnancy, a critical phase in human reproduction[1]. The disease is characterized by maternal pruritus and elevated serum bile acid concentrations (≥10μmol/L), with diagnosis typically involving assessing maternal total serum bile acids (TBA) during routine pregnancy screening in the late second and third trimesters[2]. ICP has been associated with several severe adverse pregnancy outcomes, including increased risks of spontaneous and iatrogenic preterm birth, meconium-stained amniotic fluid, fetal asphyxia, and stillbirth[3,4]. A large-scale meta-analysis of ICP cases revealed that a TBA level of 100 μmol/L is associated with an elevated risk of stillbirth, with a hazard ratio [*HR*] of 30.5 [95% *CI*: 8.83–105.30], compared to a TBA level below 40 μmol/L[5]. Several retrospective studies and case reports underscore the unpredictability of prenatal fetal death in patients with ICP after 36 weeks of pregnancy[2,6,7]. At present, oral ursodeoxycholic acid (UDCA) is the most commonly employed treatment in clinics to alleviate maternal pruritus and improve liver function. However, a recent clinical randomized controlled trial declared that treatment with UDCA does not reduce adverse perinatal outcomes in women with ICP[1], leaving its effectiveness in enhancing perinatal outcomes and the long-term health of mothers with ICP and their children inconclusive[8].

Unraveling new treatments of ICP necessitates a profound understanding of the disease's etiology. Accumulated evidence suggests the complexity of ICP etiology, which is associated with genetic, endocrine, and environmental factors[9]. Familial clustering of ICP strongly implies a genetic basis[10]. Family-based studies indicate that the heterozygous state for an *MDR3* (alias *ABCB4*) gene defect likely represents a genetic predisposition within families[11,12]. A candidate gene analysis study identified six SNPs in each of the *ABCB4* and *ABCB11* genes that demonstrated a significant association with ICP[13]. Case-control study also showed that *ABCB4*, *ABCB11,* and *ATP8B1* genes are related to intrahepatic cholestasis of pregnancy[14]. Case reports have proposed that NTCP deficiency, encoded by *SLC10A1* gene, might be a genetic factor contributing to ICP[15]. A recent genome-wide meta-association study involving 1,138 ICP cases and 153,642 controls from three Europeans studies identified eleven genes associated with ICP risk, most of which are related to hepatic functions[16].

Notably, global prevalence of ICP varies substantially among different ethnicities. ICP affects approximately 0.32% of pregnancies in the United States, 5.6% in the Latina population in Los Angele[17,18], 0.7% in the United Kingdom[17], 1.5%-4% in Chile[19,20] and 1.2% to 6.06% in China[21]. However, despite its variable geographical prevalence, factors influencing these differences are poorly understood. In addition, despite the considerably higher prevalence of the disease among East Asians, no studies of sufficient power have been conducted among the East Asian populations.

To address these gaps, we conducted the most powerful genome-wide association study (GWAS) of TBA and ICP in East Asia to date, utilizing sequencing data from the non-invasive prenatal testing (NIPT) of 101,023 Chinese pregnant women and comprehensive phenotypic records from two hospitals in Shenzhen city in South China. The discoveries were replicated in two independent Chinese datasets and compared to previous findings among the European population. Intriguingly, we found distinct etiology of ICP among Europeans and Asians. Among the ten novel loci previously not reported in the GWAS catalog, a genetic locus in 14q24.1 present exclusively among East Asians stands out, contributing to an average increase of 6.41 μmol/L in TBA per risk allele and a 15.23 times higher risk of ICP in East Asians. Further phenome-wide association and spatio-temporal evolutionary studies have linked this ICP risk locus to a putative HBV epidemic in East and Southeast Asia within the last 3000 years.

# Result

## Study design and TBA phenotypic distribution

The design of the study is summarized in **Abstract Figure**. In Shenzhen city, China, government-sponsored Non-Invasive Prenatal Tests (NIPT) were administered as a standard examination. Over the period from 2017 to 2022, we recruited 101,023 pregnant women during routine obstetric examinations across two Shenzhen hospitals (Baoan and Longgang). Each pregnant woman underwent both a NIPT test and at least one TBA test between the $13^{th}$ and $42^{nd}$ gestational weeks. After excluding outliners, 94,360 pregnancies with normal TBA levels ($< 10\mu mol/L$) were included for the GWAS of plasma TBA levels. Using the inclusion and exclusion criteria detailed in the Methods section, 4,703 pregnancies were identified as ICP cases, while 96,320 individuals served as controls for the ICP GWAS.

**Supplementary Fig. 1** illustrates a positively skewed phenotypic distribution of TBA levels, which exhibits a consistent median of 2.90 μmol/L (Interquartile Range [IQR]: [2.00, 4.20]) among normal pregnancies and a median of 14.52 μmol/L (IQR: [11.58, 23.10]) among ICP cases. The distribution of TBA levels, maternal age (29.72 ± 4.47 years old), and gestational week (33.05 ± 8.48 weeks) are similar between the two hospitals (**Supplementary Table 1)**.

We employed a well-established genetic analysis pipeline developed in our previous studies to analyze the NIPT data[22,23] and conduct the GWAS for TBA level and ICP across over 7 million sequence variants with minor allele frequency (MAF) greater than 1% in the study cohorts. We replicated the genetic findings in two independent Chinese datasets and provided meta-analysis estimates. We further conducted phenome-wide colocalization analyses to examine the genetic associations across a hundred pregnancy phenotypes. For the most significant and East Asian-specific 14q24.1 locus, we conducted spatio-temporal and natural selection evolutionary analyses to explore the frequency changes of the ICP risk alleles globally and since the Holocene.

**Genome-wide association study of TBA and ICP among the 101,023 pregnancies**

Power estimations for the meta-GWAS design indicate an 80% statistical power to detect genetic associations with an effect size ($\beta$) > 0.170 and MAF ≥ 0.05, or otherwise $\beta$ > 0.370 and MAF ≥ 0.01 for TBA, and similarly for genetic associations with an odds ratio (*OR*) > 1.140 and MAF ≥ 0.05 as well as an *OR* > 1.320 and MAF ≥ 0.01 for ICP **(Fig. S2)**. The genomic inflation (lambda) values for TBA and ICP were 1.069 and 1.018, respectively, suggesting negligible confounding of population stratification **(Fig. S3)**.

In total, we identified 1,673 and 1,281 genome-wide significant variants ($P < 5\times10^{-8}$), constituting 19 and 16 independent association signals for TBA and ICP respectively (**Table S2**). These were represented by 8 independent association loci for TBA and 5 for ICP **(Fig. 1; Table 1; Table S3)**. Among the 13 association loci, 7 and 3 were not previously reported in the GWAS catalog or Phenoscanner and were identified as novel loci (gene symbols in red in **Fig. 1**).

We assessed the fidelity of the association signals by comparing the effect estimates between the two independent hospitals (internal replication) and with two external Chinese datasets (external replication) (**Materials and Methods**). The effect estimates demonstrated high consistency between the two hospitals, where 18/19 association signals (94.8%) or 7/8 loci (87.5%) exhibited the same beta direction and surpassed Bonferroni correction significance levels for TBA in each hospital. Similarly, for ICP, all 16 association signals included in the 5 loci exhibited the same effect direction and significant *P* value after Bonferroni-correction (**Fig. S4, Table S2-S3**). The only *GREB1* gene locus (lead SNP rs10929754-T) associated with TBA did not pass the Bonferroni correction significance level in the Baoan cohort. However, it consistently showed the same effect direction and achieved nominal significance ($P$=0.0233). The effect estimates also demonstrated high consistency between our study and the two Chinese external datasets, where all signals and loci were replicated in one of the two external datasets **(Fig. S5, Table S2, Table S4)**. The

LocusZoom regional association maps for the eight loci associated with TBA and the five loci associated with ICP are presented in **Fig. S6-S7.** The above evaluations suggest high fidelity of the associations identified in our study.

**The East Asian-specific at 14q24.1 locus contributes to a 15.23-fold higher risk of ICP**

Seven out of the eight loci associated with TBA and three of the five loci associated with ICP were not previously documented in the GWAS catalog[24] and Phenoscanner[25], particularly in the current unique ICP GWAS among the European population[16]. Genes contained within the genetic loci associated with TBA levels and ICP are enriched in bile acid and bile salt transport, bile acid signaling and biosynthetic processes, as well as lipid metabolism (**Table S5**). Specifically, the *ABCB11*, *ABCG5* and *SLC10A1* loci associated with ICP are uniquely expressed in liver tissue[26] and hepatocyte cells[27], reflecting the central role of the liver organ in the development of ICP. Notably, we observed substantial discrepancies in genetic discoveries between our study and the European ICP GWAS study (**Fig. S8, Table S2 and Table S5**). Particularly, variants within the 14q24.1 locus, encompassing the *ERH*, *SLC39A9* and *SLC10A1* genes, are absent outside East Asian populations. This includes the lead SNPs rs137983251-G associated with TBA, the lead SNP rs147525203-C associated with ICP and the missense variants rs2296651-A (*p.Ser267Phe*) present in *SLC10A1* which is strongly associated with TBA and ICP risk (*MAF*=0.04, *OR*=15.23, [95%*CI*]:

15.23 [15.10, 15.36], $P = 9.23\times10^{-375}$). While rs137983251-G and rs147525203-C are in complete linkage disequilibrium (LD) with each other ($R^2$=1), rs2296651-A are in less strong LD with rs137983251-G and rs147525203-C ($R^2$=0.21)[28].

## The ICP risk variants at the 14q24.1 locus decrease HBV infection

Notably, *SLC10A1*, also known as *NTCP*, is recognized as a sodium/bile acid cotransporter and has been implicated as the cellular receptor for HBV infection[29]. Despite this, none of the GWAS studies cataloged in the GWAS catalog, Phenoscanner or PubMed has linked it to TBA or ICP[24,25], likely due to its population specificity. To elucidate the function of the East Asian-specific 14q24.1 ICP risk alleles, we scrutinized 246 phenotypes in the Biobank of Japan (BBJ)[30] and over 100 pregnancy phenotypes in our dataset (see Materials and Methods for URL). We found that lead SNPs of TBA and ICP (rs137983251 and rs147525023) and the rs2296651-A (*p.Ser267Phe*) variant were significantly associated with the traits related to Hepatitis B virus (HBV) infection. In our GWAS of six HBV traits, the 14q24.1 locus stands out as the second most significantly associated locus with HBV infection, exhibiting strong associations with Hepatitis B surface antigen (HBsAg), Hepatitis B e antibody (HBeAb) and Hepatitis B core antibodies (HBcAb) ($P<5\times10^{-8}$) (**Fig. S9**). Stacked LocusZoom plots further confirmed the shared genetic effect between TBA/ICP, and HBV/HBsAg/HBeAb/HBcAb in the 14q24.1 locus **(Fig. S10)**. Other traits affected by

the 14q24.1 locus include Alanine transaminase level during pregnancy (**Materials and Methods** for URL).

To fine map the potential causal variants contributing to both TBA/ICP and HBV infection, we conducted a colocalization analysis of TBA & ICP and six HBV-related traits on a 1Mbp region (chr14:69-70Mbp) consisting of 1,909 genetic variants (MAF>0.01). We found that TBA and ICP were both colocalized with HBV with at least one variant overlapping in their 95% credible set (PP.H4>0.75 & H4/H3>3) (**Fig. 2A-B, Fig. S11, Table S7)**. Evidence for the strongest colocalization of GWAS signals between TBA and HBV was found for *SLC39A9* (rs138089855, intron variant), and that between ICP and HBV was found for *SCL10A1* (rs2296651, *p.Ser267Phe*, missense variant). The rs138089855-C mutation and rs2296651-A mutation significantly affect immune evasion of HBV infection (*OR* ranging from 0.57 to 0.82, $P < 4.63\times10^{-10}$), reducing risks of HBsAg (*OR* ranges from 0.60 to 0.83, $P < 3.72\times10^{-5}$), HBeAb (OR ranges from 0.65 to 0.82, $P < 3.24\times10^{-7}$) and HBcAb (*OR* ranges from 0.63 to 0.79, $P < 2.24\times10^{-11}$) (**Fig. 2C-D**).

## The *p.Ser267Phe* mutation of the *SLC10A1* variant became prevalent within the last 3,000 years in the areas of East and Southeast Asia

The Geography of Genetic Variants Browser (https://popgen.uchicago.edu/ggv/) illustrates that the lead SNP of TBA (rs137983251-G) and ICP (rs147525203-G), as

well as the rs138089855-C and missense locus rs2296651-A identified in colocalization analysis, are only present in East and Southeast Asia **(Fig. S12)**. To delve further into the historical evolutionary history of ICP genetic loci, we assessed the spatio-temporal frequency differences of rs2296651-A globally and since the Holocene using the Allen Ancient Data Resource (AADR)[31]. In the Holocene era (10,000 BP before now), only one sample with the *p.Ser267Phe* mutation exists in Morocco (**Fig. 3A, Table S8**). In the Neolithic era (10,000~3,000BP), several *p.Ser267Phe* variations appear sporadically, with single samples distributed in Italy, Russia, Spain and California, USA (**Fig. 3B, Table S9**). Notably, until the Historic (<3,000BP), a large number of mutations occurred at this locus, especially within the range of 1,000 to 2,000BP (**Fig. 3C, Table S10**). Currently, the ICP risk mutations of this locus are only distributed in East Asia and Southeast Asia (**Fig. 3D, Table S11)**.

## Natural selection signals in the 14q24.1 locus

The distinctive prevalence of the 14q24.1 locus in different geographical regions may indicate natural selection. To further understand the genetic diversity in the 14q24.1 locus, we employed 1KGP high-depth sequencing data for selection analysis. We calculated several metrics based on the site-frequency spectrum in this region for three representative populations: Southern Han Chinese (CHS), Northern Europeans from Utah (CEU), and Yoruba in Ibadan, Nigeria (YRI). We observed a decrease in nucleotide diversity in chr14:69.0-69.2Mbp and around chr14:69.8-69.9 regions in

CHS compared to CEU and YRI. Tajima's *D* and Fay and Wu's *F* statistics also suggested an excess of high-frequency derived SNPs in these two regions **(Fig. S13)**.

Furthermore, we conducted a selection analysis based on population differentiation and linkage disequilibrium. The chr14:69.0-69.2Mbp region shows the highest locus-specific branch length (LBSL) statistics for CHS compared with the CEU using AFR as the reference panel, indicating a signal of positive selection in East Asia. The results of iHS and the negative XP-EHH values also support that the selection occurred in the CHS population **(Fig. S14)**.

Regarding the high linkage between lead SNPs of TBA and ICP, we explored whether there are differences in haplotype frequencies among different races. We extracted SNPs at high linkage disequilibrium with SNP rs147525203, rs137983251, rs138089855, or rs2296651, resulting in 16 SNPs which were phased into haplotypes. Having examined differences between different races, we found significant differences in haplotypes in East Asians compared to other populations. Of the top 10 haplotypes, 4 types (H4, H5, H6, and H9) were almost exclusively found in 10.8% of Asian individuals **(Fig. S15, Table S12)**. The origin of these haplotypes was not linked to known archaic ancestry. However, we cannot be ruled out that these haplotypes may have originated from unknown East and Southeast Asian archaic hominins, which are currently underrepresented in human genetic investigations[32,33,34].

## Discussion

Pregnancy constitutes a critical period for human reproduction, and ICP emerges as a high-risk liver disorder during this phase. Characterized by maternal pruritus and elevated serum bile acid concentrations[2], ICP is associated with an elevated risk of severed pregnancy outcomes, including intrauterine fetal growth restriction, preterm birth, fetal asphyxia, still birth and lower birth weight[3,4]. Standard clinical practice typically designates pregnancies diagnosed with ICP as high-risk, administering UDCA post-diagnosis, and often inducing labor toward full term. However, recent evidence from an RCT challenges the efficacy of UDCA in reducing adverse perinatal outcomes in women with ICP[1]. The high risk and lack of effective treatment of ICP underscores the need for a more comprehensive understanding of the factors leading to ICP. In the present study, we conducted the first-ever GWAS meta-analysis of TBA and ICP traits in East Asians, encompassing 101,023 participants, including 4,703 cases and 96,320 controls, with signals verified in two independent Chinese cohorts. We identified eight and three genome-wide significant associated loci for TBA and ICP, respectively, including seven and three novel association loci.

Of particular interest is the revelation that the most significant 14q24.1 locus contributes to an average odds ratio of 15.23 for ICP in Chinese pregnancies. This is much higher compared to the 1.70 odds ratio for ICP associated with the known

372 *ABCG5* and *ABCB11* loci reported in the European population. The high-risk allele in
373 the 14q24.1 locus, comprising the *SLC39A9* (lead SNP rs137983251), *ERH* (lead SNP
374 rs147525023), and *SLC10A1* (missense variant rs2296651-A, *p.Ser267Phe*) genes, is
375 uniquely present in the East and Southeast Asia and conspicuously absent in other
376 global populations. Comprehensive phenome-wide association scans based on BBJ
377 data and our own dataset revealed a significant association between the ICP-high-risk
378 allele and a protective effect against HBV infection, reflected by a decreased risk of
379 HBV infection, HBsAg, HBcAb, and HBeAb serological positivity. Spatiotemporal
380 and natural selection analyses further suggest that the prevalence of the ICP risk allele
381 dates back within the last 3000 years, putatively due to selection over resistance of
382 HBV infection.
383
384 Currently, Hepatitis B virus (HBV) infection remains a significant public health
385 concern, affecting over 296 million people worldwide chronically infected[35]. However,
386 little is known about the origin and historical spread of HBV. In 1964, American
387 physician and geneticist Blumberg et al discovered a new antigen in the serum of an
388 Australian Aborigine named Australian antigen (AuAg)[36], and subsequently identified
389 it as the surface antigen (HBsAg), marking the first specific indicator of hepatitis B
390 virus (HBV)[37]. Recent evidence of HBV's historical existence extends to mummies in
391 Korea and Italy from 400 years ago[38,39]. Moreover, in 2018, the discovery of HBV
392 genomes from ancient DNA derived from human skeletal remains fossilized in the

393 Neolithic period around 7,000 years ago was documented[40]. Our study suggests that a
394 potential HBV epidemic occurred 3,000 years ago in East and Southeast Asia, leading
395 to a balancing or positive selection of the ICP risk allele. This selection resulted in a
396 higher prevalence of ICP observed currently in these regions, shaping a distinct
397 genetic etiology of ICP between East Asians and Europeans.
398
399 What are the medical implications arising from these findings? The first takeaway is
400 that the genetic etiology of ICP differ substantially between Asian and European
401 populations. Clinical practice should take into account the ethnic ancestry of patients
402 for informed clinical decisions, with particular attention to the presence of the
403 14q24.1 ICP risk allele. Secondly, the prevalence of ICP in East and Southeast Asia is
404 an outcome of evolutionary processes. Pregnant individuals carrying the risk alleles
405 exhibited enhanced biological resistance to HBV infection and are likely descendants
406 of survivors from a historical HBV epidemic. Mothers carrying the risk allele need
407 not bear any stigma regarding their infants. Thirdly, evidence on the clinical risk of
408 ICP primarily stems from large-scale cohort-based observational studies[5], susceptible
409 to confounding by unknown factors. Given the cumulative substantial impact of TBA
410 and ICP risk alleles identified in this study, more refined causal inference regarding
411 the relationship between TBA and ICP and the short-term birth outcomes and
412 long-term offspring health, can be achieved using methods such as Mendelian
413 randomization applied to East-Asian birth cohorts[41,42].

Building upon these discoveries, more precise clinical decisions can be tailored for pregnancies, preventing unnecessary over-treatment. Additionally, novel therapeutics can be developed specifically for high-risk patients based on their genetic origins, compensating for the limited effectiveness of UDCA in reducing adverse outcomes for ICP patients. For example, as HBV vaccinations are mandatorily implemented for newborns in countries like China, it may be worthwhile to explore whether increasing NTCP activity may benefit the ICP patients in clinics while not increasing their risk for HBV infection.

## Materials and Methods

### Cohort description

The study encompassed 70,608 pregnancies recruited from Longgang District Maternity & Child Healthcare Hospital of Shenzhen City (Longgang cohort) and 50,948 pregnancies from Shenzhen Baoan Women's and Children's Hospital (Baoan cohort). These participants engaged in a routine pregnancy screening program in Shenzhen between 2017 and 2022. Pregnancies involving multiple gestations and those lacking TBA level measurements were excluded. Finally, our study comprised 101,023 pregnant women who underwent at least one TBA level assessment during gestational weeks 13 to 42.

To conduct GWAS of TBA and ICP, we integrated NIPT data with clinical phenotypic data (see **Supplementary Methods**). All participants provided written informed consent. The study received approved from the Medical Ethics Committee of the School of Public Health (Shenzhen), Sun Yat-sen University (No. 2022-021), Longgang District Maternity and Child Healthcare Hospital of Shenzhen City (LGFYYXLLL-2022-024), and Shenzhen Baoan Women's and Children's Hospital (LLSC-2021-04-01-10-KS). Data collection was also approved by the Human Genetic Resources Administration of China (HGRAC) (Baoan cohort: [2023]CJ1415, Longgang cohort:[2023]CJ1455).

## Phenotype definition

The TBA level was defined as the peak level observed for each individual during gestational weeks 13 to 42 after removing outliers (N = 94,360). ICP cases were identified as pregnancies with TBA concentrations ≥ 10μmol/L during the same gestational period. In total, 4,703 cases and 96,320 controls for ICP were included in the study.

We also obtained the hepatitis B antigen and antibody measurements from data of the pregnant screening program. The phenotype of hepatitis B virus (HBV) infection was defined based on hepatitis B antigen and antibody. Detailed information can be found in **Supplementary Methods**.

## Genome-wide association analysis

We conducted the GWAS analysis using PLINK 2.0[43]. Covariates such as gestational week, maternal age, and the top ten principal components accounting for population stratification were included in the analysis. The quantitative phenotype (TBA) underwent a rank-based transformation to achieve a normal distribution.

To integrate results from two cohorts, a meta-analysis was performed using fixed-effect models with inverse-variance weighting, employing METAL software

(version 2011-03-25)[44]. Variants with a minor allele frequency (MAF) less than 0.01 were excluded. Conditional and joint analysis (GCTA-COJO) was performed using GCTA software[45,46] to identify independent genome-wide significant signals. A stepwise model selection procedure (--cojo-slct) with a collinearity threshold of 0.2 was employed to choose independently associated SNPs, with the reference panel from the Born in Guangzhou Birth cohort (BIGCS)[42] for the LD structure of the variants. Subsequently, we delineated the range of 500kb upstream and downstream blocks of a significant signal into separate locus, considering the SNP with the lowest *P* value as the lead SNP.

**Statistic power calculation**

All statistical analyses were executed in R (version 4.2.1). Post hoc power calculations were performed to access the spectrum of effect size ($\beta$)/odds ratio (*OR*) and allele frequencies at which associations could be detected at genome-wide significance ($P < 5\times10^{-8}$) within a specified number of samples/cases and controls in the meta-analysis. The calculation was performed using the R package "genpwr"[47] (version 1.0.4) using a linear or logistic model under a genetic additive mode.

**Replication and comparison**

For internal replication, we applied stringent criteria to define a locus as replicated,

requiring the same beta direction and *P* values reaching the Bonferroni correction threshold in both the Baoan and Longgang cohorts.

In the external replication, we replicated genome-wide significant SNPs using two independent study cohorts (Baoan NIPT PLUS cohort and BIGCS cohort). SNPs meeting the following criteria were regarded as replicated: 1) they exhibited a consistent direction of effect for lead SNPs between the discovery and the replication cohorts and 2) they reached Bonferroni-corrected *P* values or passed a two-sided two-sample t-test. Further details on external replication can be found in **Supplementary methods**.

Furthermore, we compared the genetic influence of SNPs with a previously published European cohort[16]. This comparison involved assessing the direction of genetic effects and Bonferroni-corrected *P* values of lead SNPs between our meta-analysis results and the European cohort.

## PheWAS and Colocalization analysis

A PheWAS was conducted using 246 phenotypes from BBJ (https://pheweb.jp/) and 131 pregnancy phenotypes from the MONN PheWeb established by our research group (http://47.112.105.165/). Regions showing evidence of colocalization between the GWAS and GWAS signals were identified utilizing pre-defined thresholds: PP4

(posterior probability that there exists a single causal variant common to both traits) ≥ 0.75 and PP4/PP3 ≥ 3, employing the R package "coloc" (version 5.1.0.1)[48]. The prior probabilities for SNP association with either of the two traits or with both traits were set to $1 \times 10^{-4}$ and $1 \times 10^{-5}$, the default values, respectively.

## Geographical frequency distribution analysis

To examine the frequency distribution of identified SNPs across various ancient age groups, we employed ancient DNA data from the compiled 1240K dataset sourced from the Allen Ancient DNA Resource (version 54.1)[31]. The geographical map, supplemented with pie charts, was created using R packages "maps" and "mapplots" to illustrate the geographical frequency distribution.

## Modern DNA-based selection test

We performed three types of selection tests for modern DNA, encompassing analyses based on site-frequency spectrum (SFS), population differentiation, and haplotype selection to explore signals of positive or balancing selection within East Asia populations. SNPs with minor allele frequencies (MAF) < 0.01 in all five super-populations (AFR, CEU, SAS, EAS and AMR) and genetic variants lacking ancestral information were excluded. Ultimately, 18,221,282 SNPs were used for the natural selection analysis (**Supplementary Methods** for details).

## Haplotype analysis

The data used for haplotype analysis corresponds to that employed in the aforementioned natural selection analysis. SNPs with an MAF < 0.05 in all 5 super-populations and genetic variants lacking ancestral information were excluded. We extracted SNPs exhibiting high linkage disequilibrium ($R^2 > 0.20$) with rs137983251, rs138089855, rs147525203, or rs2296651 variants, resulting in 16 SNPs that were used to construct 31 haplotype types. To further streamline haplotype diversity, the rs79422091 variation was excluded to merge similar haplotypes. Consequently, 16 SNPs were selected to form 25 distinct haplotypes.

## Abbreviations

AADR, Allen Ancient DNA Resource; BRN, Brunei; CDX, Chinese Dai in Xishuangbanna; CHB, Han Chinese in Beijing; CHS, Southern Han Chinese; CI, confidence interval; COJO, conditional and joint analysis; CQ, Chongqing (China); DE, Germany; DK, Denmark; ES, Spain; FJ, Fujian (China); FSM, Federated States of Micronesia; GWAS, genome-wide association study; GX, Guangxi (China); HA, Henan (China); HBcAb, hepatitis B core antibody; HBeAb, hepatitis B e antibody; HBeAg, hepatitis B e antigen; HBsAb, hepatitis B surface antibody; HBsAg, hepatitis B surface antigen; HBV: hepatitis B virus; HCV, hepatitis C virus; HR, hazard ratio; ICP, intrahepatic cholestasis of pregnancy; ID, Indonesia; iHS, integrated Haplotype

546 Score; IS, Iceland; IT, Italy; JPN, Japan; KH, Cambodia; LB, Lebanon; LD, linkage
547 disequilibrium; MA, Morocco; MAF, minor allele frequency; MY, Malaysia; NIPT,
548 non-invasive prenatal testing; OR, odds ratio; PCs, principal components; PH,
549 Philippines; PNG, Papua New Guinea; PP.H, posterior probability for hypothesis; RU,
550 Russia; SFS, site-frequency spectrum; SNP, single nucleotide polymorphism; TBA,
551 total serum bile acids; TWN, Taiwan (China); USA, United States of America; VAN,
552 Vanuatu; VEP, Ensembl Variant Effect Predictor; VN, Vietnam; XJ, Xinjiang (China);
553 XP-EHH, Cross Population Extended Haplotype Homozygosity.
554

555 **Reference**

556 1. Chappell, L. C. *et al.* Ursodeoxycholic acid versus placebo in women with
557 intrahepatic cholestasis of pregnancy (PITCHES): a randomised controlled trial.
558 *Lancet* **394**, 849–860 (2019).
559 2. Puljic, A. *et al.* The risk of infant and fetal death by each additional week of
560 expectant management in intrahepatic cholestasis of pregnancy by gestational
561 age. *Am. J. Obstet. Gynecol.* **212**, 667.e1-667.e5 (2015).
562 3. Glantz, A., Marschall, H. U. & Mattsson, L. Å. Intrahepatic cholestasis of
563 pregnancy: Relationships between bile acid levels and fetal complication rates.
564 *Hepatology* **40**, 467–474 (2004).
565 4. Geenes, V. *et al.* Association of severe intrahepatic cholestasis of pregnancy
566 with adverse pregnancy outcomes: A prospective population-based case-control
567 study. *Hepatology* **59**, 1482–1491 (2014).
568 5. Ovadia, C. *et al.* Association of adverse perinatal outcomes of intrahepatic
569 cholestasis of pregnancy with biochemical markers: results of aggregate and
570 individual patient data meta-analyses. *Lancet* **393**, 899–909 (2019).
571 6. Alsulyman, O. M., Ouzounian, J. G., Ames-Castro, M. & Goodwin, T. M.
572 Intrahepatic cholestasis of pregnancy: perinatal outcome associated with
573 expectant management. *Am. J. Obstet. Gynecol.* **175**, 957–960 (1996).
574 7. Sentilhes, L., Verspyck, E., Pia, P. & Marpeau, L. Fetal death in a patient with
575 intrahepatic cholestasis of pregnancy. *Obstet. Gynecol.* **107**, 458–460 (2006).

8. Ovadia, C. *et al.* Ursodeoxycholic acid in intrahepatic cholestasis of pregnancy: a systematic review and individual participant data meta-analysis. *Lancet Gastroenterol. Hepatol.* **6**, 547–558 (2021).

9. Arrese, M., Macias, R. I. R., Briz, O., Perez, M. J. & Marin, J. J. G. Molecular pathogenesis of intrahepatic cholestasis of pregnancy. *Expert Rev. Mol. Med.* **10**, 1–18 (2008).

10. Lammert, F., Marschall, H. U., Glantz, A. & Matern, S. Intrahepatic cholestasis of pregnancy: molecular pathogenesis, diagnosis and management. *J. Hepatol.* **33**, 1012–1021 (2000).

11. Jacquemin, E., Cresteil, D., Manouvrier, S., Boute, O. & Hadchouel, M. Heterozygous non-sense mutation of the MDR3 gene in familial intrahepatic cholestasis of pregnancy. *Lancet* **353**, 210–211 (1999).

12. Schneider, G. *et al.* Linkage between a new splicing site mutation in the MDR3 alias ABCB4 gene and intrahepatic cholestasis of pregnancy. *Hepatology* **45**, 150–158 (2007).

13. Dixon, P. H. *et al.* A comprehensive analysis of common genetic variation around six candidate loci for intrahepatic cholestasis of pregnancy. *American Journal of Gastroenterology* vol. 109 76–84 at https://doi.org/10.1038/ajg.2013.406 (2014).

14. Wasmuth, H. E. *et al.* Intrahepatic cholestasis of pregnancy: The severe form is associated with common variants of the hepatobiliary phospholipid transporter ABCB4 gene. *Gut* **56**, 265–270 (2007).

15. Chen, R. *et al.* Intrahepatic cholestasis of pregnancy as a clinical manifestation of sodium-taurocholate cotransporting polypeptide deficiency. *Tohoku J. Exp. Med.* **248**, 57–61 (2019).

16. Dixon, P. H. *et al.* GWAS meta-analysis of intrahepatic cholestasis of pregnancy implicates multiple hepatic genes and regulatory elements. *Nat. Commun.* **13**, (2022).

17. Devin D Smith, K. M. R. Intrahepatic Cholestasis of Pregnancy. *Clin. Obstet. Gynecol.* **63**, 134–151 (2020).

18. Lee, R. H., Goodwin, T. M., Greenspoon, J. & Incerpi, M. The prevalence of intrahepatic cholestasis of pregnancy in a primarily Latina Los Angeles population. *J. Perinatol. Off. J. Calif. Perinat. Assoc.* **26**, 527–532 (2006).

19. Reyes, H. Sex hormones and bile acids in intrahepatic cholestasis of pregnancy. *Hepatology* **47**, 376–379 (2008).

20. Arrese, M. & Reyes, H. Intrahepatic cholestasis of pregnancy: A past and present riddle. *Ann. Hepatol.* **5**, 202–205 (2006).

21. Gao, X. X. *et al.* Prevalence and risk factors of intrahepatic cholestasis of pregnancy in a Chinese population. *Sci. Rep.* **10**, 16307 (2020).

615 22. Liu, S. *et al.* Genomic Analyses from Non-invasive Prenatal Testing Reveal
616 Genetic Associations, Patterns of Viral Infections, and Chinese Population
617 History. *Cell* **175**, 347-359.e14 (2018).
618 23. Liu, S. *et al.* Utilizing Non-Invasive Prenatal Test Sequencing Data Resource
619 for Human Genetic Investigation. *bioRxiv* (2023).
620 24. Sollis, E. *et al.* The NHGRI-EBI GWAS Catalog: knowledgebase and
621 deposition resource. *Nucleic Acids Res.* **51**, D977–D985 (2023).
622 25. Kamat, M. A. *et al.* PhenoScanner V2: An expanded tool for searching human
623 genotype-phenotype associations. *Bioinformatics* **35**, 4851–4853 (2019).
624 26. Thul, P. J. & Lindskog, C. The human protein atlas: A spatial map of the human
625 proteome. *Protein Sci.* **27**, 233–244 (2018).
626 27. Consortium, Gte. The GTEx Consortium atlas of genetic regulatory effects
627 across human tissues. *Physiol. Behav.* **369**, 1318–1330 (2020).
628 28. Machiela, M. J. & Chanock, S. J. LDlink: A web-based application for
629 exploring population-specific haplotype structure and linking correlated alleles
630 of possible functional variants. *Bioinformatics* **31**, 3555–3557 (2015).
631 29. Park, J. H. *et al.* Structural insights into the HBV receptor and bile acid
632 transporter NTCP. *Nature* **606**, 1027–1031 (2022).
633 30. Sakaue, S. *et al.* A cross-population atlas of genetic associations for 220 human
634 phenotypes. *Nat. Genet.* **53**, 1415–1424 (2021).

635 31. Mallick, S. *et al.* The Allen Ancient DNA Resource (AADR): A curated
636 compendium of ancient human genomes. *bioRxiv Prepr. Serv. Biol.* (2023)
637 doi:10.1101/2023.04.06.535797.
638 32. Yang, M. A. *et al.* Ancient DNA indicates human population shifts and
639 admixture in northern and southern China. *Science (80-. ).* **369**, 282–288
640 (2020).
641 33. Wang, C. C. *et al.* Genomic insights into the formation of human populations in
642 East Asia. *Nature* **591**, 413–419 (2021).
643 34. Wang, T. *et al.* Human population history at the crossroads of East and
644 Southeast Asia since 11,000 years ago. *Cell* **184**, 3829-3841.e21 (2021).
645 35. Kramvis, A. *et al.* A roadmap for serum biomarkers for hepatitis B virus:
646 current status and future outlook. *Nat. Rev. Gastroenterol. Hepatol.* **19**,
647 727–745 (2022).
648 36. Blumberg, B. S., Alter, H. J. & Visnich, S. A "New" Antigen in Leukemia Sera.
649 *JAMA* **191**, 541–6 (1965).
650 37. Prince, A. M. Relation of Australia and SH antigens. *Lancet* **292**, 462–463
651 (1968).
652 38. Kahila Bar-Gal, G. *et al.* Tracing hepatitis B virus to the 16th century in a
653 Korean mummy. *Hepatology* **56**, 1671–80 (2012).
654 39. Patterson Ross, Z. *et al.* The paradox of HBV evolution as revealed from a 16th
655 century mummy. *PLoS Pathog.* **14**, e1006750 (2018).

656 40. Krause-Kyora, B. *et al.* Neolithic and medieval virus genomes reveal complex
657 evolution of hepatitis B. *Elife* **7**, e36666 (2018).
658 41. Qiu, X. *et al.* The Born in Guangzhou Cohort Study (BIGCS). *Eur. J.*
659 *Epidemiol.* **32**, 337–346 (2017).
660 42. Huang, S. *et al. Whole genome sequencing and analysis of 4,053 individuals in*
661 *trios and mother-infant duos from the Born in Guangzhou Cohort Study*.
662 *Research Square* (2022).
663 43. Purcell, S. *et al.* PLINK: A tool set for whole-genome association and
664 population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
665 44. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: Fast and efficient meta-analysis
666 of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
667 45. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA : A Tool for
668 Genome-wide Complex Trait Analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
669 46. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary
670 statistics identifies additional variants influencing complex traits. *Nat. Genet.*
671 **44**, 369–S3 (2012).
672 47. Moore, C. M., Jacobso, S. A., Fingerlin, T. E. & Health, N. J. Power and
673 Sample Size Calculations for Genetic Association Studies in the Presence of
674 Genetic Model Mis-Specification. **84**, 256–271 (2019).

675 48. Giambartolomei, C. *et al.* Bayesian Test for Colocalisation between Pairs of
676 Genetic Association Studies Using Summary Statistics. *PLoS Genet.* **10**,
677 e1004383 (2014).

678

679

## Figure legends

**Abstract Figure: Genetic basis and evolutionary history of intrahepatic cholestasis of pregnancy in East Asia.**

TBA: Total bile Acid. ICP: Pregnancy intrahepatic cholestasis. Refer to the main text for the illustration.

**Fig. 1. Meta-analysis of Genome-wide association study for total bile acid (TBA) and intrahepatic cholestasis of pregnancy (ICP).**

(A) Manhattan plot for TBA GWAS meta-analysis, encompassing 94,360 samples. (B) Manhattan plot for ICP GWAS meta-analysis, involving 4,703 cases and 96,320 controls. Chromosomes are ordered on the x-axis and the $-\log_{10}(P)$ values for the association tests are shown on the y-axis. Horizontal dashed lines delineate the genome-wide significance threshold ($P = 5\times10^{-8}$, in grey). Eight and five independent loci achieved genome-wide significance ($P < 5\times10^{-8}$) with TBA and ICP, respectively. Labels in black denote known loci and labels in red highlight novel loci.

**Fig. 2. GWAS-GWAS colocalization of TBA & ICP with HBV and forest plots for the shared SNPs with HBV-related traits.**

(A) The purple diamond in the plot indicates the colocalized SNP, rs138089855,

699 shared between TBA and HBV. (B) Another SNP, rs2296651, is identified as the
700 colocalized SNP between ICP and HBV. The color coding on the plot signifies
701 the $R^2$ measure of linkage disequilibrium of the colocalized SNP. The forest plot
702 illustrates the genetic effect of (C) rs138089855-C allele and (D) rs2296651-A allele
703 at 14q24.1 on the meta-analysis for HBV and five HBV antigen & and antibody status.
704 The error bars represent the 95% confidence interval of the odds ratio (*OR*).
705

706 **Fig. 3. Temporal and geographical changes of allele frequency of rs2296651.**

707 The geographical frequency distribution of SNP rs2296651 is depicted in (A) the
708 Holocene age (> 10,000BP), (B) the Neolithic age (10,000~3,000BP), (C) the Historic
709 (<3,000BP), and (D) the present-day global populations, utilizing the "1240K" dataset
710 ("Allen Ancient DNA Resource", version 54.1). The ancestral allele (reference allele)
711 rs2296651-G is represented in orange and the derived allele (alternative allele A, the
712 ICP risk allele) is shown in blue. The color of the pie chart border indicates the
713 detailed archaeological period in the plot in (C) (i.e., red: 0~1,000BP, black:
714 1,000~2,000BP, green: 2,000~3,000BP). Plots (C) and (D) exclusively display
715 populations where rs2296651-A is present. The frequency or quantity of the A allele
716 in different populations is shown in parentheses. Detailed allele frequencies are
717 provided in **Supplementary Table 8-11**.

Phenotype

Genotype

## Cohort

13th to 42nd gestational weeks

**TBA**
(TBA<10μmol/L)

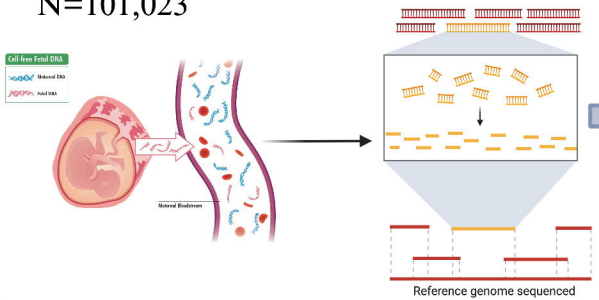
- **Baoan** (N=38,663)
- **Longgang** (N=55,697)

**ICP**
(Diagnosed by if TBA≥10μmol/L)

- **Baoan** (1,931/39,593)
- **Longgang** (2,772/56,727)

## Genome sequencing from NIPT

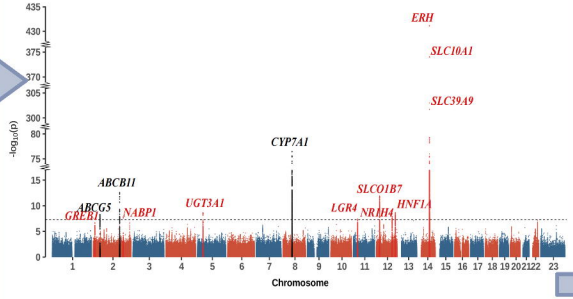
N=101,023




## Genetic determinants

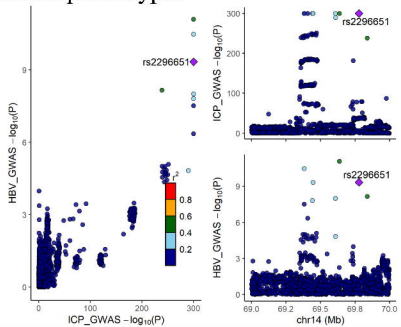
### GWAS & meta-GWAS

7,300,273 variants (MAF>0.01)
8 loci identified with TBA and 5 loci with ICP




### Phenome-wide colocalization

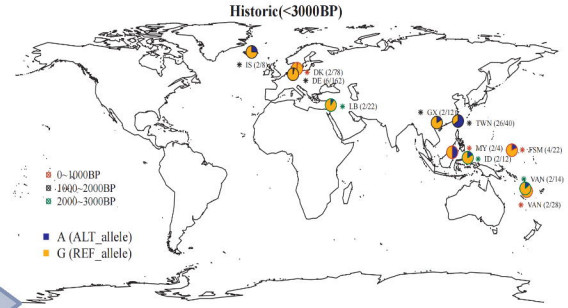
The 14q24.1 locus was associated with HBV and its related phenotypes




## Evolution

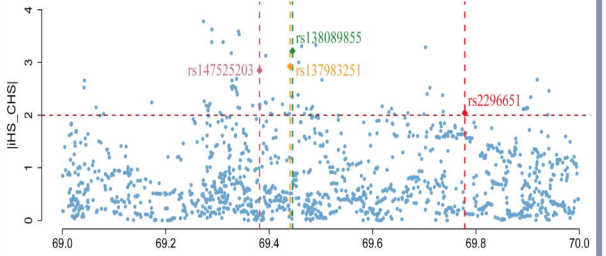
### Spatio-temporal frequency change

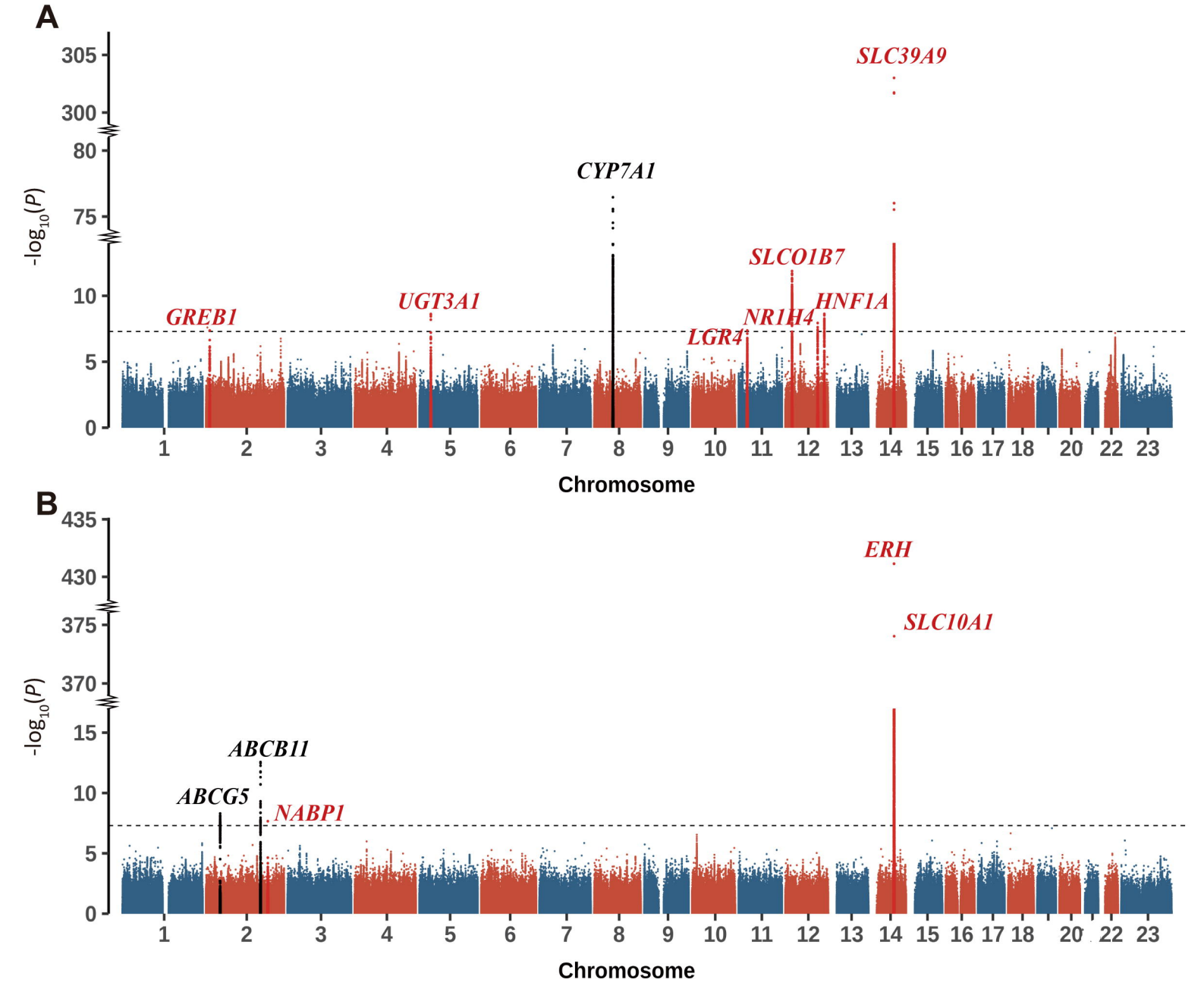
ICP risk allele frequencies throughout world and since Holocene

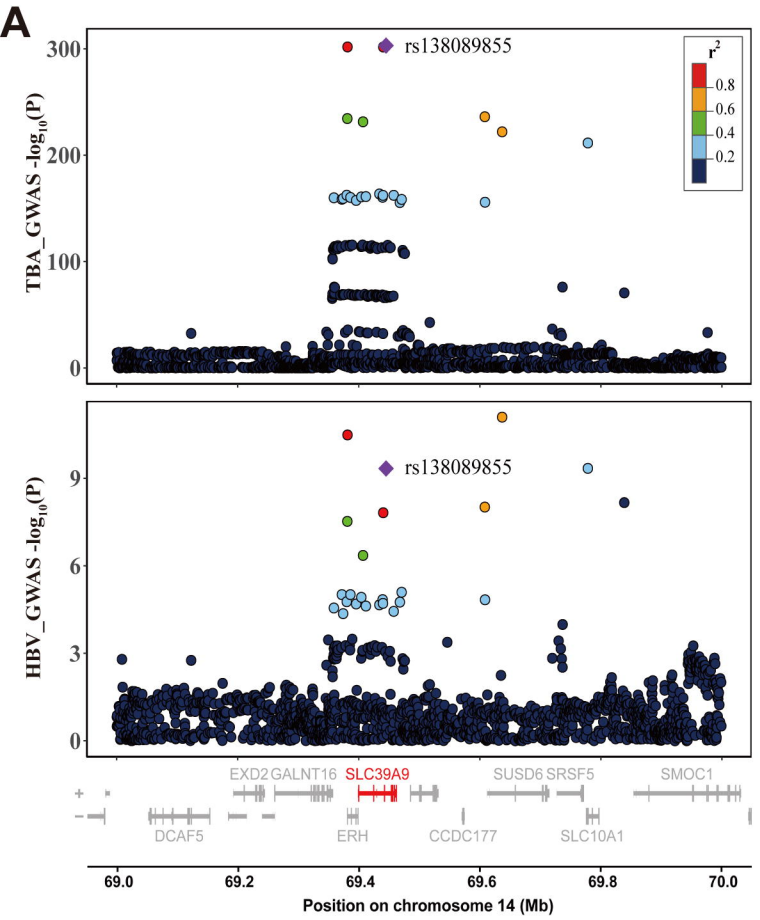


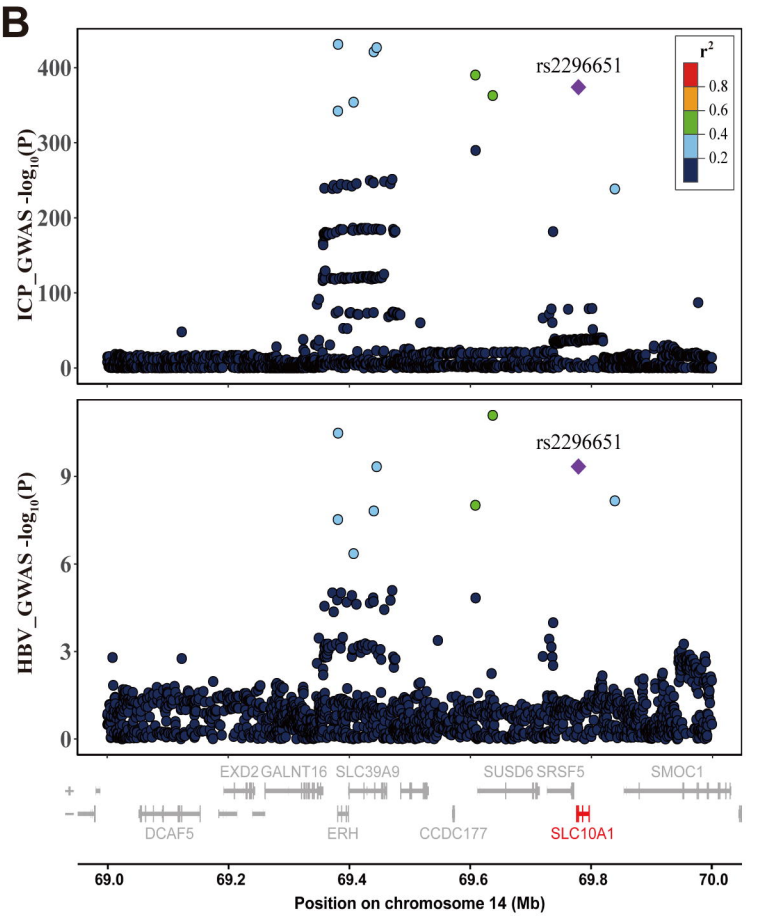

### Natural selection

Natural selection analysis of ICP risk alleles in 14q24.1



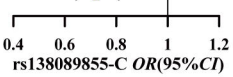

A
305
300
80
75
10
5
0
-log10(P)
SLC39A9
CYP7A1
SLCO1B7
HNF1A
NR1H4
LGR4
UGT3A1
GREB1
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
20
22
23
Chromosome
B
435
430
375
370
15
10
5
0
-log10(P)
ERH
SLC10A1
ABCB11
ABCG5
NABP1
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
20
22
23
Chromosome



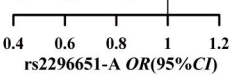

C

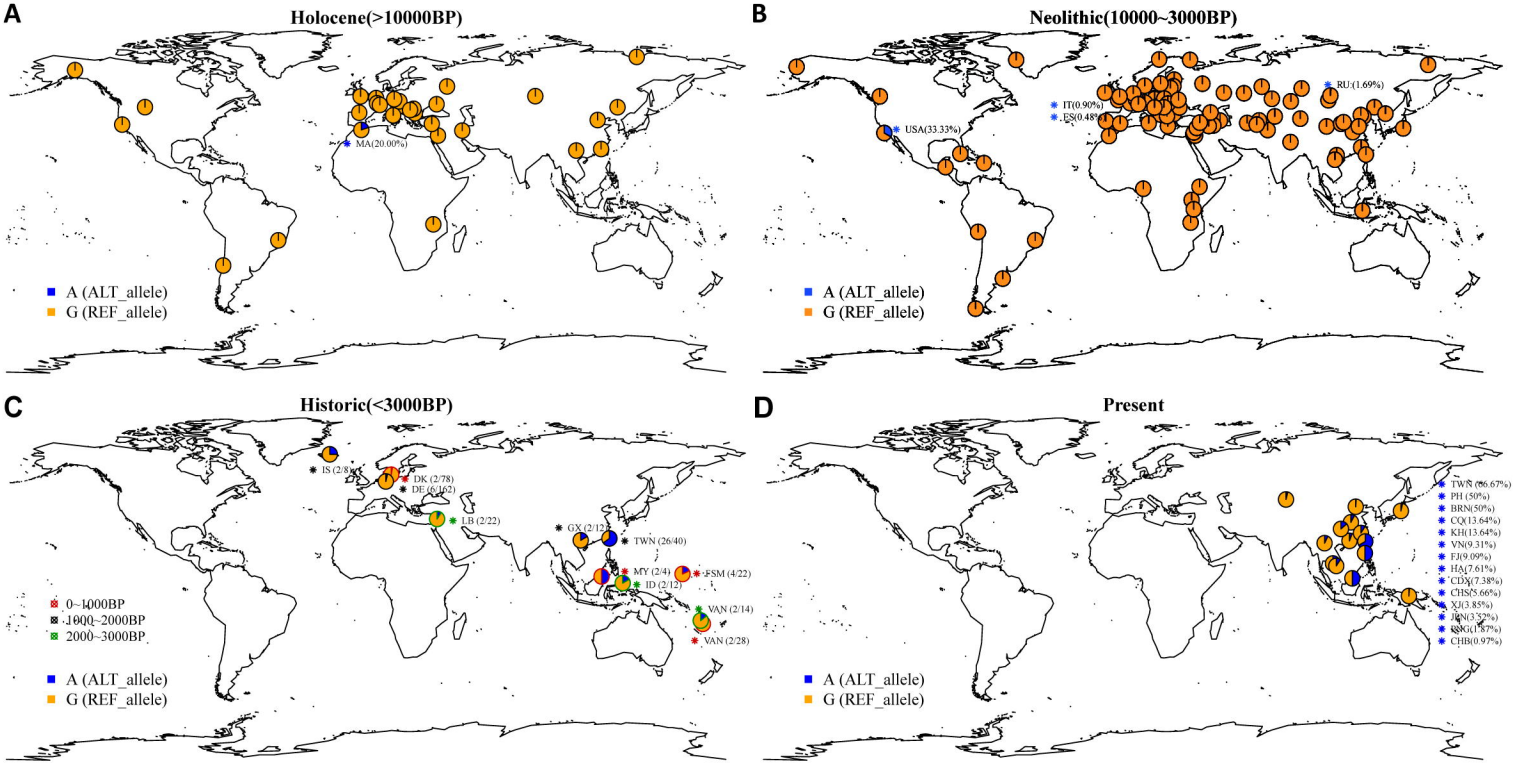
| Pheno | Case | Control | *OR*(95%*CI*) | *P* |
|---|---|---|---|---|
| HBsAb | 63,627 | 30,834 | 1.03(0.97~1.10) | 3.34E-01 |
| HBeAg | 1,817 | 92,644 | 0.80(0.65~1.00) | 4.91E-02 |
| HBeAb | 16,117 | 78,344 | 0.76(0.70~0.82) | 4.34E-11 |
| HBV | 21,770 | 22,662 | 0.75(0.68~0.82) | 4.63E-10 |
| HBcAb | 23,618 | 70,843 | 0.74(0.69~0.79) | 9.10E-17 |
| HBsAg | 7,575 | 86,866 | 0.72(0.64~0.81) | 2.28E-08 |




D

| Pheno | Case | Control | *OR*(95%*CI*) | *P* |
|---|---|---|---|---|
| HBsAb | 63,627 | 30,834 | 0.98(0.89~1.07) | 6.53E-01 |
| HBeAb | 16,117 | 78,344 | 0.73(0.65~0.82) | 3.24E-07 |
| HBsAg | 7,575 | 86,866 | 0.70(0.60~0.83) | 3.72E-05 |
| HBcAb | 23,618 | 70,843 | 0.70(0.63~0.78) | 2.24E-11 |
| HBV | 21,770 | 22,662 | 0.65(0.57~0.75) | 4.57E-10 |
| HBeAg | 1,817 | 92,644 | 0.61(0.44~0.86) | 4.17E-03 |




A
Holocene(>10000BP)
MA(20.00%)
A (ALT_allele)
G (REF_allele)
B
Neolithic(10000~3000BP)
USA(33.33%)
IT(0.90%)
ES(0.48%)
RU(1.69%)
A (ALT_allele)
G (REF_allele)
C
Historic(<3000BP)
IS (2/8)
DK (2/78)
DE (6/162)
LB (2/22)
GX (2/12)
TWN (26/40)
MY (2/4)
ID (2/12)
FSM (4/22)
VAN (2/14)
VAN (2/28)
0~1000BP
1000~2000BP
2000~3000BP
A (ALT_allele)
G (REF_allele)
D
Present
TWS (86.67%)
PH (50%)
BRN(50%)
CQ(13.64%)
KH(13.64%)
VN(9.31%)
FJ(9.09%)
HAJ(7.61%)
CDX(7.38%)
CHS(5.66%)
XJ(3.85%)
JSN(3.52%)
SNG(1.87%)
CHB(0.97%)
A (ALT_allele)
G (REF_allele)