nature communications

Article

https://doi.org/10.1038/s41467-026-73122-z

Genome-wide meta-analysis identifies genetic drivers of bile acid metabolism in intrahepatic cholestasis of pregnancy

Received: 21 August 2025

Accepted: 2 May 2026

Published online: 25 May 2026

Check for updates

Jaakko S. Tyrmi[1,2] ✉, Juha Karjalainen[3], Samvida S. Venkatesh[4], Clara Benoit-Pilven[3], Susanna M. Lemmelä[3], Søren Brunak[5], Bitten Aagaard[6], Mie T. Bruun[7], Christian Erikstrup[8], Henrik Ullum[9], Ole B. Pedersen[10], Karina Banasik[5,11], Erik Sørensen[12], Christina Mikkelsen[12], Michael Schwinn[12], Arni Sturluson[13], Sisse R. Ostrowski[12,14], FinnGen*, Estonian Biobank research team*, DBDS Genomic Consortium*, Johannes Kettunen[2], Henriette Svarre Nielsen[11,14], Mette Nyegaard[9,14], David Westergaard[11], Thorunn Rafnar[13], Patrick Sulem[13], Kari Stefansson[13], Priit Palta[15], Triin Laisk[15], Taru Tukiainen[3,16], Cecilia M. Lindgren[4], Mark Daly[3,17], Aki S. Havulinna[3,18,21] & Hannele M. Laivuori[1,3,19,20,21]

Intrahepatic cholestasis of pregnancy, which affects 0.2–2% of pregnancies, is characterized by pruritus, increased aminotransferase activity and elevated serum bile acids. Previous studies have implicated liver-enriched genes in intrahepatic cholestasis of pregnancy. We conducted a meta-analysis of intrahepatic cholestasis of pregnancy genome-wide association studies in the FinnGen study, deCODE, Estonian Biobank, the Danish Blood Donor Study and Copenhagen Hospital Biobank with 4,738 women with prior ICP and 436,834 female controls. The analysis found 26 genome-wide significant associations of which 10 were novel. Genes in the associated loci were prioritized using lead SNP expression quantitative trait loci associations and colocalization analysis to assess potential causality. The associated loci implicate bile acid synthesis, LDL cholesterol, and lipid metabolism. Additionally, comorbidity, genetic correlation and polygenic risk score analyses further indicated a link between intrahepatic cholestasis of pregnancy and pancreatitis, suggesting shared genetic underpinnings.

Intrahepatic cholestasis of pregnancy (ICP) is a poorly understood liver disease with resolution after delivery and high recurrence risk in subsequent pregnancies[1]. ICP is characterized by pruritus, increased aminotransferase activity and elevated serum bile acids. It is linked to increased risk of fetal complications including preterm labor, fetal asphyxia, meconium-stained amniotic fluid, and stillbirth[2] and associated with gestational diabetes and preeclampsia[3]. Familial predisposition with over tenfold risk increase to ICP observed in sibling studies[4,5] suggests a genetic underpinning, as further demonstrated by the development of promising polygenic risk score for ICP[6]. Reproductive hormones are suggested to trigger the genetic susceptibility resulting in cholestasis and elevated serum bile acids[7]. Previously implicated candidate loci harbor genes regulating bile acid synthesis, transport and cholesterol homeostasis[6,8–10]. Here, we performed a meta-analysis of genome-wide association studies (GWAS) of 4738 women with ICP and 436,834 parous female controls without a history of ICP to

A full list of affiliations appears at the end of the paper. *Lists of authors and their affiliations appear at the end of the paper. ✉e-mail: jaakko.tyrmi@tuni.fi

further enhance our understanding of the biological pathways contributing to this condition and examine its associated comorbidities.

Results

GWAS meta-analysis uncovers 10 novel ICP-associated loci

We set out to dissect the genetic susceptibility to ICP by conducting a meta-analysis (Fig. 1) of ICP GWASs conducted in FinnGen, deCODE, Estonian Biobank (EstBB) and The Danish Blood Donor Study (DBDS) & Copenhagen Hospital Biobank (CHB), with a total of 4738 ICP cases and 436,834 women without a history of ICP as controls. Details of the study cohorts are presented in Supplementary Note 1.

Figure 2 shows the genetic associations of ICP in a meta-analysis of all four studies. Meta-analysis results did not show evidence of genomic inflation ($\lambda GC = 1.043$) or evidence of heterogeneity. In total, our fine-mapping and missense variant identification approach identified 27 index variants (Table 1, Supplementary Data 1) at genome-wide significance ($p < 5 \times 10^{-8}$) residing in 26 loci, 10 of which had not been associated with ICP in earlier GWASs[6,8–10]. Fine-mapping results suggested a single signal in 25 loci, with one locus in chromosome 7 at 7q21 having two separate signals with lead variants rs58238559 and rs59934297 (Supplementary Figs. 2–6). Most genome-wide significant variants associated with ICP were intronic, although the proportion of intergenic and ncRNA intronic variants was higher than expected by chance (Fig. 2d).

We used the online platform FUMA[11] to add annotation to GWAS results and perform gene-based association and enrichment tests (Supplementary Fig. 7). We further tested the association signals in each locus for gene expression colocalization (eQTL) in the most relevant tissues identified in ICP pathophysiology: the liver, whole blood and pancreas (Supplementary Data 2–3) in GTEx v8 dataset. eQTL colocalization identified 24 significant genes across 11 loci in liver tissue, 20 genes across 7 loci in whole blood, and 21 genes across 9 loci in the pancreas.

Furthermore, we identified splicing quantitative trait loci (sQTL, Supplementary Data 4) and protein quantitative loci (pQTL) using

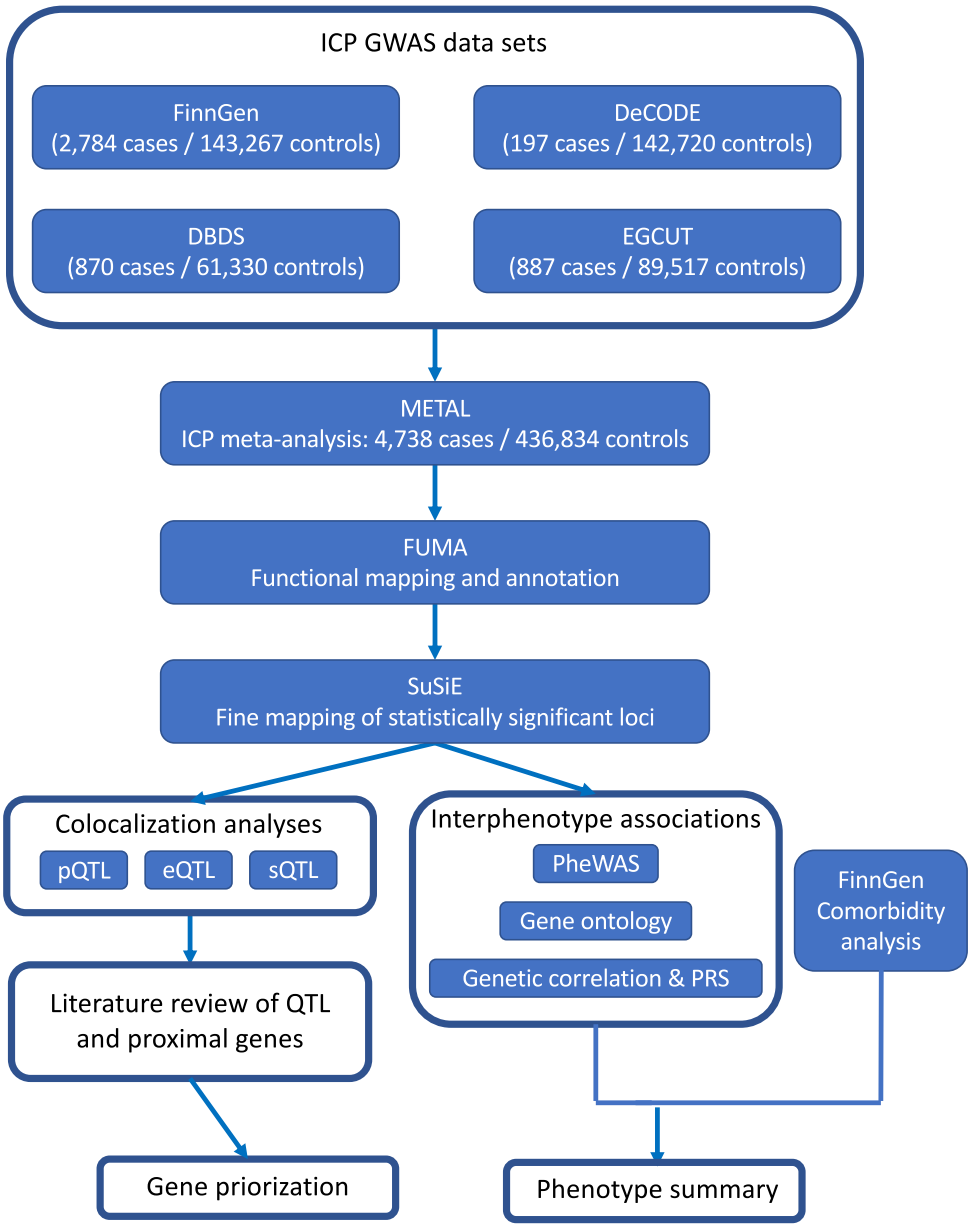

Fig. 1 | Flow chart visualizing the analysis steps. GWAS data from four cohorts (FinnGen, deCODE, DBDS and EGCUT) were combined using METAL. Associated loci were annotated with FUMA and fine-mapped using SuSiE. Downstream analyses included QTL colocalization (eQTL, pQTL, sQTL), PheWAS, gene ontology, genetic correlation and polygenic risk score (PRS) analyses, together with FinnGen comorbidity analyses. These approaches were integrated for gene prioritization.

https://doi.org/10.1038/s41467-026-73122-z

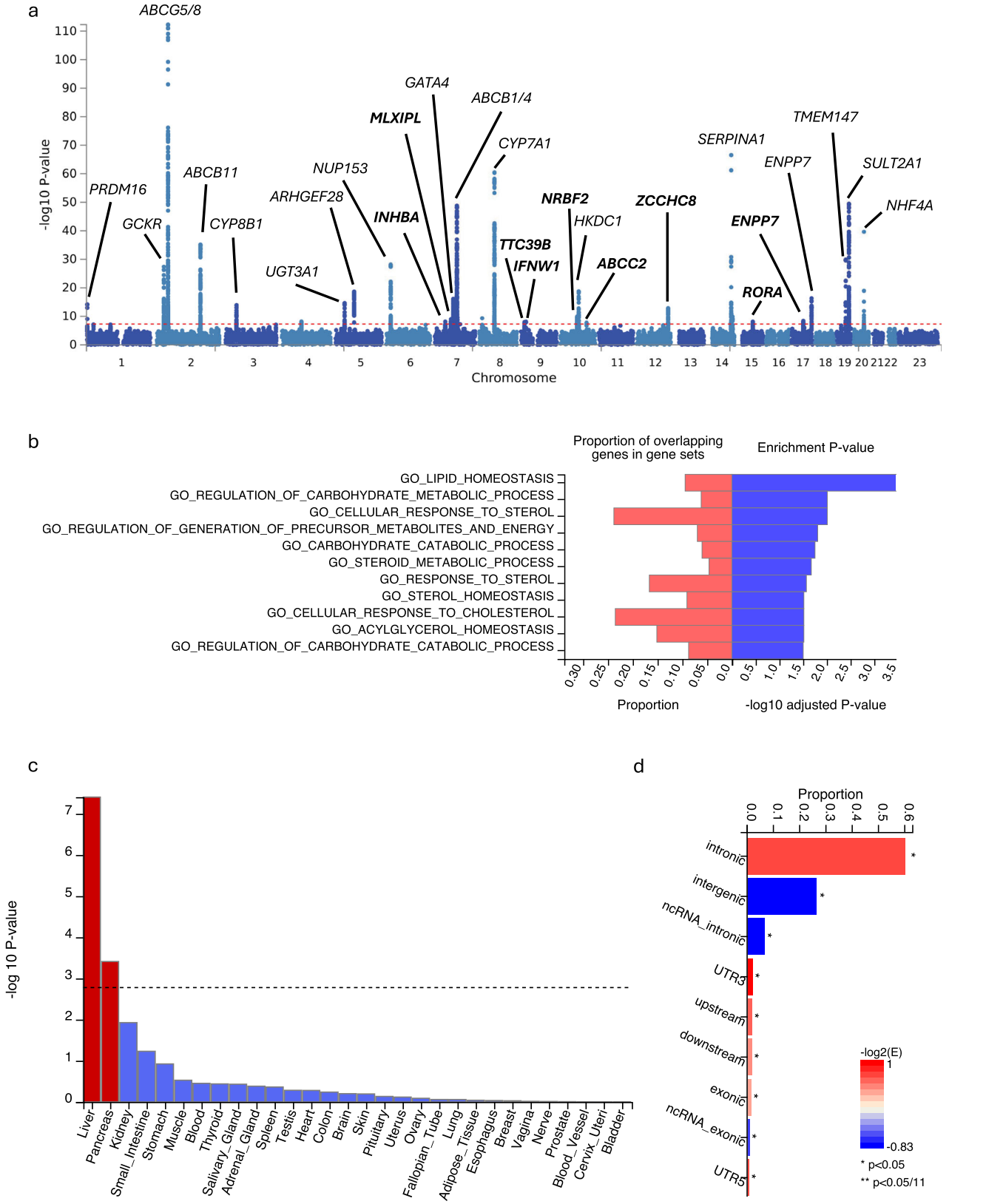

Fig. 2 | Variant summary and gene-set enrichment analysis of ICP GWAS meta-analysis results. a Manhattan plot of ICP meta-analysis results. Novel loci are bolded. **b** Gene Ontology pathway enrichment analysis using MAGMA gene-based association analysis. **c** Tissue-specific expression enrichment analysis results using MAGMA gene-property analysis in FUMA. **d** Functional annotation of variants in LD with the lead variants. Enrichment was tested using a two-sided Fisher's exact test as implemented in FUMA, with Bonferroni correction for multiple comparisons. Source data are provided in Supplementary Data 1, 7 and 8.

https://doi.org/10.1038/s41467-026-73122-z

Table 1 | Lead variants within each credible set in the meta-analysis

Cytoband	Chr:Pos	Lead variant rsID	EA/OA	EAF	P-value	OR (95 % CI)	Candidate gene	Nearest gene	Mis-sense	eQTL	MAG-MA	Lit. rev.
1p36	chr1:3084718	rs2788095	C/A	0.23	6.7E-15	0.82 (0.77-0.86)	PRDM16	■		■	■	■
2p23	chr2:27508073	rs1260326	T/C	0.65	4.1E-28	1.28 (1.23-1.34)	GCKR	■	L		■	■
2p21	chr2:43844604	rs4148211	A/G	0.43	5.8E-113	0.61 (0.59-0.64)	ABCG5/8	■	L		■	■
2q31	chr2:169001056	rs6733156	C/T	0.07	7.6E-36	0.57 (0.53-0.63)	ABCB11	■			■	■
3p22	chr3:42872353	rs12494055	T/C	0.49	1.2E-14	0.85 (0.81-0.88)	CYP8B1	■			■	■
5p13	chr5:35989887	rs73076175	A/G	0.14	2.3E-15	1.27 (1.20-1.34)	UGT3A1	■	LD		■	
5q13	chr5:73623753	rs1981810	A/G	0.76	3.2E-19	0.80 (0.77-0.84)	ARHGEF28	■		■		
6p22	chr6:17705013	rs9396793	T/C	0.40	1.1E-28	0.78 (0.75-0.82)	NUP153	■			■	■
7p14	**chr7:41788929**	**rs10234444**	**G/A**	**0.13**	**8.1E-09**	**1.22 (1.14-1.30)**	**INHBA**	■				■
7q11	**chr7:73507279**	**rs79862839**	**T/C**	**0.12**	**7.7E-17**	**1.33 (1.24-1.42)**	**MLXIPL**		LD		■	
7q21	chr7:87452957	rs58238559	T/C	0.03	1.1E-17	0.49 (0.42-0.58)	ABCB1/4	■	L		■	■
7q21	chr7:87487153	rs59934297	T/G	0.13	1.7E-49	0.60 (0.56-0.64)	ABCB1/4	■	LD		■	■
8q12	chr8:58485902	rs10504255	G/A	0.65	5.5E-61	0.70 (0.67-0.73)	CYP7A1	■			■	■
9p22	**chr9:15304784**	**rs686030**	**C/A**	**0.87**	**1.2E-08**	**1.21 (1.13-1.29)**	**TTC39B**	■		■		
9p21	**chr9:21128732**	**rs12553010**	**G/C**	**0.28**	**8.7E-09**	**1.15 (1.09-1.20)**	**IFNW1**	■				
10q21	**chr10:63338981**	**rs6479894**	**T/C**	**0.86**	**1.1E-10**	**1.26 (1.17-1.35)**	**NRBF2**		■		■	
10q22	chr10:69225511	rs2394529	G/C	0.73	1.9E-19	1.25 (1.19-1.31)	HKDC1	■		■	■	■
10q24	**chr10:99765280**	**rs143176769**	**C/T**	**0.05**	**1.7E-08**	**1.31 (1.19-1.44)**	**ABCC2**	■				
12q24	**chr12:122164766**	**rs10773287**	**A/G**	**0.17**	**1.6E-13**		**CHC8**			■		
14q32.13	chr14:94378610	rs28929474	C/T	0.02	2.8E-67	2.83 (2.52-3.18)	SERPINA1	■			■	■
14q32.2	**chr14:99571781**	**rs2021095**	**T/C**	**0.57**	**1.2E-10**	**1.15 (1.10-1.20)**	**CCDC85C**	■			■	
15q22	**chr15:60640638**	**rs11071547**	**C/T**	**0.42**	**8.8E-09**	**1.13 (1.09-1.18)**	**RORA**	■		■		
17q21	**chr17:47658340**	**rs8078686**	**T/C**	**0.48**	**4.3E-09**	**0.88 (0.85-0.92)**	**TBKBP1**		LD	■	■	
17q25	chr17:79732281	rs34491636	A/G	0.23	5.1E-17	1.23 (1.17-1.30)	ENPP7	■			■	■
19q13.12	chr19:35528620	rs4805129	T/C	0.64	1.2E-30	0.78 (0.74-0.81)	TMEM147			■	■	■
19q13.33	chr19:47873738	rs212100	T/C	0.85	3.7E-50	1.61 (1.52-1.72)	SULT2A1	■		■	■	■
20q13	chr20:44413724	rs1800961	C/T	0.04	2.4E-40	1.88 (1.71-2.06)	HNF4A	■				■

Chr:Pos: Chromosome and position of the lead variant of the locus (according to SuSie fine-mapping); EA/OA: Effect allele / Other allele; EAF: Effect allele frequency; OR (95 % CI): Odds Ratio of the effect allele and it's 95% confidence interval; Nearest gene: Gene nearest to the lead variant; Candidate gene: The gene assumed driving the association signal based on functional evidence and earlier literature. Nearest gene: Candidate gene is nearest gene; Missense: missense or linked variant is missense in the candidiate gene (Supplementary Data 2); eQTL: Candidate gene implicated by eQTL analysis; MAGMA: Candidate gene implicated by MAGMA; Lit. rev.: candidate gene supported by literature review.
Loci not reported in earlier studies are bolded.

deCode SomaScan platform (Supplementary Data 5) and UKBB Olink platform (Supplementary Data 6). We additionally conducted a literature search to aid prioritization of genes within each locus and to provide biological context.

Gene enrichment analysis using MAGMA[12] (Supplementary Fig. 7 Supplementary Data 7–8) confirmed the over-representation of genes expressed in the liver and, notably, in the pancreas (Fig. 2C), a finding not reported in earlier studies.

Several associations are driven by genes related to bile acid metabolism

Several novel loci (7p14, 7q11, 10q24) and previously associated loci (1p36, 3p22, 5p13) harbor genes implicated in bile acid metabolism. The chromosome 1 locus at 1p36 (rs2788095) harbors *PRDM16*, implicated by our eQTL. This gene has been associated with the response to high bile acid concentrations and shown to contribute to fat loss and hypothermia, features often linked to cholestasis[13]. The locus also contains *PRKCZ*, previously associated with bile acid transport[14,15]. Similarly, the nearest gene to lead variant rs12494055 at 3p22, *CYP8B1*, is a well-known regulator of biliary balance that contributes to maintaining the equilibrium of bile acids[16].

At locus 5p13, the lead variant rs73076175 is most proximal to *UGT3A1*, an enzyme that catalyzes reactions where lipophilic substrates are conjugated with N-acetylglucosamine to increase water solubility and enhance excretion. *UGT3A1* plays a role in the metabolism of estradiol and bile acids[17]. This locus also contains a closely linked (r^2 = 0.96) missense variant, rs3756669 (OR = 1.26, 95 % CI = 1.19–1.33, $p = 6.43*10^{-15}$), also known as C121G, which renders *UGT3A1* catalytically inactive[17]. The ICP risk allele C has also been associated with increased estradiol levels in a previous GWAS[18].

The locus at 7p14 (rs10234444) harbors multiple genes, with *INHBA* being the most proximal. *INHBA* encodes a subunit of TGF-beta, which has a central role in liver functioning and has been suggested to affect bile acid metabolism[19,20]. Additionally, *INHBA* is involved in the regulation of follicle-stimulating hormone, which has been proposed to play a role in cholangiocyte development[21]. The lead variant rs79862839 at 7q11 is in linkage disequilibrium with two missense variants (rs35332062, $r^2 = 0.87$, $p = 4.5 * 10^{-15}$; rs3812316, $r^2 = 0.85$, $p = 2.44 * 10^{-15}$) in *MLXIPL*, also known as *ChREBP*. This gene regulates carbohydrate metabolism in the liver, affects bile acid synthesis, and plays a role in the pathophysiology of non-alcoholic fatty liver disease[22]. *MLXIPL* is also located within the critical region associated with Williams syndrome, a rare microdeletion syndrome that impacts

the endocrine system, suggesting a possible shared genetic mechanism[23].

The locus at 10q24 (rs143176769) is associated with the bilirubin transporter *ABCC2*, which has previously been implicated in ICP through candidate gene studies[24] but had not reached genome-wide significance in earlier GWASs. The variant rs11071547 on chromosome 15 at 15q22 is an intron and eQTL for *RORA*. This gene modulates the expression of bile acid synthesis enzymes *CYP7B1* and another well-known ICP associated gene *SULT2A1*, and contributes more broadly to lipid metabolism[25–28].

Overall, our results are consistent with earlier studies of ICP, most notably the study by Dixon et al.[8], which identified 11 risk loci in genes central to bile acid synthesis and transport (Table 1). In contrast, a recent large GWAS in Chinese women by Liu et al.[10] reported a strong East Asian–specific association at 14q24.1, harboring *ERH* and *SLC10A1*, a locus not observed in our analysis and absent from European populations, highlighting possible population-specific genetic contributions to ICP.

Other notable loci

Various other loci have been identified with roles in different biological processes or conditions related to ICP. The locus at 9p22 (rs686030) is near the gene *TTC39B*, which is implicated in cholesterol homeostasis and cholelithiasis[29]. The association at 10q21 (rs6479894) is exonic to *JMJD1C* and is in LD ($r^2 = 0.63$) with missense variant rs10761725 ($p = 8.78*10^{-7}$) in *NRBF2* – a gene also identified in the sQTL analysis (Supplementary Data 4). *JMJD1C* is linked to cholelithiasis[29], while *NRBF2* is a liver-specific transcription factor involved in lipid metabolism[30,31]. The chromosome 17 locus 17q21 (rs8078686) has a tightly linked missense variant (rs4794057, $r^2 = 0.70$, $p = 4.8*10^{-8}$) in *TBKBP1*, known to play an important role in immunological responses, hepatic lipid metabolism and non-alcoholic fatty liver disease[32]. This locus appears to have highly pleiotropic effects, as expression colocalization, sQTL, and pQTL analyses implicated several potential targets, notably a trans effect on *ENPP7*, which has been suggested as a candidate gene for ICP in previous studies[8]. These findings support a broader regulatory influence of this locus beyond local effects.

The lead SNPs associate with several blood biomarkers in UKBB

To further explore the functional implications of ICP-associated variants, we examined their effects on blood biomarkers in UKBB. Across 35 biomarkers, we identified 149 significant associations (Bonferroni-adjusted $p = 7.1 * 10^{-5}$), highlighting pleiotropic links between ICP genetic risk and metabolic traits (Fig. 3, Supplementary Data 9).

ICP associated variants have been shown to have different effects in the UKBB data to the levels of several blood biomarkers[33] implicated in ICP (Fig. 3). For example rs1260326 at 2p23 in *GCKR* and rs79862839 at 7q11 in *MLXIPL* both increasing risk for ICP have opposite effects on triglycerides, sex hormone binding globulin, gamma glutamyl transpeptidase, C-reactive protein and alkaline phosphatase[34]. Majority of ATP-binding cassette variants are in the same cluster, consistent with their shared functional role.

ICP associates with several clinical comorbidities in FinnGen

For each lead SNP (Table 1) we performed a phenome-wide association study (PheWAS) lookup in FinnGen (Fig. 4, Supplementary Data 10), including both men and women. FinnGen PheWAS and GWAS catalog findings for the lead variants indicated strong cholelithiasis and cholecystectomy associations. GWAS Catalog PheWAS analysis suggested that ten of the fine-mapped lead variants were also associated with cholelithiasis, nine with cholecystectomy, and several with weight, height, type 2 diabetes. Interestingly, eight variants were specific to ICP: rs10234444, rs10773287, rs12494055, rs12553010, rs2021095, rs34491636, rs58238559, rs73076175. PheWAS conducted in UKBB did show associations for the variants, albeit only rs73076175 related with liver and bile acid related functions (Supplementary Data 11).

We conducted clinical comorbidity analysis adjusted for age in the FinnGen study by looking at a selection of incident traits (standard FinnGen endpoint definitions) which occurred during or after the index pregnancy (first pregnancy with ICP in cases and first pregnancy in parous controls, Fig. 5, Supplementary Data 12). The association of ICP with two other common pregnancy complications, preeclampsia (HR [95% CI] = 2.10 [1.81–2.43], $P = 3.08x10^{-22}$) and gestational diabetes (HR [95% CI] = 4.00 [3.64–4.41], $P = 1.10x10^{-179}$), is in agreement with the results of the Swedish Medical Birth Register study[3], which was our main basis for selecting the incident traits. Women with a history of ICP had an increased risk for developing cholelithiasis, cholecystitis, cholangitis and other biliary diseases[35]. ICP was also associated with polycystic ovary syndrome, thyroid disorders, type 2 diabetes, Crohn's disease (but not ulcerative colitis), polyarthropathies, systemic connective tissue disease and autoimmune diseases in general as reported in previous studies[3,35].

To further examine the genetic pleiotropy, we constructed a polygenic risk score (PRS) for ICP, regression coefficients from a meta-analysis excluding FinnGen from the cohorts (Fig. 6, Supplementary Data 13). The PRS was strongly associated with ICP (OR = 1.54 per SD, $p < 10^{-16}$). The PRS was also associated with cholelithiasis (OR = 1.11 per SD, $p = 4.52x10^{-71}$) in women, and OR = 1.06 per SD, $p = 3.46x10^{-16}$ in men-only analysis. Pancreatic diseases were significantly (OR = 1.06 per SD, $p = 0.0003$) associated with ICP PRS indicating possible shared genetic underpinnings of the diseases. These results were also reflected by genetic correlations (r_g) calculated using LDSC in between similar FinnGen endpoints as PRS, showing r_g of 0.31 (95 % CI 0.13-0.50, $p = 0.01$, Supplementary Data 14, Supplementary Fig. 8).

Potential genetic links between ICP and pancreatitis

As the comorbidity, polygenic risk associations, genetic correlations and FUMA Gene Ontology (GO) enrichment analysis pointed towards the possible importance of pancreatitis, we further reviewed the nearby genes in all associated loci considering evidence pointing to their potential contribution to this inflammation. The locus at 1p36 had significant trans-pQTL hits with *PRSS3* and *CTRC* (Supplementary Data 5-6). *PRSS3* has little previous evidence supporting a role in ICP, but association with pancreatitis has been suspected[36]. *CTRC* has also been shown to harbor variants contributing to pancreatitis[37]. Furthermore, at 2p21 the candidate gene *ABCG8* has been associated with pancreatitis[38,39]. PheWAS analyses conducted in FinnGen or UKBB did not indicate significant association of ICP-associated variants and pancreatitis.

Strengths and limitations

Our study has a large sample size compared to previous studies examining the genetic basis of ICP, allowing us to discover several novel associated loci. The integration of fine-mapping, colocalization, and biomarker analyses, and careful literature review of the proposed causal genes strengthen the biological relevance of the findings. However, the focus on European ancestry populations limits generalizability, and functional validation is needed to confirm causal mechanisms. Future studies in diverse populations and experimental models are required to refine these insights and explore their clinical applications.

Discussion

In summary, this study identifies 26 genetic loci associated with ICP, of which 10 are novel, expanding our understanding of the disease's

https://doi.org/10.1038/s41467-026-73122-z

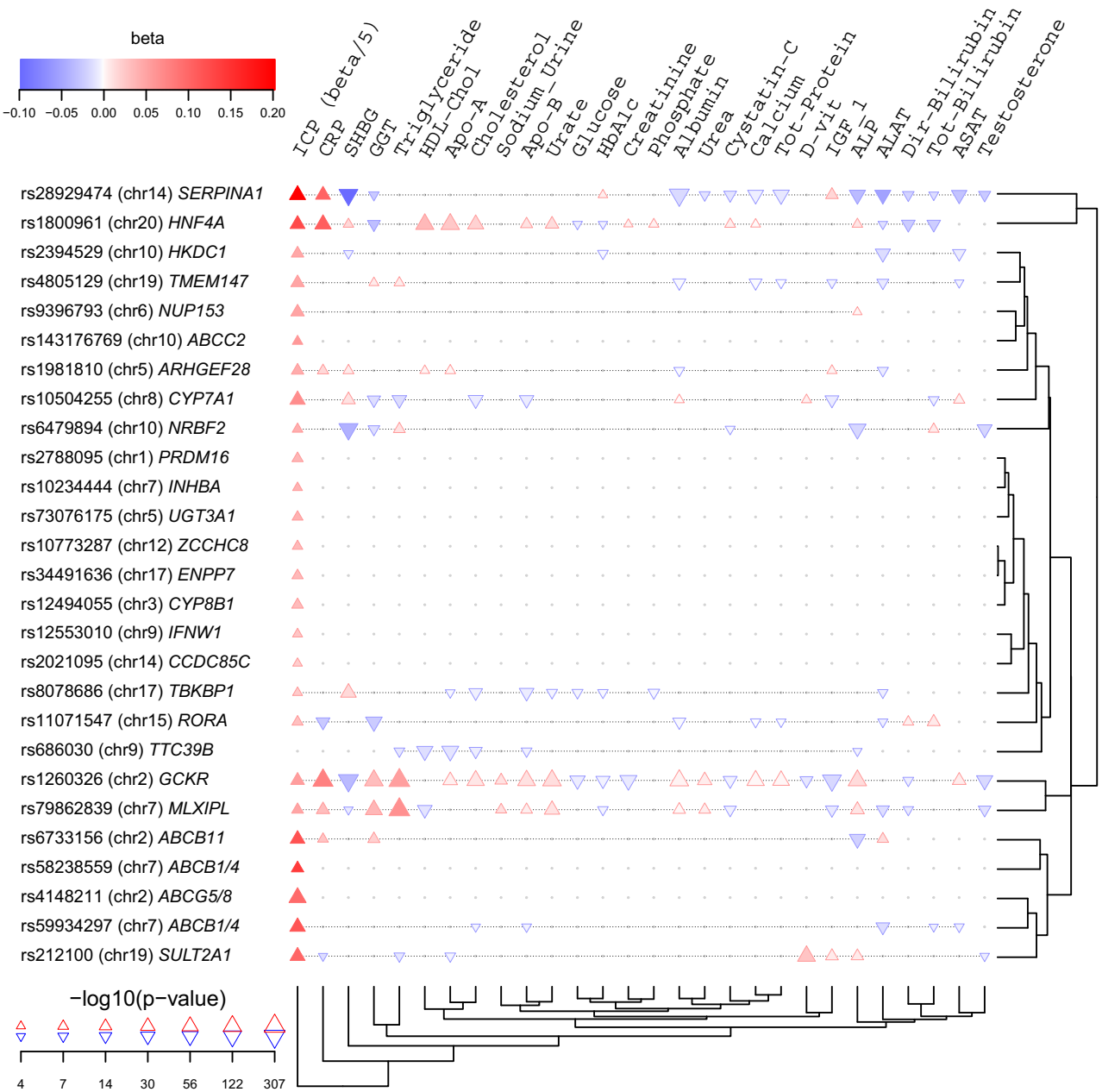

Fig. 3 | Analysis of 27 ICP-associated variants (or their proxies) subset from GWAS summary statistics of 35 biomarkers in Sinnott-Armstrong et al.[33] for both sexes. Effects are expressed per ICP risk allele (Table 1), and for comparison, the variants' associations with ICP (beta/5, from Fig. 4) are also shown. Variants and traits have been ordered by hierarchical clustering to show common patterns. We used a stringent *p*-value cutoff: 0.05 / (27 SNPs * 35 biomarkers) = 5.29×10^{-5}. In total we found 149 significant associations. Source data are provided in Supplementary Data 9.

genetic architecture. These findings highlight the crucial roles of bile acid, cholesterol, and lipid metabolism in the pathophysiology of ICP and suggest potential genetic links to pancreatitis. Comorbidity analyses revealed that ICP frequently co-occurs with inflammatory conditions of the liver and pancreas, autoimmune diseases, pregnancy complications, and polycystic ovary syndrome. Genetic analyses also implicated several liver- and gallbladder-related diseases, including cholelithiasis, cholecystitis, acute pancreatitis, and nonalcoholic fatty liver disease.

Notably, comorbidity analysis indicated association with pancreatitis. Gallstone disease is a shared risk factor in intrahepatic cholestasis and pancreatitis. Pre-pregnancy gallstone disease is associated with increased odds of subsequent ICP[40], and gallstone migration into the common bile duct is a frequent cause of acute pancreatitis[41]. Here, the association between pancreatitis and ICP was further supported by PRS associations and enrichment of ICP-associated gene expression in the pancreas. Several associated loci, including *ABCG8*, *CTRC*, and *PRSS3*, have also been implicated in pancreatitis. To our knowledge, this genetic overlap has not been reported in previous genetic studies of ICP. Nonetheless, the specific genetic link remains inconclusive and requires further validation.

Taken together, our findings lay a strong foundation for future research into the pathophysiology of ICP and its related comorbidities.

https://doi.org/10.1038/s41467-026-73122-z

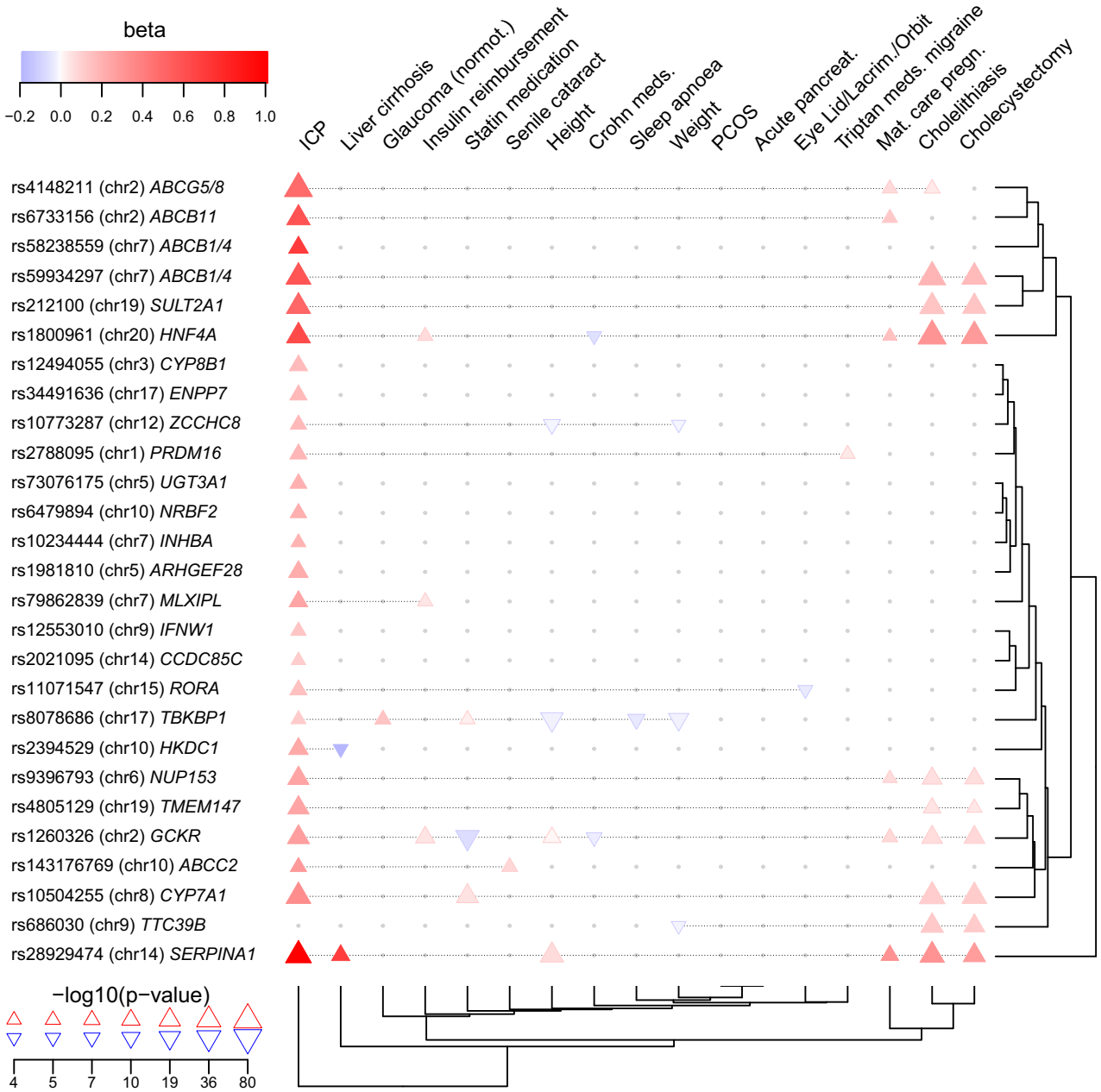

Fig. 4 | PheWAS of the ICP meta-analysis lead SNPs in the FinnGen release 11. In this release, PheWAS analyses have been done for 2440 FinnGen endpoints, leading to Bonferroni-corrected p-value cutoff 0.05/2440 = 2.05×10^{-5}.Variants and traits have been ordered by hierarchical clustering to show common patterns. Many additional associations that were specific only to *GCKR* or *SERPINA1* were omitted in this figure for clarity and are given in the Supplementary Data 10. Depending on the phenotype, the range of cases is 1470–155745 and range of controls is 159020–448864. PheWAS data is available at https://r11.finngen.fi.

Methods

Study cohorts and association analyses

All research was conducted in accordance with relevant ethical regulations. Study protocols for each cohort were approved by the appropriate national and institutional ethics committees and data protection authorities, and all participants provided informed consent. Detailed cohort descriptions including within cohort association analysis workflows are described in Supplementary Note. In brief, ICP cases were identified using International Classification of Diseases (ICD) codes O26.6 (ICD-10), 6467 (ICD-9) and 639.0 (ICD-8). Controls were restricted to parous women without a diagnosis of ICP to ensure phenotypic comparability with cases and to avoid potential bias related to pregnancy exposure. Association analysis was then performed in each cohort using either SAIGE[42] or REGENIE[43] software while adjusting for age and at least 10 principal components, and any cohort specific variables.

Meta-analyses and fine-mapping

Summary statistics obtained from these analyses in each cohort were then used to perform inverse variance–weighted meta-analysis with the METAL software[44]. Genome-wide significance was set to $P < 5*10^{-8}$. To further characterize the loci we performed fine-mapping using SuSie[45]. We obtained the 99 % credible sets from the workflow and examined these results in our downstream analyses.

Gene-based and gene-set enrichment analyses were performed using MAGMA as implemented in FUMA platform[11]. Gene-level association statistics were computed using the SNP-wise mean model,

https://doi.org/10.1038/s41467-026-73122-z

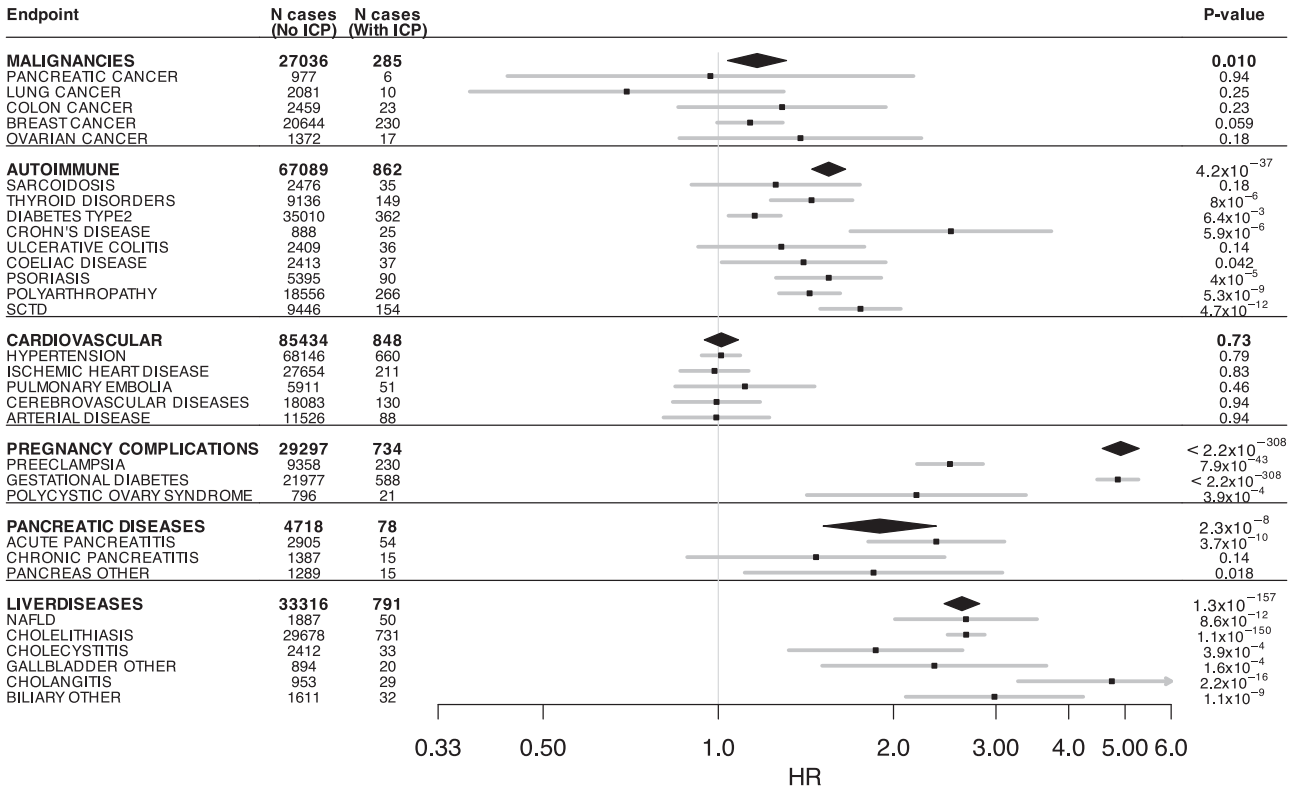

Fig. 5 | Incident comorbidity analysis in FinnGen. Cox regression was used to model the effect of ICP on incident comorbidities, with ICP preceding the comorbity. Age was used as the time scale. For ICP cases, baseline age was defined as 6 months prior to first ICP diagnosis (approximating pregnancy onset), and for controls as 9 months prior to first delivery. Follow-up continued until the first occurrence of the comorbidity, death, or end of follow-up. Age at a comorbitity was taken from the first recorded event of the comorbidity in question. Source data are provided in Supplementary Data 12.

incorporating linkage disequilibrium estimated from a population-matched reference panel. Competitive gene-set enrichment analysis was then conducted using 4728 curated gene sets and 6166 Gene Ontology terms from MSigDB[46], testing whether genes within each set were more strongly associated with ICP than genes outside the set. This approach tests enrichment relative to the genome-wide background while accounting for gene size, SNP density, and LD structure.

To further prioritize genes we utilized functional annotations from GenBank[47] and UniProt[48]. To supplement these database resources, we conducted a comprehensive literature review to identify prior studies relevant to the genes of interest within 2 Mbp window around the lead variants.

PheWAS analyses

The 27 ICP-associated variants were tested for associations with 2440 endpoints in the FinnGen release 11. Thus we used Bonferroni-corrected *p*-value threshold $0.05/2440 = 2.05 \times 10^{-5}$. Selected endpoints having associations with multiple variants are shown in Fig. 5. Hierarchical clustering of traits and SNPs were based on euclidean distances of the regression coefficients (betas), using R function "hclust", and plotted as dendrograms. R version 4.4.3 was used for plotting. Similarly, we conducted pheWAS in UKBB, using a suggestive *p*-value threshold of 1.0×10^{-5}.

Furthermore, the 27 ICP-associated variants were tested for association with 33 biomarkers and sex hormones in up to 392,091 unrelated individuals of non-Finnish European ancestry in UKBB in sex-combined strata using PLINK2. The biomarkers tested were all log-transformed and adjusted according to the protocol in Sinnott-Armstrong et al.[33]. No statin adjustment was made for cholesterol. Bonferroni corrected *p*-value threshold of 5.29×10^{-5} was used (0.05 / (27 * 35)). Hierarchical clustering of traits and SNPs were based on euclidean distances of the regression coefficients (betas), using R-function "hclust", and plotted as dendrograms. R version 4.4.3 was used for plotting.

Polygenic risk score association analyses

Based on meta-analysis regression coefficients obtained after re-running the GWAS while omitting FinnGen, we constructed a polygenic risk score (PRS) for ICP using the PRS-CS pipeline[49] implemented in FinnGen. The analysis was based on FinnGen Data Freeze 13 with 218277 men and 281909 women. Logistic regression, adjusted for the FinnGen baseline age, was used to model the odds ratio of one SD of PRS to the ICP comorbidities. The ICP PRS was first adjusted for the first 5 PCs, FinnGen study cohort, and genotyping batch effects by linear regression. The standardized regression residuals were used in the analyses as the PRS. To gain statistical power and reduce multiple testing burden, we restricted the analyses to those endpoints having at least 250 cases. For each endpoint, controls were those non-cases which were not excluded from controls according to the FinnGen control exclusion rules.

Clinical comorbidity analyses

The analysis was done using FinnGen release 13. Cox regression was used to model comorbidity-free survival starting from pregnancy index-age in ICD-cases ($N = 672$) and parous female controls ($N = 39765$). The index age was extrapolated to the assumed beginning of pregnancy as 6 months prior to first ICP diagnosis in cases, or 9 months prior to first delivery in controls. Age was used as the time scale, so the follow-up ended at comorbidity event or in case of no event, to death or end of follow-up, if the person was still alive at that point. For each incident comorbidity endpoint analysis, the respective prevalent comorbidity cases were excluded, as well as any non-cases which were excluded from controls according to the FinnGen control

https://doi.org/10.1038/s41467-026-73122-z

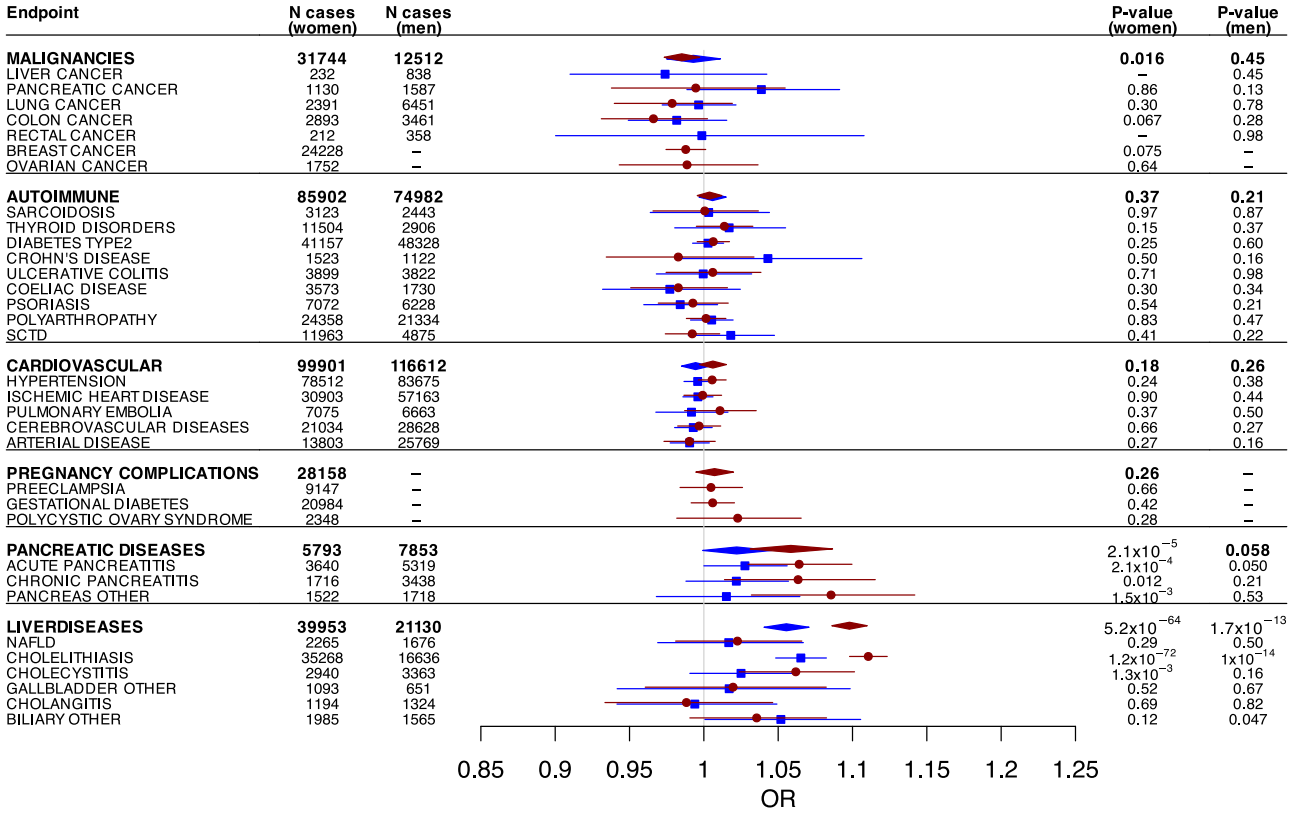

Fig. 6 | Associations of the meta-analysis–based ICP polygenic risk score (PRS) with ICP comorbidities, shown separately for women (red) and men (blue). Logistic regression models were adjusted for FinnGen baseline age. Source data are provided in Supplementary Data 13.

exclusion rules. The analyses were restricted to those endpoints which had at least 5 cases among the women who had ICP.

QTL colocalization and annotation

We performed cis eQTL colocalization in GTEx v8 dataset for our fine-mapped loci using coloc[50] and focused on signals where posterior probability for a single causal variant [PP4] was >0.8. We also tested whether the variant itself and variants in high LD (r^2 > 0.8) could affect the coding sequence of genes or their splicing based on the Variant Effect Predictor tool[51], as described previously[29]. Furthermore, we tested whether the variant itself and variants in high LD (r^2 > 0.8) affected levels of 4,907 proteins measured by SomaScan in deCODE genetics data and levels of 2,931 proteins measured by Olink in the UKBB[52]. Finally, we tested whether the variant itself and variants in high LD (r^2 > 0.8) was a top (most significantly associated) variant per Mb bin and per experimental factor ontology (EFO) term reported in the NHGRI-EBI GWAS Catalog (downloaded 18 August 2022).

Furthermore, using the fine-mapped variants we identified and protein quantitative loci (pQTL) using deCode SomaScan platform and UKBB Olink platform. Splice site QTLs were searched for lead using Gtex v8 data.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The summary statistics from the ICP meta-are available via GWAS Catalog using accession number GCST90832988. Individual-level genotype and phenotype data are available under controlled access due to legal, ethical and data protection restrictions, as they contain potentially identifiable health and genetic information. Formal approval is required to access these data; details are available at https://www.finngen.fi/en/access_results. Access to data from deCODE genetics, the Estonian Biobank, and the Danish Blood Donor Study/Copenhagen Hospital Biobank is subject to approval by the respective biobank governance bodies and national regulatory authorities in accordance with local legislation. Genome-wide summary statistics used for genetic correlation analyses are available from the MRC-IEU GWAS database (https://gwas.mrcieu.ac.uk/). GTEx v8 data are available from the GTEx Portal (https://gtexportal.org/). Source data underlying figures are provided with this paper.

Code availability

No custom software central to the study was developed. All analyses were performed using publicly available, open-source software.

References

1. Roediger, R. & Fleckenstein, J. Intrahepatic Cholestasis of Pregnancy: Natural History and Current Management. Semin. Liver Dis. **41**, 103–108 (2021).
2. Ovadia, C. et al. Association of adverse perinatal outcomes of intrahepatic cholestasis of pregnancy with biochemical markers: results of aggregate and individual patient data meta-analyses. The Lancet **393**, 899–909 (2019).
3. Wikström Shemer, E. A. et al. Intrahepatic cholestasis of pregnancy and cancer, immune-mediated and cardiovascular diseases: A population-based cohort study. J. Hepatol. **63**, 456–461 (2015).
4. Dixon, P. & Williamson, C. The molecular genetics of intrahepatic cholestasis of pregnancy. Obstet. Med. **1**, 65–71 (2008).

5. Eloranta, M.-L., Heinonen, S., Mononen, T. & Saarikoski, S. Risk of obstetric cholestasis in sisters of index patients. Clin. Genet. **60**, 42–45 (2001).
6. Pujol Gualdo, N. et al. Atlas of genetic and phenotypic associations across 42 female reproductive health diagnoses. Nat. Med. https://doi.org/10.1038/s41591-025-03543-8 (2025) .
7. Dixon, P. H. & Williamson, C. The pathophysiology of intrahepatic cholestasis of pregnancy. Clin. Res. Hepatol. Gastroenterol. **40**, 141–153 (2016).
8. Dixon, P. H. et al. GWAS meta-analysis of intrahepatic cholestasis of pregnancy implicates multiple hepatic genes and regulatory elements. Nat. Commun. **13**, 4840 (2022).
9. Karjalainen, M. K. et al. Genome-wide characterization of circulating metabolic biomarkers. Nature **628**, 130–138 (2024).
10. Liu, Y. et al. Genetic study of intrahepatic cholestasis of pregnancy in Chinese women unveils East Asian etiology linked to historic HBV epidemic. J. Hepatol. S0168827824027089 (2024) https://doi.org/10.1016/j.jhep.2024.11.008.
11. Watanabe, K., Taskesen, E., Van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. Nat. Commun. **8**, 1826 (2017).
12. De Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLOS Comput. Biol. **11**, e1004219 (2015).
13. Zhou, W. et al. Pathological bile acid concentrations in chronic cholestasis cause adipose mitochondrial defects. JHEP Rep **5**, 100714 (2023).
14. Chen, F., Ellis, E., Strom, S. C. & Shneider, B. L. ATPase Class I Type 8B Member 1 and Protein Kinase C Induce the Expression of the Canalicular Bile Salt Export Pump in Human Hepatocytes.
15. Frankenberg, T. et al. The membrane protein ATPase class I type 8B member 1 signals through protein kinase C zeta to activate the farnesoid X receptor. Hepatology **48**, 1896–1905 (2008).
16. Pathak, P. & Chiang, J. Y. L. Sterol 12α-Hydroxylase Aggravates Dyslipidemia by Activating the Ceramide/mTORC1/SREBP-1C Pathway via FGF21 and FGF15. Gene Expr **19**, 161–173 (2019).
17. Mackenzie, P. I. et al. Identification of UDP Glycosyltransferase 3A1 as a UDP N-Acetylglucosaminyltransferase. J. Biol. Chem. **283**, 36205–36210 (2008).
18. Schmitz, D. et al. Genome-wide Association Study of Estradiol Levels and the Causal Effect of Estradiol on Bone Mineral Density. J. Clin. Endocrinol. Metab. **106**, e4471–e4486 (2021).
19. Fabregat, I. et al. TGF -β signalling and liver disease. FEBS J **283**, 2219–2232 (2016).
20. Matsubara, T. et al. TGF-β-SMAD3 signaling mediates hepatic bile acid and phospholipid metabolism following lithocholic acid-induced liver injury. J. Lipid Res. **53**, 2698–2707 (2012).
21. Hall, C. et al. Regulators of Cholangiocyte Proliferation. Gene Expr **17**, 155–171 (2017).
22. Régnier, M., Carbinatti, T., Parlati, L., Benhamed, F. & Postic, C. The role of ChREBP in carbohydrate sensing and NAFLD development. Nat. Rev. Endocrinol. **19**, 336–349 (2023).
23. Kozel, B. A. et al. Williams syndrome. (2022).
24. Dixon, P. H. et al. An expanded role for heterozygous mutations of ABCB4, ABCB11, ATP8B1, ABCC2 and TJP2 in intrahepatic cholestasis of pregnancy. Sci. Rep. **7**, 11823 (2017).
25. Huang, J. et al. Kinetic analysis of bile acid sulfation by stably expressed human sulfotransferase 2A1 (SULT2A1). Xenobiotica **40**, 184–194 (2010).
26. Ou, Z. et al. Regulation of the Human Hydroxysteroid Sulfotransferase (SULT2A1) by RORα and RORγ and Its Potential Relevance to Human Liver Diseases. Mol. Endocrinol. **27**, 106–115 (2013).
27. Ren, S. et al. Regulation of oxysterol 7α-hydroxylase (CYP7B1) in the rat. Metabolism **52**, 636–642 (2003).
28. Wada, T. et al. Identification of Oxysterol 7α-Hydroxylase (Cyp7b1) as a Novel Retinoid-Related Orphan Receptor α (RORα) (NR1F1) Target Gene and a Functional Cross-Talk between RORα and Liver X Receptor (NR1H3). Mol. Pharmacol. **73**, 891–899 (2008).
29. Ferkingstad, E. et al. Genome-wide association meta-analysis yields 20 loci associated with gallstone disease. Nat. Commun. **9**, 5101 (2018).
30. De Vries, P. S. et al. Multiancestry Genome-Wide Association Study of Lipid Levels Incorporating Gene-Alcohol Interactions. Am. J. Epidemiol. **188**, 1033–1054 (2019).
31. Gellert-Kristensen, H. et al. Identification and Replication of Six Loci Associated With Gallstone Disease. Hepatology **70**, 597–609 (2019).
32. Huh, J. Y. & Saltiel, A. R. Roles of IκB kinases and TANK-binding kinase 1 in hepatic lipid metabolism and nonalcoholic fatty liver disease. Exp. Mol. Med. **53**, 1697–1705 (2021).
33. Sinnott-Armstrong, N. et al. Genetics of 35 blood and urine biomarkers in the UK Biobank. Nat. Genet. **53**, 185–194 (2021).
34. Zhan, Y., Xu, T., Chen, T. & Wang, X. Intrahepatic cholestasis of pregnancy and maternal dyslipidemia: a systematic review and meta-analysis. Acta Obstet. Gynecol. Scand. **101**, 719–727 (2022).
35. Hämäläinen, S., Turunen, K., Mattila, K. J., Kosunen, E. & Sumanen, M. Intrahepatic cholestasis of pregnancy and comorbidity: A 44-year follow-up study. Acta Obstet. Gynecol. Scand. **98**, 1534–1539 (2019).
36. Masamune, A. et al. Genetic background is different between sentinel and recurrent acute pancreatitis. J. Gastroenterol. Hepatol. **26**, 974–978 (2011).
37. Zhou, J. & Sahin-Tóth, M. Chymotrypsin C mutations in chronic pancreatitis. J. Gastroenterol. Hepatol. **26**, 1238–1246 (2011).
38. Bourgault, J. et al. Proteome-Wide Mendelian Randomization Identifies Causal Links Between Blood Proteins and Acute Pancreatitis. Gastroenterology **164**, 953–965.e3 (2023).
39. Sakaue, S. et al. A cross-population atlas of genetic associations for 220 human phenotypes. Nat. Genet. **53**, 1415–1424 (2021).
40. Marschall, H.-U., Wikström Shemer, E., Ludvigsson, J. F. & Stephansson, O. Intrahepatic cholestasis of pregnancy and associated hepatobiliary disease: A population-based cohort study. Hepatology **58**, 1385–1391 (2013).
41. Frossard, J.-L., Steer, L. M. & Pastor, C. M. Acute pancreatitis. Lancet **371**, 143–152 (2008).
42. Zhou, W. et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. Nat. Genet. **50**, 1335–1341 (2018).
43. Mbatchou, J. et al. Computationally efficient whole-genome regression for quantitative and binary traits. Nat. Genet. **53**, 1097–1103 (2021).
44. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics **26**, 2190–2191 (2010).
45. Wang, G., Sarkar, A., Carbonetto, P. & Stephens, M. A Simple New Approach to Variable Selection in Regression, with Application to Genetic Fine Mapping. J. R. Stat. Soc. Ser. B Stat. Methodol. **82**, 1273–1300 (2020).
46. Liberzon, A. et al. Molecular signatures database (MSigDB) 3.0. Bioinformatics **27**, 1739–1740 (2011).
47. Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Sayers, E. W. GenBank. Nucleic Acids Res **44**, D67–D72 (2016).
48. The UniProt Consortium UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. **47**, D506–D515 (2019).
49. Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A. & Smoller, J. W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat. Commun. **10**, 1776 (2019).

50. Giambartolomei, C. et al. Bayesian Test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. **10**, e1004383 (2014).
51. McLaren, W. et al. The ensembl variant effect predictor. Genome Biol **17**, 122 (2016).
52. Eldjarn, G. H. et al. Large-scale plasma proteomics comparisons through genetics and disease associations. Nature **622**, 348–358 (2023).

Acknowledgements

We want to acknowledge the participants and investigators of the FinnGen study. The FinnGen project is funded by two grants from Business Finland (HUS 4685/31/2016 and UH 4386/31/2016) and the following industry partners: AbbVie Inc., AstraZeneca UK Ltd, Biogen MA Inc., Bristol Myers Squibb (and Celgene Corporation & Celgene International II Sàrl), Genentech Inc., Merck Sharp & Dohme LCC, Pfizer Inc., GlaxoSmithKline Intellectual Property Development Ltd., Sanofi US Services Inc., Maze Therapeutics Inc., Janssen Biotech Inc, Novartis AG, and Boehringer Ingelheim International GmbH. Following biobanks are acknowledged for delivering biobank samples to FinnGen: Auria Biobank (www.auria.fi/biopankki), THL Biobank (www.thl.fi/biobank), Helsinki Biobank (www.helsinginbiopankki.fi), Biobank Borealis of Northern Finland (https://www.ppshp.fi/Tutkimus-ja-opetus/Biopankki/Pages/Biobank-Borealis-briefly-in-English.aspx), Finnish Clinical Biobank Tampere (www.tays.fi/en-US/Research_and_development/Finnish_Clinical_Biobank_Tampere), Biobank of Eastern Finland (www.ita-suomenbiopankki.fi/en), Central Finland Biobank (www.ksshp.fi/fi-FI/Potilaalle/Biopankki), Finnish Red Cross Blood Service Biobank (www.veripalvelu.fi/verenluovutus/biopankkitoiminta), Terveystalo Biobank (www.terveystalo.com/fi/Yritystietoa/Terveystalo-Biopankki/Biopankki/) and Arctic Biobank (https://www.oulu.fi/en/university/faculties-and-units/faculty-medicine/northern-finland-birth-cohorts-and-arctic-biobank). All Finnish Biobanks are members of BBMRI.fi infrastructure (https://www.bbmri-eric.eu/national-nodes/finland/). Finnish Biobank Cooperative -FINBB (https://finbb.fi/) is the coordinator of BBMRI-ERIC operations in Finland. The Finnish biobank data can be accessed through the Fingenious® services (https://site.fingenious.fi/en/) managed by FINBB. Full list of FinnGen investigators is provided in Supplementary Data 15. We thank the participants of the DBDS Genomic Consortium for their contributions. List of DBDS Genomic Consortium members is provided in Supplementary Data 16. We want to acknowledge the participants of the Estonian Biobank for their contributions. The Estonian Genome Center analyses were carried out in the High Performance Computing Center, University of Tartu. This work was supported by the Estonian Research Council grant PRG1291 (Systematic phenome-wide search for genetic modulators in health and disease) and Ministry of Education and Research Centres of Excellence grant TK214 (Centre of Excellence for Personalized Medicine). The Estonian Biobank Research Team was responsible for data collection, genotyping, quality control and imputation. List of Estonian Biobank Research Team is provided in Supplementary Data 17. The authors declare no relevant funding.

Author contributions

C.M.L., M.D., A.S.H., H.M.L. conceptualized the study. J.S.T., J.Ka., S.S.V., C.B., P.S., T.T. and A.S.H. analysed data and generated results. J.S.T., S.S.V., J.Ke., M.N., P.S., C.M.L., A.S.H. and H.M.L interpreted the results. J.S.T., M.N., T.L., T.T., C.M.L., A.S.H. and H.M.L. wrote the original manuscript. H.S.N., D.W., T.R., P.S., P.P., T.L., M.D., A.S.H.,H.M.L., FinnGen, Estonian Biobank research team and DBDS Genomic Consortium provided data. S.M.L., S.B., B.A., M.T.B., C.E., H.U., O.B.P., K.B., E.S., C.M., M.S., A.S., S.R.O., H.S.N., D.W., P.S., K.S. and P.P. provided additional study resources. All authors contributed to revising the content of the manuscript and approved the final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-026-73122-z.

Correspondence and requests for materials should be addressed to Jaakko S. Tyrmi.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

[1]Center for Child, Adolescent and Maternal Health Research, Faculty of Medicine and Health Technology, Tampere, Tampere University, Tampere, Finland. [2]Systems epidemiology, Research Unit of Population Health, Faculty of Medicine, University of Oulu and Biocenter Oulu, Oulu, Finland. [3]Institute for Molecular Medicine Finland, FIMM, Helsinki Institute of Life Science - HiLIFE, University of Helsinki, Helsinki, Finland. [4]Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK. [5]Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen, Denmark. [6]Department of Clinical Immunology, Aalborg University Hospital, Aalborg, Denmark. [7]Department of Clinical Immunology, Odense University Hospital, Odense, Denmark. [8]Department of Clinical Immunology, Aarhus University Hospital, Aarhus, Denmark. [9]Statens Serum Institut, Copenhagen, Denmark. [10]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark. [11]Department of Obstetrics and Gynecology, Copenhagen University Hospital Hvidovre, Copenhagen, Denmark. [12]Department of Clinical Immunology, Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark. [13]deCODE genetics/Amgen, Reykjavik, Iceland. [14]Department of Health Science and Technology, Aalborg University,

Aalborg, Denmark. [15]Estonian Genome Center, Institute of Genomics, University of Tartu, Tartu, Estonia. [16]Institute of Clinical Medicine, School of Medicine, Faculty of Health Sciences, University of Eastern Finland, Kuopio, Finland. [17]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA. [18]Department of Computing, University of Turku, Turku, Finland. [19]Department of Obstetrics and Gynecology, Tampere University Hospital, The Wellbeing Services County of Pirkanmaa, Tampere, Finland. [20]Medical and Clinical Genetics, University of Helsinki and Helsinki University Hospital, Helsinki, Finland. [21]These authors contributed equally: Aki S. Havulinna, Hannele M. Laivuori. ✉e-mail: jaakko.tyrmi@tuni.fi

FinnGen

Jaakko S. Tyrmi[1,2]✉, Juha Karjalainen[3], Susanna M. Lemmelä[3], Johannes Kettunen[2], Taru Tukiainen[3,16], Aki S. Havulinna[3,18,21] & Hannele M. Laivuori[1,3,19,20,21]

Estonian Biobank research team

Priit Palta[15] & Triin Laisk[15]

DBDS Genomic Consortium

Søren Brunak[5], Bitten Aagaard[6], Mie T. Bruun[7], Christian Erikstrup[8], Henrik Ullum[9], Ole B. Pedersen[10], Erik Sørensen[12], Christina Mikkelsen[12], Michael Schwinn[12], Sisse R. Ostrowski[12,14] & Mette Nyegaard[9,14]