

# Genome-wide association study identifies five risk loci for pernicious anemia and implicates the role of HLA-DR15 haplotype

Triin Laisk<sup>1, #</sup>, Maarja Lepamets<sup>1, 2</sup>, Reedik Mägi<sup>1</sup>

<sup>1</sup> Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, Estonia

<sup>2</sup> Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia

# corresponding author T. Laisk, [triin.laisk@ut.ee](mailto:triin.laisk@ut.ee)

Keywords: genome-wide association study; pernicious anemia, B12 deficiency

## Abstract

Pernicious anemia is a rare condition characterized by vitamin B12 deficiency anemia due to lack of intrinsic factor, often caused by autoimmune gastritis. Patients with pernicious anemia have a higher incidence of other autoimmune disorders, such as type 1 diabetes, vitiligo and autoimmune thyroid issues. Therefore, the disease has a clear autoimmune basis, although the genetic susceptibility factors have thus far remained poorly studied. We conducted a genome-wide association study meta-analysis in 2,166 cases and 659,516 European controls from population-based biobanks and identified genome-wide significant signals in or near the *PTPN22* (rs6679677,  $p=1.91 \times 10^{-24}$ , OR=1.63), *PNPT1* (rs12616502,  $p=3.14 \times 10^{-8}$ , OR=1.70), *HLA-DQB1* (rs28414666,  $p=1.40 \times 10^{-16}$ , OR=1.38), *IL2RA* (rs2476491,  $p=1.90 \times 10^{-8}$ , OR=1.22) and *AIRE* (rs74203920,  $p=2.33 \times 10^{-9}$ , OR=1.83) genes, thus providing the first robust associations between pernicious anemia and genetic risk factors. We further mapped the susceptibility in the HLA region to the HLA-DR15 haplotype. Analysis of associated diagnoses and disease trajectories confirm the association between pernicious anemia and thyroid issues, vitiligo, gastritis, stomach cancer, osteoporosis and other diagnoses.

## Main text

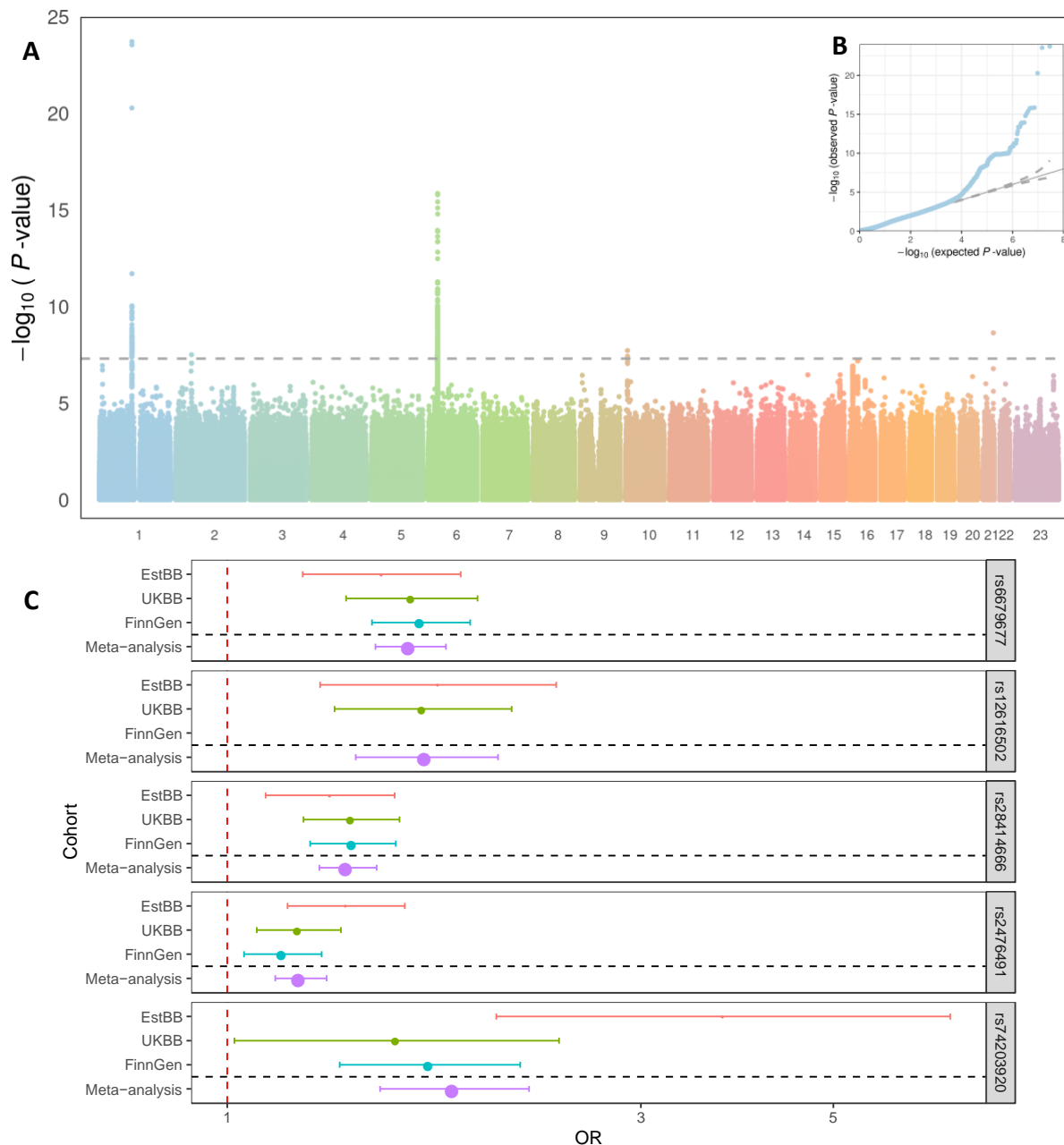
B12 deficiency anemia due to intrinsic factor deficiency, also known as pernicious anemia, is characterized by impaired B12 uptake caused by lack of intrinsic factor, a substance produced by epithelial cells of the stomach lining. Intrinsic factor normally binds B12 and facilitates absorption in the intestinal tract. Pernicious anemia is often caused by autoimmune damage to the stomach lining (autoimmune gastritis) in which case the gastric epithelial lining is damaged or destroyed<sup>1</sup>. The prevalence of pernicious anemia is around 0.1% in populations of European ancestry; however, it is more common in older people (~2% in >60-year-olds), and believed to be less prevalent in Asian populations<sup>2</sup>. Symptoms of pernicious anemia range from fatigue to

megaloblastic anemia and neurological abnormalities (peripheral numbness, paresthesia, ataxia) in more serious cases<sup>3</sup>.

Pernicious anemia is a complex disease with familial clustering, with a clear autoimmune basis and higher incidence of other autoimmune diseases, such as autoimmune thyroid conditions<sup>4</sup>, vitiligo<sup>5</sup>, type 1 diabetes<sup>6</sup> in both patients with pernicious anemia and their relatives<sup>7</sup>. Although studies focusing on HLA serotypes have been conducted for pernicious anemia, the results have been conflicting<sup>7</sup> and currently there is no clear consensus on the HLA alleles or other genetic risk factors predisposing to pernicious anemia.

We conducted a genome-wide association study (GWAS) of vitamin B12 deficiency due to lack of intrinsic factor to evaluate the contribution of genetic variation to the etiology of this disease in a combined dataset of 2,166 cases and 659,516 controls from three large population based biobanks - Estonian Biobank (EstBB)<sup>8</sup>, UK Biobank (UKBB)<sup>9</sup>, and FinnGen study. In the EstBB, individuals with pernicious anemia were identified using the ICD-10 code D51.0, resulting in 378 cases and 138,207 controls for analysis. Association testing was carried out with SAIGE 0.38 software<sup>10</sup>, adjusting for sex, year of birth and 10 PCs. Analysis in the EstBB was carried out under ethical approval 1.1-12/624 from the Estonian Committee on Bioethics and Human Research and data release N05 from the EstBB. GWAS summary statistics for the UKBB analysis including White British participants were downloaded from the UKBB PheWeb (<http://pheweb.sph.umich.edu/SAIGE-UKB/about>). Similarly, cases had been identified using the ICD-10 code D51.0 (754 cases, 390,026 controls) and SAIGE had been used for association testing, adjusting for sex, birth year and the first 4 PCs. Summary statistics for FinnGen study were obtained from the publicly available R3 release PheWeb ([https://www.finnngen.fi/en/access\\_results](https://www.finnngen.fi/en/access_results)). In the FinnGen study, we used the endpoint 'Vitamin B12 deficiency anemia', which included all the subcodes in the ICD10 D51 diagnosis group (1,034 cases, 131,283 controls). Similarly to other cohorts, SAIGE had been used for association testing, adjusting for sex, age, 10 PCs and genotyping batch. For meta-analysis, we used fixed-effects meta-analysis implemented in GWAMA<sup>11</sup>. Additional details available in **Methods** section.

In our GWAS meta-analysis we identify five genome-wide significant ( $p < 5 \times 10^{-8}$ ) associations (Table 1, Figure 1, Supplementary Figure 1) on 1p13.2 (lead signal rs6679677,  $p = 1.91 \times 10^{-24}$ ), 2p16.1 (rs12616502,  $p = 3.14 \times 10^{-8}$ ), 6p21.32 (rs28414666,  $p = 1.40 \times 10^{-16}$ ), 10p15.1 (rs2476491,  $p = 1.90 \times 10^{-8}$ ), and 21q22.3 (rs74203920,  $p = 2.33 \times 10^{-9}$ ), with similar effect estimates (effect heterogeneity was measured using Cochran's test, p-values ranging from 0.05 to 0.83) in all analyzed cohorts (Figure 1), except for the lead variant on chromosome 2, which is absent in the Finnish data.



**Figure 1. Results of the pernicious anemia GWAS meta-analysis.** A) Manhattan and B) QQ plot; C) Forest plot of effect estimates for lead variants associated with pernicious anemia. The odds ratios (dots) and 95% confidence intervals (error bars) are shown for all included cohorts and meta-analysis. The size of the dot is proportional to the effective sample size, (calculated as  $4/((1/N_{\text{cases}})+(1/N_{\text{controls}}))$ ). Pernicious anemia is defined as ICD10 code D51.0 in EstBB and UKBB and as D51 (vitamin B12 deficiency anemia) in FinnGen.

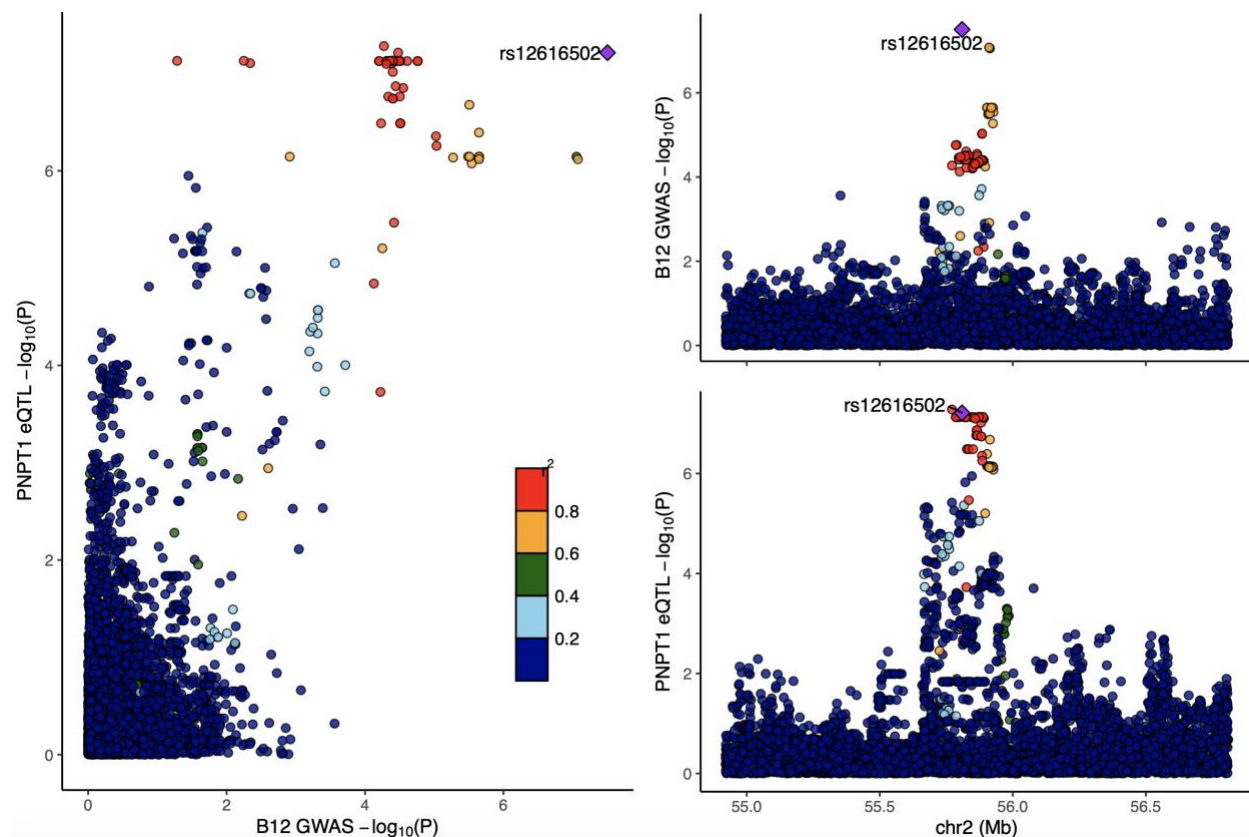
**Table 1. Summary of GWAS meta-analysis results for pernicious anemia**

Lead variant	chr:pos*	EA/NEA	EAF	p-value	OR (95% CI)
rs6679677	1:114303808	A/C	EstBB=0.140 UKBB=0.1 FinnGen=0.147	$1.91 \times 10^{-24}$	1.63 (1.48-1.79)
rs12616502	2:55809015	A/G	EstBB=0.055 UKBB=0.051 FinnGen=NA	$3.14 \times 10^{-8}$	1.70 (1.41-2.05)
rs28414666	6:32626451	G/A	EstBB=0.767 UKBB=0.790 FinnGen=0.684	$1.40 \times 10^{-16}$	1.38 (1.28-1.49)
rs2476491	10:6095410	A/T	EstBB=0.688 UKBB=0.700 FinnGen=0.755	$1.90 \times 10^{-8}$	1.22 (1.14-1.30)
rs74203920	21:45714294	T/C	EstBB=0.018 UKBB=0.015 FinnGen=0.038	$2.33 \times 10^{-9}$	1.83 (1.5-2.29)

\*Positions according to GRCh37

The lead SNP on chr1, rs6679677, is in high LD ( $r^2=0.96$ ) with a non-synonymous variant in exon 12 of the *PTPN22* gene (rs2476601,  $p=2.82 \times 10^{-24}$ ). *PTPN22* is a known immune regulator gene and this particular variant has been associated with several autoimmune diseases, including rheumatoid arthritis, Crohn's disease, systemic lupus erythematosus, vitiligo, autoimmune thyroid conditions, type 1 diabetes and others (Supplementary Table 1 and 3, Supplementary Figure 2).

On chromosome 2, the lead signal rs12616502 is in high LD with exonic non-synonymous variants in *CCDC104* ( $r^2=0.94$ , rs1045920) and *PNPT1* ( $r^2=0.94$ , rs7594497), and this region has been previously associated with both vitiligo, hypothyroidism and myelodysplastic syndrome (Supplementary Tables 1 and 3, Supplementary Figure 2). Colocalization analysis demonstrated that pernicious anemia has a common causal variant with the expression of *PTPN1* in thyroid tissue in GTEx v8 dataset (Figure 2, Supplementary Table 2). *PTPN1* is further supported as a candidate causal gene in this locus by data from mouse knockouts, as *Pnpt1tm1a(KOMP)Wtsi/Pnpt1*+ mice exhibit increased mean corpuscular volume (MCV) together with increased mean corpuscular hemoglobin (MCH)<sup>12</sup>. Increased red blood cell MCV is a common feature in macrocytic anemias, both megaloblastic (caused by B12 deficiency and pernicious anemia) and nonmegaloblastic (caused by diseases such as myelodysplastic syndrome and hypothyroidism)<sup>13</sup>.



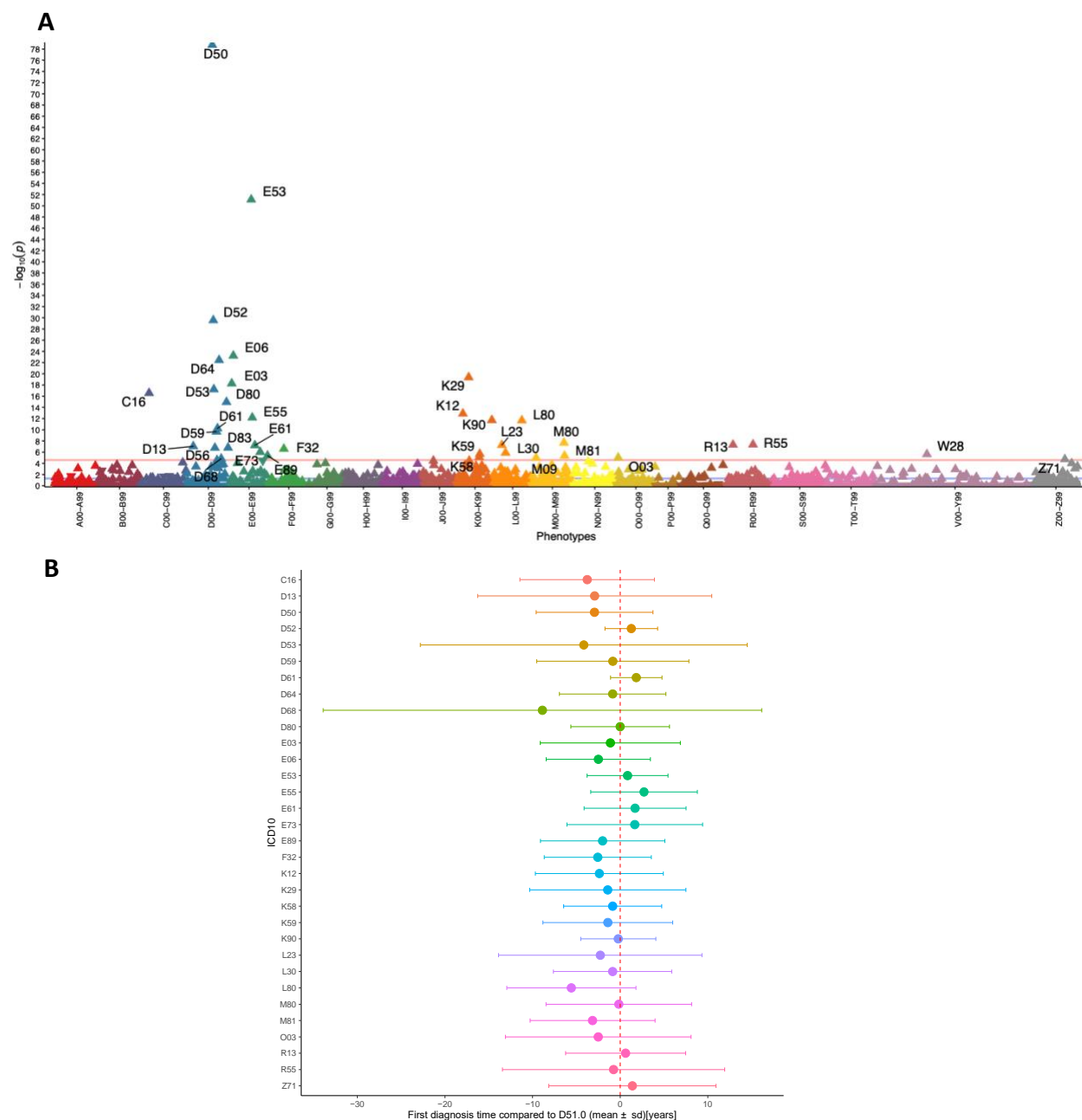
**Figure 2. GWAS signature for pernicious anemia co-localizes with the eQTL signal for *PNPT1* in thyroid tissue in GTEx v8 dataset (posterior probability for shared causal variant PP4=0.87).** Left – comparison on  $-\log_{10}$  p-values from pernicious anemia GWAS-meta-analysis (x-axis) and eQTL analysis(y-axis). Right – locus plots for the *PNPT1* locus from GWAS meta-analysis (top) and eQTL analysis (bottom). For all plots, LD is colored with respect to the GWAS lead signal rs12616502 (labeled).

The third signal on chromosome 6 is located in the HLA region, a common hub for autoimmune condition associations, downstream *HLA-DQB1*. To further clarify the association signal in the HLA region, we used data on a total of 249 imputed HLA alleles at four-digit level available for 183 cases and 67,505 controls in the EstBB. Three tested alleles passed the Bonferroni corrected threshold of association ( $0.05/249=2 \times 10^{-4}$ ), including HLA-DRB1\*15:01 ( $p=1.3 \times 10^{-6}$ ), HLA-DQB1\*06:02 ( $p=1.3 \times 10^{-6}$ ) and HLA-DQA1\*01:02 ( $p=2.8 \times 10^{-5}$ ). The DRB1\*15:01-DQB1\*06:02-DQA1\*01:02 combination forms the HLA-DR15 haplotype, which is a reported risk factor for multiple sclerosis<sup>14</sup>. HLA-DR15 belongs under the HLA-DR2 group, which has been associated with pernicious anemia in a 1981 study<sup>15</sup>.

On chromosome 10, the sentinel variant rs2476491 is intronic to *IL2RA*. *IL2RA* encodes the interleukin-2 receptor alpha chain, thus being involved in regulating regulatory T-cells and immune tolerance, as regulatory T cells suppress autoreactive T-cells. Accordingly, this locus has previously been associated with multiple sclerosis, juvenile idiopathic arthritis, vitiligo and hypothyroidism (Supplementary Table 1).

Finally, the association on chromosome 21, rs74203920, is a missense variant in the *AIRE* gene, a known autoimmune regulator. Interestingly, mutations in *AIRE* are a known cause of autoimmune polyendocrinopathy syndrome type 1 (APS-1), which is a rare autosomal recessive syndrome, that sometimes includes pernicious anemia among other components<sup>16</sup>.

We used the individual level data in EstBB to evaluate the association between pernicious anemia and other diseases (defined by ICD-10 codes). According to our analysis, individuals with pernicious anemia have more diagnoses of other anemias and vitamin deficiencies (Figure 3, Supplementary Table 4), thyroid problems (thyroiditis and hypothyroidism), gastrointestinal tract diagnoses (gastritis, malignant neoplasm of stomach, intestinal malabsorption, irritable bowel syndrome), but also of vitiligo, dermatitis, osteoporosis and spontaneous abortion. Evaluation of disease timeline shows that while there is considerable individual variation, on average, thyroid issues (ICD10 E03 other hypothyroidism and E06 thyroiditis, Figure 3), intestinal problems (K12 stomatitis, K29 gastritis and K58 irritable bowel syndrome), spontaneous abortion (O03), depression (F32) and vitiligo (L80) are diagnosed before pernicious anemia, while nutritional deficiencies are diagnosed after (D52 folate deficiency, E53 deficiency of other B vitamins, E55 vitamin D deficiency, E61 deficiency of other nutrient elements). Majority of these diagnoses reflect the etiology of pernicious anemia (gastritis) or symptoms (skin problems, syncope and collapse, depression, stomatitis), known comorbidities (vitiligo, thyroid issues<sup>4,5</sup>), or diseases where pernicious anemia is a known risk factor (such as osteoporosis<sup>17</sup> and stomach cancer<sup>18</sup>). The association with spontaneous abortion is interesting, as although there is some evidence B12 deficiency and pernicious anemia could cause recurrent miscarriage<sup>19–21</sup>, the data is scarce and the link with spontaneous miscarriage has not been explored in depth.



**Figure 3. Association plot for ICD10 codes associated with pernicious anemia diagnosis in Estonian Biobank.** A) Each triangle in the plot represents one ICD10 main code and the direction of the triangle represents direction of effect – upward-pointing triangles show increased prevalence of diagnosis code in pernicious anemia cases. Red line – Bonferroni-corrected significance level ( $2.5 \times 10^{-5}$ ). B) Time at first diagnosis of associated diagnosis code compared to time of pernicious anemia (D51.0) diagnosis. Plot shows mean difference in years, error bars correspond to standard deviation. Red dashed vertical line represents D51.0 diagnosis.



This GWAS meta-analysis focused on individuals with pernicious anemia identified from population-based biobanks, thus more detailed analysis on the effect these genetic risk factors have on disease severity or subphenotypes is unfortunately not possible. Although the used summary statistics for FinnGen cohort are for a broader vitamin B12 deficiency phenotype definition compared to other two cohorts, we see no significant differences in effect estimates for the reported lead variants. The analyses included participants of European ancestry, however, analyses in other populations are warranted, since the prevalence of pernicious anemia differs depending on racial background, being more common in people with African or European ancestry, especially from Scandinavia and UK<sup>2</sup>. The analysis of associated phenotypes could potentially provide clinically useful insights into pernicious anemia disease trajectories and offer information for patient management; however, at the moment it only includes participants of the EstBB, with limited follow-up time, therefore further studies are needed.

In summary, our analysis of 2,166 cases and 659,516 controls identifies robust risk loci for pernicious anemia in or near candidate genes with a known role in autoimmune conditions (*PTPN22*, *HLA*, *IL2RA*, *AIRE*). We further narrow the signal in the HLA region to the HLA-DR15 haplotype and propose *PNPT1* as the most likely causal gene in the 2p16.1 locus. The associations between the identified loci and other autoimmune conditions, such as type 1 diabetes, vitiligo, and autoimmune thyroid conditions help to clarify the link between pernicious anemia and its common comorbidities. Analysis of associated diagnoses and disease trajectories confirm the association between pernicious anemia and thyroid issues, vitiligo, gastritis, stomach cancer, osteoporosis and other diagnoses, but also between pernicious anemia and spontaneous abortion.

## Methods

### Cohorts

#### *Estonian Biobank*

The Estonian Biobank is a population-based biobank with over 200,000 participants. The 150K data freeze was used for the analyses described in this paper. All biobank participants have signed a broad informed consent form. Individuals with pernicious anemia were identified using the ICD-10 code D51.0, and all biobank participants who did not have this diagnosis were considered as controls. Information on ICD codes is obtained via regular linking with the national Health Insurance Fund and other relevant databases<sup>8</sup>.



All EstBB participants have been genotyped at the Core Genotyping Lab of the Institute of Genomics, University of Tartu, using Illumina GSAv1.0, GSAv2.0, and GSAv2.0\_EST arrays. Samples were genotyped and PLINK format files were created using Illumina GenomeStudio v2.0.4. Individuals were excluded from the analysis if their call-rate was < 95% or if sex defined based on heterozygosity of X chromosome did not match sex in phenotype data. Before imputation, variants were filtered by call-rate < 95%, HWE p-value < 1e-4 (autosomal variants only), and minor allele frequency < 1%. Variant positions were updated to b37 and all variants were changed to be from TOP strand using GSAMD-24v1-0\_20011747\_A1-b37.strand.RefAlt.zip files from <https://www.well.ox.ac.uk/~wrayner/strand/> webpage. Prephasing was done using Eagle v2.3 software<sup>22</sup> (number of conditioning haplotypes Eagle2 uses when phasing each sample was set to: --Kpbwt=20000) and imputation was done using Beagle v.28Sep18.793<sup>23</sup> with effective population size ne=20,000. Population specific imputation reference of 2297 WGS samples was used.

Association analysis was carried out using SAIGE (v0.38)<sup>10</sup> software implementing mixed logistic regression model with LOCO option, using sex, year of birth and 10 PCs as covariates in step I.

### *UK Biobank*

The UK Biobank (UKBB) is a prospective cohort of 502,637 individuals aged 37-73 recruited in 2006-2010 from across the UK, who completed detailed questionnaires regarding socio-demographic and lifestyle characteristics and their medical history and had a clinical assessment. Additional information about medical conditions (both existing at baseline and occurring during follow-up) has been obtained through linking with hospital admission and mortality data. Full details of the study have been reported in Sudlow et al<sup>9</sup>. Publicly available GWAS summary statistics downloaded from the UKBB PheWeb [<http://pheweb.sph.umich.edu/SAIGE-UKB/about>] were used for the analysis. Briefly, the PheWeb includes GWAS summary statistics for ICD code-based traits extracted from electronic health records. Phenotypes have been classified into 1,403 broad PheWAS codes, including pernicious anemia (PheCode 281.11), defined using the ICD-10 code D51.0 and excluding other anemias under the PheCodes 280-285.99. Genetic analyses have been carried out using SAIGE<sup>10</sup>.

### *FinnGen*

FinnGen is a public-private partnership project combining data from Finnish biobanks and electronic health records from different registries. After a one-year embargo, the FinnGen

summary stats are available for download. In this study, we used the results from the FinnGen release R3, which includes data from 135,638 individuals and more than 1,800 disease endpoints. FinnGen individuals have been genotyped with Illumina and Affymetrix arrays and imputed to the population-specific SISu v3 imputation reference panel. Genetic association testing has been carried out with SAIGE<sup>10</sup>. The FinnGen disease endpoint “Vitamin B12 deficiency anemia” included all individuals with the ICD10 D51 diagnosis as cases. For more information on genotype data, disease endpoints and GWAS analyses, please see <https://finngen.gitbook.io/documentation/>.

### **GWAS meta-analysis**

We extracted all genetic variants with a rs-number from the summary statistics of the three participating cohorts and conducted an inverse of variance weighted fixed-effects meta-analysis without genomic control using GWAMA<sup>11</sup>. A total of 30,907,385 variants were included in the meta-analysis. Genome-wide significance was set to  $p < 5 \times 10^{-8}$ .

### **HLA allele imputation in the EstBB**

Imputation of HLA alleles from SNP data was carried out at the Broad Institute using the SNP2HLA tool<sup>24</sup>.

### **Colocalization**

We conducted colocalization analyses to detect shared causal variants between pernicious anemia and gene expression using COLOC (v.3.2.1) R package<sup>25</sup> and GWAS meta-analysis summary statistics. We set the prior probabilities to  $p_1=1 \times 10^{-4}$ ,  $p_2=1 \times 10^{-4}$ ,  $p_{12}=5 \times 10^{-6}$  as suggested by Wallace (2020)<sup>26</sup> and used the COLOC version which takes regression coefficients and their variance into account.

In the analysis we compared our significant GWAS loci to all eQTL Catalogue (<https://www.ebi.ac.uk/eqtl/>) RNA-seq datasets (excluding Lepik et al. 2017<sup>27</sup> due to sample overlap) containing QTLs for gene expression, exon expression, transcript usage and txrevise event usage; eQTL Catalogue microarray datasets containing QTLs for gene expression (excluding Kasela et al 2017<sup>28</sup>); and GTEx v8 datasets containing QTLs for gene expression (see Methods: <https://www.ebi.ac.uk/eqtl/Methods/>).

We lifted the GWAS summary statistics over to hg38 build to match the eQTL Catalogue. For each genome-wide significant ( $p < 5 \times 10^{-8}$ ) GWAS locus we extracted the 1Mbp radius of its top hit from QTL datasets and ran the colocalization analysis for those eQTL Catalogue traits that

had at least one cis-QTL within this region with  $p < 1 \times 10^{-6}$ . We considered two signals to colocate if the posterior probability for a shared causal variant was 0.8 or higher. All results with a PP4 > 0.8 can be found in Supplementary Table 2.

Results were visualized with the *LocusCompareR* library<sup>30</sup>.

### Look-up of phenome-wide associations in GWAS catalog and with PhenoScanner v2

FUMA v1.3.6a<sup>29</sup> was used to compare the genome-wide significant lead signals and markers in high LD with these markers against the results in the GWAS catalogue. The results of this look-up are presented in Supplementary Table 1.

PhenoScanner v2<sup>31,32</sup> was used for look-up of phenotype associations for the GWAS lead variants in previous GWAS studies. PhenoScanner query was done using the rsid-s of GWAS lead variants and the *phenoscanner* R package (<https://github.com/phenoscanner/phenoscanner>). Query results were filtered to keep one association per variant per trait, keeping studies from newer or larger studies. Descriptions of Experimental Factor Ontology (EFO) terms and classification of EFO broad categories were obtained from the GWAS Catalog. Missing categories were added by manually searching the EMBL-EBI EFO webpage ([www.ebi.ac.uk/efo/](http://www.ebi.ac.uk/efo/)). For visualization of PhenoScanner results, parent categories with fewer results were grouped into larger categories and a heatmap was created using the *pheatmap* library in R 3.6.1. and a modified script from ([https://github.com/LappalainenLab/spiromics-covid19-eqtl/blob/master/eqtl/summary\\_phenoscanner\\_lookup.Rmd](https://github.com/LappalainenLab/spiromics-covid19-eqtl/blob/master/eqtl/summary_phenoscanner_lookup.Rmd)). The results of this look-up are presented in Supplementary Table 3.

### Mouse phenotypes

We used the Mouse Genome Database<sup>12</sup> (<http://www.informatics.jax.org>) to evaluate the *PNPT1* effect on phenotype in mouse models.

### Analysis of associated phenotypes in EstBB

Using the individual level data in the EstBB, we conducted an analysis to find ICD10 diagnosis codes associated with the D51.0 diagnosis. We tested the association between pernicious anemia status (defined as ICD10 D51.0) and other ICD10 codes using logistic regression and adjusting for sex, age and 10 PCs. Bonferroni correction was applied to select statistically significant associations (Number of tested ICD main codes – 1,944, corrected p-value threshold –  $2.5 \times 10^{-5}$ ).

<sup>5</sup>). Results were visualised using the *PheWas* library (<https://github.com/PheWAS/PheWAS>). All analyses were carried out in R 3.6.1. Diagnosis timeline of pernicious anemia and significantly associated diagnoses was evaluated by calculating the time interval (in years) between the first date of pernicious anemia diagnosis (D51.0) and first date of associated disease for each individual. Results were visualized by plotting the means and standard deviations of calculated intervals using R 3.6.1. The results of this analysis are presented in Supplementary Table 4.

### **Data availability**

Used UKBB and FinnGen summary statistics can be browsed and downloaded from UKBB PheWeb and FinnGen PheWeb, respectively. Full meta-analysis summary statistics will be made available upon publication.

### **Web resources**

UKBB PheWeb: <http://pheweb.sph.umich.edu/SAIGE-UKB/about>; FinnGen Freeze 3 PheWeb: <http://r3.finnngen.fi>; Mouse Genome Database: <http://www.informatics.jax.org>; GTEx Portal: <https://gtexportal.org/home/>; eQTL Catalogue: <https://www.ebi.ac.uk/eql/>

### **Declaration of Interests**

The authors declare no competing interests.

### **Acknowledgements**

TL is supported by the Estonian Research Council grant MOBTP155. ML and RM are supported by the Estonian Research Council grant PRG687. Computations were performed in the High Performance Computing Center, University of Tartu. We want to acknowledge the participants and investigators of the FinnGen and UKBB studies. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 10/01/20.

### **References**

1. Lenti, M.V., Rugge, M., Lahner, E., Miceli, E., Toh, B.H., Genta, R.M., De Block, C., Hershko, C., and Di Sabatino, A. (2020). Autoimmune gastritis. *Nat. Rev. Dis. Prim.* 6, 1–19.

2. Stabler, S.P., and Allen, R.H. (2004). Vitamin B12 deficiency as a worldwide problem. *Annu. Rev. Nutr.* 24, 299–326.
3. Bizzaro, N., and Antico, A. (2014). Diagnosis and classification of pernicious anemia. *Autoimmun. Rev.* 13, 565–568.
4. Ness-Abramof, R., Nabriski, D.A., Braverman, L.E., Shilo, L., Weiss, E., Reshef, T., Shapiro, M.S., and Shenkman, L. (2006). Prevalence and evaluation of B12 deficiency in patients with autoimmune thyroid disease. *Am. J. Med. Sci.* 332, 119–122.
5. Dahir, A.M., and Thomsen, S.F. (2018). Comorbidities in vitiligo: comprehensive review. *Int. J. Dermatol.* 57, 1157–1164.
6. Kahaly, G.J., and Hansen, M.P. (2016). Type 1 diabetes associated autoimmunity. *Autoimmun. Rev.* 15, 644–648.
7. Banka, S., Ryan, K., Thomson, W., and Newman, W.G. (2011). Pernicious anemia - Genetic insights. *Autoimmun. Rev.* 10, 455–459.
8. Leitsalu, L., Haller, T., Esko, T., Tammesoo, M.-L., Alavere, H., Snieder, H., Perola, M., Ng, P.C., Mägi, R., Milani, L., et al. (2015). Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. *Int. J. Epidemiol.* 44, 1137–1147.
9. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., et al. (2015). UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med.* 12, e1001779.
10. Zhou, W., Nielsen, J.B., Fritsche, L.G., Dey, R., Gabrielsen, M.E., Wolford, B.N., LeFaive, J., VandeHaar, P., Gagliano, S.A., Gifford, A., et al. (2018). Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat. Genet.* 50, 1335–1341.
11. Mägi, R., and Morris, A.P. (2010). GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* 11, 288.
12. Bult, C.J., Blake, J.A., Smith, C.L., Kadin, J.A., Richardson, J.E., Anagnostopoulos, A., Asabor, R., Baldarelli, R.M., Beal, J.S., Bello, S.M., et al. (2019). Mouse Genome Database (MGD) 2019. *Nucleic Acids Res.* 47, D801–D806.
13. Nagao, T., and Hirokawa, M. (2017). Diagnosis and treatment of macrocytic anemias in adults. *J. Gen. Fam. Med.* 18, 200–204.
14. Schmidt, H., Williamson, D., and Ashley-Koch, A. (2007). HLA-DR15 haplotype and multiple sclerosis: A HuGE review. *Am. J. Epidemiol.* 165, 1097–1109.
15. Ungar, B., Mathews, J.D., Tait, B.D., and Cowling, D.C. (1981). HLA-DR patterns in

pernicious anaemia. *Br. Med. J. (Clin. Res. Ed)*. 282, 768–770.

16. Bruserud, Ø., Oftedal, B.E., Wolff, A.B., and Husebye, E.S. (2016). AIRE-mutations and autoimmune disease. *Curr. Opin. Immunol.* 43, 8–15.
17. Goerss, J.B., Kim, C.H., Atkinson, E.J., Eastell, R., O’Fallon, W.M., and Melton, L.J. (1992). Risk of fractures in patients with pernicious anemia. *J. Bone Miner. Res.* 7, 573–579.
18. Ye, W., and Nyrén, O. (2003). Risk of cancers of the oesophagus and stomach by histology or subsite patients hospitalised for pernicious anaemia. *Gut* 52, 938–941.
19. Bennett, M. (2001). Vitamin B 12 deficiency, infertility and recurrent fetal loss. *J. Reprod. Med. Obstet. Gynecol.* 46, 209–212.
20. Reznikoff-Etiévant, M.F., Zittoun, J., Vaylet, C., Pernet, P., and Milliez, J. (2002). Low Vitamin B 12 level as a risk factor for very early recurrent abortion. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 104, 156–159.
21. Hübner, U., Alwan, A., Jouma, M., Tabbaa, M., Schorr, H., and Herrmann, W. (2008). Low serum vitamin B12 is associated with recurrent pregnancy loss in Syrian women. *Clin. Chem. Lab. Med.* 46, 1265–1269.
22. Loh, P.R., Danecek, P., Palamara, P.F., Fuchsberger, C., Reshef, Y.A., Finucane, H.K., Schoenherr, S., Forer, L., McCarthy, S., Abecasis, G.R., et al. (2016). Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* 48, 1443–1448.
23. Browning, S.R., and Browning, B.L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* 81, 1084–1097.
24. Jia, X., Han, B., Onengut-Gumuscu, S., Chen, W.M., Concannon, P.J., Rich, S.S., Raychaudhuri, S., and de Bakker, P.I.W. (2013). Imputing Amino Acid Polymorphisms in Human Leukocyte Antigens. *PLoS One* 8,.
25. Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace, C., and Plagnol, V. (2014). Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS Genet.* 10, e1004383.
26. Wallace, C. (2020). Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. *PLoS Genet.* 16,.
27. Lepik, K., Annilo, T., Kukuškina, V., Kisand, K., Kutalik, Z., Peterson, P., and Peterson, H. (2017). C-reactive protein upregulates the whole blood expression of CD59 - an integrative analysis. *PLoS Comput. Biol.* 13,.
28. Kasela, S., Kisand, K., Tserel, L., Kaleviste, E., Remm, A., Fischer, K., Esko, T., Westra, H.J., Fairfax, B.P., Makino, S., et al. (2017). Pathogenic implications for autoimmune

mechanisms derived by comparative eQTL analysis of CD4 + versus CD8 + T cells. *PLoS Genet.* 13,.

29. Watanabe, K., Taskesen, E., van Bochoven, A., and Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* 8, 1826.

30. Liu, B., Gloudemans, M.J., Rao, A.S., Ingelsson, E., and Montgomery, S.B. (2019). Abundant associations with gene expression complicate GWAS follow-up. *Nat. Genet.* 51, 768–769.

31. Staley, J.R., Blackshaw, J., Kamat, M.A., Ellis, S., Surendran, P., Sun, B.B., Paul, D.S., Freitag, D., Burgess, S., Danesh, J., et al. (2016). PhenoScanner: A database of human genotype-phenotype associations. *Bioinformatics* 32, 3207–3209.

32. Kamat, M.A., Blackshaw, J.A., Young, R., Surendran, P., Burgess, S., Danesh, J., Butterworth, A.S., Staley, J.R., and Kelso, J. (2019). PhenoScanner V2: An expanded tool for searching human genotype-phenotype associations. *Bioinformatics* 35, 4851–4853.