Title
Evidence in the UK Biobank for the underdiagnosis of erythropoietic protoporphyria

Journal
Genetics in Medicine, 23(1)

ISSN
1098-3600

Authors
Dickey, Amy K
Quick, Corbin
Ducamp, Sarah
et al.

Publication Date
2021

DOI
10.1038/s41436-020-00951-8

Peer reviewed

HHS Public Access

Author manuscript

Genet Med. Author manuscript; available in PMC 2021 March 02.

Published in final edited form as:

Genet Med. 2021 January ; 23(1): 140–148. doi:10.1038/s41436-020-00951-8.

Evidence in the UK Biobank for the underdiagnosis of erythropoietic protoporphyria

Amy K. Dickey, MD[1,2], **Corbin Quick, PhD**[3], **Sarah Ducamp, PhD**[2,4], **Zhaozhong Zhu, ScD**[6], **Yen-Chen A. Feng, PhD**[5,6,7], **Hetanshi Naik, PhD**[8], **Manisha Balwani, MD**[8], **Karl E. Anderson, MD**[9], **Xihong Lin, PhD**[6], **John E. Phillips, PhD**[10], **Lina Rebeiz**[1,14], **Herbert L. Bonkovsky, MD**[11], **Brendan M. McGuire, MD**[12], **Bruce Wang, MD**[13], **Daniel I. Chasman, PhD**[2,14], **Jordan W. Smoller, MD ScD**[2,5,6,7], **Mark D. Fleming, MD DPhil**[2,4,†], **David C. Christiani, MD**[1,2,6,†]

[1]Deparment of Medicine, Division of Pulmonary and Critical Care Medicine, Massachusetts General Hospital, Boston, MA.

[2]Harvard Medical School, Boston, MA.

[3]Department of Biostatistics, Harvard School of Public Health, Boston, MA.

[4]Department of Pathology, Boston Children's Hospital, Boston, MA.

[5]Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA.

[6]Department of Epidemiology, Harvard School of Public Health, Boston, MA.

[7]Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA.

[8]Department of Genetics and Genomic Sciences, Mount Sinai Hospital, New York City, NY.

[9]Division of Gastroenterology, The University of Texas Medical Branch at Galveston, Galveston, TX.

[10]Hematology Department, University of Utah School of Medicine, Salt Lake City, UT.

[11]Department of Medicine, Section on Gastroenterology & Hepatology, Wake Forest Baptist Health, Winston-Salem, NC.

[12]Division of Gastroenterology and Hepatology, University of Alabama at Birmingham, Birmingham, AL.

[13]Department of Medicine, Division of Gastroenterology, University of California San Francisco, San Francisco, CA.

Corresponding Author: Amy Dickey, 206-681-0218, adickey@mgh.harvard.edu.
[†]These authors contributed equally to this work

Conflicts of Interest

AKD and LR report serving as a consultant for Alnylam Pharmaceuticals. KEA reports grants and personal fees from Alnylam, Recordati Rare Diseases, and Mitsubishi-Tanabe Pharma. HLB reports serving as a consultant for Alnylam, Recordati Rare Diseases, Mitsubishi Tanabe North America, and Clinuvel and receiving grant support from Anlylam, Gilead Sciences, and Mitsubishi Tanaba Pharma. JEP reports serving as a consultant for Alnylam, Recordati, Mitsubishi Tanabe Pharma, and Agios. MDF is on the Scientific Advisory Board of Disc Medicine. All of the competing interests listed above are outside/unrelated to this work. CQ, HN, SD, ZZZ, YCAF, MJW, XHL, DIC, JWS, and DCC have no conflicts of interest to report.

[14]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA.

Abstract

Purpose: Erythropoietic protoporphyria (EPP), characterized by painful cutaneous photosensitivity, results from pathogenic variants in ferrochelatase (FECH). For 96% of patients, EPP results from coinheriting a rare pathogenic variant in *trans* of a common hypomorphic variant c.315–48T>C (minor allele frequency 0.05). The estimated prevalence of EPP derived from the number of diagnosed individuals in Europe is 0.00092%, but this may be conservative due to underdiagnosis. No study has estimated EPP prevalence using large genetic datasets.

Methods: Disease-associated *FECH* variants were identified in the UK Biobank, a dataset of 500,953 individuals including 49,960 exome sequences. EPP prevalence was then estimated. The association of *FECH* variants with EPP-related traits was assessed.

Results: Analysis of pathogenic *FECH* variants in the UK Biobank provides evidence that EPP prevalence is 0.0059% (95% CI: 0.0042%−0.0076%), 1.7–3.0 times more common than previously thought in the UK. In homozygotes for the common c.315–48T>C *FECH* variant, there was a novel decrement in both erythrocyte mean corpuscular volume (MCV) and hemoglobin.

Conclusion: The prevalence of EPP has been underestimated secondary to underdiagnosis. The common c.315–48T>C allele is associated with both MCV and hemoglobin, an association that could be important both for those with and without EPP.

Keywords

Erythropoietic protoporphyria (EPP); ferrochelatase (FECH); prevalence; mean corpuscular volume (MCV); anemia

Introduction

Erythropoietic protoporphyria (EPP [MIM: 177000]) results from pathogenic variants in ferrochelatase (FECH), the last enzyme in the heme biosynthetic pathway, causing accumulation of the light-sensitive molecule protoporphyrin IX (PPIX) in erythrocytes and secondarily in the plasma and biliary system.[1] Although ~4% of EPP patients have two rare pathogenic *FECH* variants, the classical molecular alteration present in ~96% of EPP patients is a rare pathogenic *FECH* variant in *trans* of a common intronic *FECH* variant c.315–48T>C (rs2272783, historically called IVS3–48T>C), which is known to increase the use of an aberrant splice site. This combination decreases FECH enzymatic activity to ~35%.[2] Although light tolerance varies among individuals, EPP is fully penetrant after considering the contribution of the hypomorphic c.315–48T>C variant.[2–4] Between 2–10% of patients with EPP symptoms instead have X-linked protoporphyria (XLPP [MIM: 300752]), in which the EPP phenotype and biochemical changes result from gain-of-function variants in delta-aminolevulinate synthase-2 (*ALAS2*), the first enzyme of erythroid heme biosynthesis.[5–8] Clinically, patients with all forms of protoporphyria experience severe life-long painful cutaneous photosensitivity.[1] While patients may also experience iron deficiency anemia and gallstones, the most life-threatening complication is rapidly

progressive cholestatic liver failure, which is often fatal unless a liver transplant is performed.[1,9]

Based on the incidence of EPP in Europe, the prevalence of EPP was estimated to be 0.00092%.[10] The calculated prevalence ranged from zero in Poland to 0.00277% in Norway.[10] The prevalence in the UK was estimated to be 0.00254%.[10] It is unclear to what extent these differences represent actual disparities in disease prevalence or simply differences in the likelihood of reaching a diagnosis in the respective countries. As patients generally experience a delay in diagnosis of more than a decade, the rate of EPP diagnosis may not be an adequate representation of the true disease incidence and prevalence.[10–13] No study to date has estimated the prevalence of EPP using large genetic datasets.

A better understanding of EPP prevalence, if truly underdiagnosed, could encourage efforts to decrease the barriers that patients currently face prior to diagnosis and to develop novel therapeutics. Because a new treatment for EPP called afamelanotide has recently been FDA-approved, these efforts are increasingly critical in order to bring this therapy to individuals who could benefit from it.[14,15]

Materials and Methods

Ethics Statement

All clinical data was de-identified, requiring only data use agreements. This analysis obtained exemption from the Partners Healthcare IRB and was performed in accordance with relevant guidelines and regulations.

Pathogenic *FECH* variants in the UK Biobank

The UK Biobank is a dataset of 500,953 individuals aged 40–69, including both clinical and genetic data.[16] Genotypes were assessed for all 500,953 participants using the UK Biobank Affymetrix SNP array (Table 1). All of the variants in the array were directly genotyped and not determined by imputation. Exome sequencing was performed on 49,960 of the participants, but evidence for large deletions was not assessed. The functional equivalent (FE) exome format was analyzed in this study, and the *FECH* gene was not within the UK Biobank regions that experienced FE exome data quality concerns.

The American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) recommendations were applied for the determination of variant pathogenicity, as described in the supplemental methods.[17–19] The estimation of the genetic prevalence of EPP was performed as described in the supplemental methods, and code availability is addressed in the supplemental methods.

The Porphyrias Consortium Dataset

The Porphyrias Consortium is a consortium of six university sites in the US studying porphyria and additional satellite sites active in porphyria research.[20] The Consortium's Longitudinal Study of the Porphyrias (clinicaltrials.gov Identifier: NCT01561157) collects genetic and clinical data on patients at these sites. The *FECH* gene and *ALAS2* exon 11 were sequenced in all of the EPP patients, as previously described.[7] Notably, all XLPP-

causing *ALAS2* pathogenic variants are in exon 11.[5] Gene dosage analysis to evaluate for large deletions was performed on many but not all of the participants.

The Partners Biobank

The Partners Biobank is an ongoing research project that collects clinical, biochemical, and genetic data from patients seen at Massachusetts General Hospital, Brigham and Women's Hospital, and other Partners Institutions. It currently contains clinical and biochemical data from 100,818 patients and genetic data in the form of a SNP array for 36,422 patients.[16]

The Women's Genome Health Study

The Women's Genome Health Study (WGHS) is a dataset of genomic and biochemical data evaluating >25,000 healthy American women, with follow-up for over 23 years for major health events.[21] Variants in *FECH* were identified in a subset of 22,618 WGHS participants with both European ancestry and genotype data available from the Exome Chip v1.1.

Clinical and biochemical associations with *FECH* variants in the UK Biobank

For the evaluation of associations between pathogenic *FECH* variants and EPP-related traits, the variant collapsing method to construct gene-based tests of association was used, as described in the supplemental methods.[22]

Statistics

To test for categorical effects, values in each *FECH* variant group were compared to those with no *FECH* variants by t-test in a linear regression model. The linear effect for 0, 1, and 2 c.315–48T>C variants was also determined by linear regression. Meta-analysis was performed by inverse variance weighting (IVW). To evaluate for associations with hemoglobin <12.5 g/dL and with various diagnoses, logistic regression models were fit using glm in R v3.5.3. All analyses were adjusted for the top 10 genetic principal components. P-values for each of the comparisons are listed in the supplemental table 3. Confidence intervals for allele frequencies and the genetic prevalence of EPP were calculated using non-parametric bootstrapping. Further details are provided in the supplemental methods.

Results

Frequency of pathogenic FECH variants

Among the 49,960 exome sequences available in the UK Biobank, 54 individuals had *FECH* variants that are pathogenic for EPP when in *trans* of the c.315–48T>C variant, with pathogenicity determined according to the American College of Medical Genetics (ACMG) criteria (Table 1, Table 2, Supplemental Table 1).[17] Furthermore, among the 38,841 unrelated individuals of European ancestry, 0.118% (95% CI: 0.0849%−0.152%) of individuals (46 total) had one variant that is pathogenic for EPP when in *trans* of c.315–48T>C. The allele frequency of c.315–48T>C in this population was 0.0439. Consequently, the estimated genetic prevalence of EPP in individuals of European descent in the UK was 0.0052% (95% CI: 0.0036%, 0.0068%; see Supplemental Methods).[4] There were no

statistically significant differences between the allele frequencies in the UK Biobank, gnomAD, the Partners Biobank, and the WGHS datasets, suggesting that the UK Biobank is not an outlier in terms of the frequencies of pathogenic *FECH* variants.

Calculating EPP prevalence based on exomic variant frequency alone would lead to an underestimate of EPP prevalence for two reasons. First, it has been described that a significant portion of pathogenic FECH variants found in *trans* of c.315–48T>C are large deletions (10% in the UK EPP dataset), which are not assessed in the UK biobank and difficult to detect (Figure 1).[8,23] Secondly, the analysis is predicated on having observed any particular variant in a well-characterized EPP patient; however, 4.1% of known EPP patients in the Porphyrias Consortium dataset and a published UK EPP dataset have private missense or splice regions variants not found in other published EPP patient datasets. In the absence of an EPP diagnosis, those variants would not have been identified as pathogenic. Notably, the UK Biobank contains a number of *FECH* variants of uncertain significance, including a total of 37 missense variants in 143 individuals that are rare (minor allele frequency < 0.0001) and predicted to be deleterious by SIFT and probably damaging by PolyPhen. These analytical biases were corrected for as described in the supplemental methods, a correction achieved by comparing the frequencies of variant types in the UK Biobank, the Porphyrias Consortium dataset, and in a previously published dataset of EPP patients in the UK (Figure 1).[8] Including this estimate of unidentified pathogenic *FECH* variants, the calculated EPP genetic prevalence increased to 0.0059% (95% CI: 0.0042%−0.0076%), which suggests that EPP may be 2.3 (95% CI:1.7–3.0) times more common than previously thought in the UK (detailed calculations provided in supplemental methods).

This estimate may still be an underestimate because it is only the estimate for EPP patients with a rare pathogenic variant in *trans* of c.315–48T>C and does not account for the possibility of two rare variants causing EPP, which is described in 2–4% of cases.[7,8] The UK Biobank contains nine variants that have previously been found in at least one EPP patient, but only in *trans* of another rare allele. Only three of these met criteria for pathogenic/likely pathogenic according to the ACMG criteria (Supplemental Table 1, 2).[24] Additional enzymatic data for the six other variants would not change the EPP prevalence determination in this study because if pathogenic, it is still not clear if these variants can cause EPP in *trans* of c.315–48T>C or if all combinations of these rare variants can cause EPP. Notably, there were no *ALAS2* variants pathogenic for XLPP in the dataset, nor were there *ALAS2* exon 11 truncation variants that would be expected to cause XLPP.

Given the small number of rare pathogenic variants in the dataset, only the phase of the p.(Pro334Leu) and c.315–48T>C variants in the exome sequences could be computationally predicted, identifying a *trans* orientation in two out of three individuals possessing both of these variants. Therefore, these two individuals are expected to have EPP. This is the same number of individuals predicted to have EPP in the exome data (Table 2).

In the UK Biobank, only 3 individuals carried an ICD10 code diagnosis of “hereditary erythropoietic porphyria,” which includes EPP, XLPP, and congenital erythropoietic porphyria (CEP [MIM: 263700]) (Table 3). If all have EPP, the estimated prevalence of EPP among unrelated individuals of European descent in the UK Biobank is 7.5 (95% CI: 5.3–

9.6) times greater than that of diagnosed EPP in the same population (0.00026%). CEP is extremely rare with an estimated prevalence of <0.0001%; consequently, a diagnosis of CEP is far less likely than a diagnosis of EPP or XLPP.[25] While none of these three participants had identified pathogenic *FECH* or *ALAS2* variants other than c.315–48T>C in FECH, few pathogenic loci were assessed as none of these individuals received exome sequencing.

The hypomorphic c.315–48T>C allele distribution among major ethnicities in the UK Biobank

Similar to previously published data, those of Chinese descent in the UK Biobank have the highest prevalence of c.315–48T>C, with 42.1% heterozygous and 8.8% homozygous individuals (Figure 2).[2] However, in those with British ancestry, the percentages of c.315–48T>C heterozygotes and homozygotes are 8.3% and 0.2%, respectively. After stratification for ethnicity, the occurrence of c.315–48T>C was found to comply with Hardy-Weinberg Equilibrium (data not shown). Due to the limited sample sizes available for non-European ethnic groups in the UK Biobank, there was insufficient information to assess for population differences in the frequencies of rare pathogenic *FECH* variants. However, differences in c.315–48T>C were significant among ethnicities (P<2.2e-16).

Clinical and biochemical associations with *FECH* variants in the UK Biobank

Among unrelated individuals of European ancestry in the UK Biobank, the c.315–48T>C variant was significantly associated with decreased erythrocyte mean corpuscular volume (MCV) (P=3.84e-5, Table 3). When the MCVs of those heterozygous or homozygous for c.315–48T>C were compared to those with no *FECH* variants separately, each comparison revealed a statistically significant decrease in MCV (P=0.0005, P=0.009, respectively, Table 3). In those who were compound heterozygotes for c.315–48T>C and a rare pathogenic *FECH* variant, MCV was significantly decreased as compared to those with no *FECH* variants, despite there being only 13 compound heterozygotes among individuals of European ancestry (P=0.009, Table 3). However, none of these individuals had diagnosis of EPP or an ICD-10 diagnosis code associated with photosensitivity. Additional computational corrections for variants in the SNP array that are associated with iron deficiency, namely rs4820268 in transmembrane protease serine 6 (TMPRSS6) and rs3811647 in transferrin (TF), did not affect these associations (data not shown).[26] Common *FECH* variants that are genetically linked with c.315–48T>C, including IVS1–23C>T, were also analyzed for associations with MCV, and none of these associations were significant after correcting for c.315–48T>C (data not shown).

Regarding other biochemical and clinical features, hemoglobin was lower in homozygotes for c.315–48T>C as compared to those with no *FECH* variants (P= 0.017, Table 3). Furthermore, in c.315–48T>C homozygotes there was a 1.4-fold increased prevalence of individuals with hemoglobin <12.5 g/dL (P=0.003). There was no statistically significant association between the presence of a *FECH* variant and the diagnosis of anemia, gallstones, or liver disease (Table 3). There was no significant association between c.315–48T>C and MCV or between c.315–48T>C and hemoglobin in the much smaller Partners Biobank dataset, a dataset of individuals presenting for clinical care within the Partners Healthcare System (Table 3). The number of unrelated individuals of European descent in the Partners

Biobank was 6.8% that of the UK Biobank (25,696 versus 379,390). Additionally, the hemoglobin of individual of European descent in the Partners Biobank was lower than in the UK Biobank dataset of healthy volunteers ($p< 2.2e-16$). In a second subset of 63,670 unrelated individuals of European descent in the UK Biobank, a group who were excluded from the first subset of 379,390 individuals due to genetic relatedness, there was again a significant association between MCV and c.315–48T>C (Table 3).

A meta-analysis of the Partners Biobank and two unrelated ethnic populations within the UK Biobank confirmed a significant linear association between MCV and c.315–48T>C (Table 3). In the meta-analysis, homozygotes for c.315–48T>C similarly had significantly lower MCV and hemoglobin and higher prevalence of hemoglobin < 12.5 mg/dL versus those with no *FECH* variants (Table 3). The second subset of unrelated individuals of European descent in the UK Biobank could not be included in this meta-analysis, as all of the individuals in the second subset were genetically related to those in the first.

Discussion

The frequency of pathogenic *FECH* variants in the UK Biobank provides evidence that EPP is underdiagnosed. Based on allele frequencies of c.315–48T>C and rare pathogenic *FECH* variants in the UK Biobank, the estimated genetic prevalence of EPP in individuals of European ancestry in the UK is 0.0052% (95% CI: 0.0036%, 0.0068%), and this estimate increases to 0.0059% (95% CI: 0.0042%−0.0076%) when correcting for pathogenic variants that are not reported in the UK Biobank exomic sequences and for variants that cannot be identified as pathogenic due to insufficient clinical data. This calculated estimate is 2.3 (95% CI:1.7–3.0) times higher than the estimate in the UK, which is based on the rate of diagnosis of EPP.[10] Furthermore, the estimated genetic prevalence of EPP among unrelated individuals of European descent in the UK Biobank is 7.5 (95% CI: 5.3–9.6) times higher than the prevalence of diagnosed EPP in the same population (0.00026%). The significant decrement in MCV in the 13 compound heterozygotes identified in Table 3, which was not present in individuals with one rare pathogenic *FECH* variant, suggests that some of these individuals likely have undiagnosed EPP, or possibly subclinical EPP. In addition to shedding light on the prevalence of EPP, this study also demonstrates a novel association between the c.315–48C>T variant and both MCV and hemoglobin.

This study's calculation of EPP prevalence is predicated on complete penetrance of the disorder, after accounting for the role of c.315–48T>C. Decreased penetrance in EPP cannot be excluded, but is probably rare, if present at all, as it has not previously been reported. Among more than 155 family cohorts of EPP patients that have been published in the literature, no occurrence of a non-penetrant disease-associated genotype has been reported.[2,27,28] Only one case of subclinical EPP has been described in the literature, in an individual with p.(Pro334Leu) in *trans* of c.315–48T>C. This individual had anemia and elevated PPIX, but no reported photosensitivity (Supplemental Table 1).[29] Although none of the 13 individuals with both a pathogenic variant and c.315–48T>C in this study had a diagnosis code associated with photosensitivity, it may be expected that undiagnosed adult EPP patients would no longer seek medical attention for their symptoms, as even diagnosed EPP patients report negative experiences telling physicians about their symptoms.[13] Furthermore,

because there may be no visible skin changes despite severe pain, physicians may not diagnose it as a form of photosensitivity.[30] In EPP, there are likely environmental or genetic factors affecting a patient's degree of photosensitivity that have yet to be identified, which could also influence the likelihood of diagnosis. Future studies to better understand the variable expressivity in EPP could pave the way for new treatments.

This study provides evidence that EPP is underdiagnosed, which should encourage efforts to decrease the many barriers that EPP patients face in their attempt to find a diagnosis. Firstly, because few physicians know about EPP and because of the minimal visible skin changes in some patients despite severe pain, there exists a large barrier to a physician's consideration of EPP in the differential diagnosis, even among specialists.[30] This could be remedied through solutions such as routine genomic sequencing and electronic medical record clinician decision support tools, as well as increased physician awareness of the disease. A new treatment for EPP called afamelanotide was approved in Europe a few years ago and recently approved by the US FDA; consequently, efforts to identify patients is increasingly important in order to bring this and future new therapies to individuals who could benefit from them.[14,15] A second barrier in the US is in the laboratory diagnosis of EPP, which must include measurement of erythrocyte total and metal-free protoporphyrin. Some major commercial labs, such as Quest and LabCorp, provide a "free erythrocyte protoporphyrin" test, which is actually a zinc protoporphyrin test.[31] Because zinc protoporphyrin is often normal in EPP, the diagnosis may be missed.[2,31] Because EPP has been considered rare, there has to date been insufficient motivation to address this problem.

Only two studies have described a role for the hypomorphic c.315–48T>C allele in patients outside of classical EPP, studies that described an incomplete EPP phenotype in four individuals homozygous for this allele.[32,33] However, the presence of a deletion or cryptic intronic variant was not excluded; cryptic intronic variants were recently discovered in four individuals homozygous for c.315–48T>C and with the EPP phenotype.[27] Interestingly, based on the computationally-assessed pattern of growth of c.315–48T>C in Asian populations over time, there is evidence for positive selection for the variant in this population.[2] In other genetic diseases, such as sickle cell anemia [MIM: 603903], a protective effect of pathogenic variants has been demonstrated for particular clinical outcomes, accounting for the persistence of the variants in the population, so a similar selective advantage of c.315–48T>C is possible, perhaps related to anemia.[34] Notably, *in vitro* studies revealed that red cells from EPP patients were resistant to malaria, an effect that was secondary to FECH deficiency and not PPIX accumulation, as this resistance to malaria was not present in red cells from individuals with XLPP.[34]

This study is the first to demonstrate a significant association between the c.315–48T>C allele and either MCV or hemoglobin. A decrement in MCV and hemoglobin was not reproduced in individuals with one rare disease-causing allele, likely due to the small sample size. In the Partners Biobank, no significant association was detected between MCV and c.315–48T>C, possibly again an effect of dataset size and decreased statistical power. Population differences between the two datasets could provide another explanation, with the Partner's Biobank participants selecting for preferentially unhealthy subjects who have lower MCV and hemoglobin, as well as having more inter-patient variability in these

hematological characteristics due to a variety of medical conditions. Nonetheless, the collective evaluation of these datasets by a meta-analysis strengthens the evidence for an association between c.315–48T>C and both MCV and hemoglobin (Table 3).

The mechanism of reduced MCV and hemoglobin in individuals with c.315–48T>C is uncertain. There is a poorly defined clinical association of EPP with iron deficiency, which is seen in 20–50% of patients despite the appropriate regulation of both hepcidin and iron absorption.[1,35,36] This iron deficiency is unexpected because although both iron and metal-free PPIX are substrates of FECH, only metal-free PPIX accumulates. Unfortunately, iron deficiency could not be tested in this study because serum iron parameters were not measured in the UK Biobank cohort. Apart from systemic iron deficiency, a possible explanation for reduced MCV is through a decrease in Mitoferrin-1 (MFRN1), which transports iron into the mitochondria for use by FECH. One study has reported a strong correlation between FECH activity and the expression of the FECH-complexed mitochondrial protein Mitoferrin-1 (MFRN1) among individuals with EPP, XLPP, and healthy controls; consequently, decreased MFRN1 could have a role in the observed association between c.315–48T>C and MCV.[37,38] Alternatively, it is possible that the variant causes microcytosis through another mechanism, such as a slight reduction in heme synthesis, which could result in decreased hemoglobin synthesis through heme-dependent regulatory pathways such as heme-regulated elongation initiation factor kinase 1A (EIFK1A, also known as heme-regulated inhibitor, HRI). Future *in vitro* studies should investigate and confirm the role of FECH deficiency in microcytosis mechanistically. The association between c.315–48T>C and both hemoglobin and MCV suggests that slight FECH deficiency predisposes to anemia without engendering an overt EPP phenotype. The extent to which FECH, along with other clinical or genetic factors, may predispose to clinically important anemia outside of EPP should be the topic of future investigations.

This study has several limitations. Because erythrocyte metal-free protoporphyrin levels could not be performed to conclude whether or not participants have EPP, decreased penetrance cannot be excluded. Regarding the nine variants that have been observed in EPP in *trans* of another rare variant but not c.315–48T>C, further clinical studies to ascertain which combinations of these variants can result in EPP may provide a better understanding of EPP prevalence. In addition, the correction for large deletions, likely cryptic variants, and private missense variants are predictions, and thus subject to error. Furthermore, this study had limited power to detect associations with EPP-related traits due to the small number of rare pathogenic *FECH* variants. Regarding the association between c.315–48T>C and both MCV and hemoglobin, ideally this would have been corroborated using another dataset, and preferably one measuring iron. However, there were no other datasets available of a similar size for comparison that include both the genetic and biochemical data of relatively healthy individuals. An effect of unknown confounders or variants genetically linked with c.315–48T>C is possible. Additionally, because the UK Biobank evaluates disproportionately healthy participants in the UK, our observations may not apply equally to clinical features in other populations.

A strength of this study is that it evaluates for the first time the prevalence of pathogenic *FECH* variants in a large exomic dataset and uses this data to estimate of EPP prevalence.

Because of the evidence for EPP underdiagnosis found in this study, increased efforts to decrease barriers to diagnosis are essential, especially now that an effective therapy has been FDA-approved. As this study primarily analyzed individuals of European descent in the UK, further research is needed to evaluate EPP prevalence and underdiagnosis in other ethnicities.[39] Furthermore, this is the first study to suggest a role for c.315–48T>C in erythrocytes outside of classical EPP, which should be further evaluated. EPP is a life-altering condition that limits quality of life with risk of hepatic complications that can be fatal. Although a new therapy has been developed for EPP, which represents a significant advancement in the field, the clinical impact of any treatment will be muted if a large percentage of EPP patients remain undiagnosed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

I would like to thank Heidi Rehm, PhD for her assistance with *FECH* variant interpretation and Meg Min-Jung Wang, PhD for her assistance with accessing the UK Biobank. This research has been conducted using the UK Biobank Resource under Application Number 47222.[16] AKD is supported in part by Massachusetts General Hospital's Healthcare Transformation Laboratory. The Porphyrias Consortium (U54 DK083909) is a part of the NCATS RDCRN, which is funded through a collaboration between NCATS and the NIDDK. The WGHS is supported by the NHLBI (HL043851 and HL080467) and the NCI (CA047988 and UM1CA182913), with funding for genotyping provided by Amgen.

References:

1. Balwani M, Naik H, Anderson KE, et al. Clinical, biochemical, and genetic characterization of north American patients with erythropoietic protoporphyria and x-linked protoporphyria. JAMA Dermatology. 2017;153(8):789–796. doi:10.1001/jamadermatol.2017.1557 [PubMed: 28614581]
2. Gouya L, Martin-Schmitt C, Robreau AM, et al. Contribution of a common single-nucleotide polymorphism to the genetic predisposition for erythropoietic protoporphyria. Am J Hum Genet. 2006;78(1):2–14. doi:10.1086/498620 [PubMed: 16385445]
3. Gouya L, Puy H, Robreau AM, et al. The penetrance of dominant erythropoietic protoporphyria is modulated by expression of wildtype FECH. Nat Genet. 2002;30(1):27–28. doi:10.1038/ng809 [PubMed: 11753383]
4. Balwani M, Bloomer J, Desnick R, Porphyrias Consortium of the NIH-Sponsored Rare Diseases Clinical Research Network. Erythropoietic Protoporphyria, Autosomal Recessive.; 1993.
5. Ducamp S, Schneider-Yin X, De Rooij F, et al. Molecular and functional analysis of the C-terminal region of human erythroid-specific 5-aminolevulinic synthase associated with X-linked dominant protoporphyria (XLDPP). Hum Mol Genet. 2013;22(7):1280–1288. doi:10.1093/hmg/dds531 [PubMed: 23263862]
6. Whatley SD, Ducamp S, Gouya L, et al. C-terminal deletions in the ALAS2 gene lead to gain of function and cause X-linked dominant protoporphyria without anemia or iron overload. Am J Hum Genet. 2008;83(3):408–414. doi:10.1016/j.ajhg.2008.08.003 [PubMed: 18760763]
7. Balwani M, Doheny D, Bishop DF, et al. Loss-of-Function Ferrochelatase and Gain-of-Function Erythroid-Specific 5-Aminolevulinate Synthase Mutations Causing Erythropoietic Protoporphyria and X-Linked Protoporphyria in North American Patients Reveal Novel Mutations and a High Prevalence of X-Lin. Mol Med. 2013;19(1):26–29. doi:10.2119/molmed.2012.00340 [PubMed: 23364466]
8. Whatley SD, Mason NG, Holme SA, Anstey AV, Elder GH, Badminton MN. Molecular epidemiology of erythropoietic protoporphyria in the U.K. Br J Dermatol. 2010;162(3):642–646. doi:10.1111/j.1365-2133.2010.09631.x [PubMed: 20105171]

9. Anstey AV, Hift RJ. Liver disease in erythropoietic protoporphyria: Insights and implications for management. Gut. 2007;56(7):1009–1018. doi:10.1136/gut.2006.097576 [PubMed: 17360790]
10. Elder G, Harper P, Badminton M, Sandberg S, Deybach JC. The incidence of inherited porphyrias in Europe. J Inherit Metab Dis. 2013;36(5):849–857. doi:10.1007/s10545-012-9544-4 [PubMed: 23114748]
11. Frank J, Poblete-Gutiérrez P. Delayed diagnosis and diminished quality of life in erythropoietic protoporphyria: results of a cross-sectional study in Sweden. J Intern Med. 2011;269(3):270–274. doi:10.1111/j.1365-2796.2010.02283.x [PubMed: 21332583]
12. Lala SM, Naik H, Balwani M. Diagnostic Delay in Erythropoietic Protoporphyria. J Pediatr. 2018;202:320–323.e2. doi:10.1016/j.jpeds.2018.06.001 [PubMed: 30041937]
13. Pitfalls Dickey A. and proposed solutions for patient communication about erythropoietic protoporphyria: A survey of parents and adult patients. J Am Acad Dermatol. 2019;81(5):1204–1207. doi:10.1016/j.jaad.2019.04.010 [PubMed: 30978424]
14. Langendonk JG, Balwani M, Anderson KE, et al. Afamelanotide for Erythropoietic Protoporphyria. N Engl J Med. 2015;373(1):48–59. doi:10.1056/NEJMoa1411481 [PubMed: 26132941]
15. Biolcati G, Marchesini E, Sorge F, Barbieri L, Schneider-Yin X, Minder EI. Long-term observational study of afamelanotide in 115 patients with erythropoietic protoporphyria. Br J Dermatol. 2015;172(6):1601–1612. doi:10.1111/bjd.13598 [PubMed: 25494545]
16. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562(7726):203–209. doi:10.1038/s41586-018-0579-z [PubMed: 30305743]
17. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405–424. doi:10.1038/gim.2015.30 [PubMed: 25741868]
18. Abou Tayoun AN, Pesaran T, DiStefano MT, et al. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. Hum Mutat. 2018;39(11):1517–1524. doi:10.1002/humu.23626 [PubMed: 30192042]
19. ClinGen Sequence Variant Interpretation Recommendation for in Trans Criterion (PM3)-Version 1.0 Working Group Page: Https://Clinicalgenome.Org/Working-Groups/Sequence-Variant-Interpretation/ SVI Recommendation for in Trans Criterion (PM3)-Version 1.0.; 2019 https://clinicalgenome.org/working-groups/sequence-variant-interpretation/. Accessed January 9, 2020.
20. Bonkovsky HL, Maddukuri VC, Yazici C, et al. Acute porphyrias in the USA: features of 108 subjects from porphyrias consortium. Am J Med. 2014;127(12):1233–1241. doi:10.1016/j.amjmed.2014.06.036 [PubMed: 25016127]
21. Ridker PM, Chasman DI, Zee RYL, et al. Rationale, design, and methodology of the Women's Genome Health Study: A genome-wide association study of more than 25 000 initially healthy American women. Clin Chem. 2008;54(2):249–255. doi:10.1373/clinchem.2007.099366 [PubMed: 18070814]
22. Li B, Leal SM. Methods for Detecting Associations with Rare Variants for Common Diseases: Application to Analysis of Sequence Data. Am J Hum Genet. 2008;83(3):311–321. doi:10.1016/j.ajhg.2008.06.024 [PubMed: 18691683]
23. Whatley SD, Mason NG, Holme SA, Anstey AV, Elder GH, Badminton MN. Gene Dosage Analysis Identifies Large Deletions of the FECH Gene in 10% of Families with Erythropoietic Protoporphyria. J Invest Dermatol. 2007;127:2790–2794. doi:10.1038/sj.jid.5700924 [PubMed: 17597821]
24. Elder GH, Gouya L, Whatley SD, Puy H, Badminton MN, Deybach J-C. The molecular genetics of erythropoietic protoporphyria. Cell Mol Biol (Noisy-le-grand). 2009;55(2):118–126. http://www.ncbi.nlm.nih.gov/pubmed/19656460. Accessed October 18, 2019. [PubMed: 19656460]
25. Karim Z, Lyoumi S, Nicolas G, Deybach JC, Gouya L, Puy H. Porphyrias: A 2015 update. Clin Res Hepatol Gastroenterol. 2015;39(4):412–425. doi:10.1016/j.clinre.2015.05.009 [PubMed: 26142871]

26. Benyamin B, Ferreira MAR, Willemsen G, et al. Common variants in TMPRSS6 are associated with iron status and erythrocyte volume. Nat Genet. 2009;41(11):1173–1175. doi:10.1038/ng.456 [PubMed: 19820699]
27. Chiara M, Primon I, Tarantini L, et al. Targeted resequencing of FECH locus reveals that a novel deep intronic pathogenic variant and eQTLs may cause erythropoietic protoporphyria (EPP) through a methylation-dependent mechanism. Genet Med. 2020;22(1):35–43. doi:10.1038/s41436-019-0584-0 [PubMed: 31273344]
28. Parker M, Corrigall AV., Hift RJ, Meissner PN. Molecular characterization of erythropoietic protoporphyria in South Africa. Br J Dermatol. 2008;159(1):182–191. doi:10.1111/j.1365-2133.2008.08580.x [PubMed: 18460026]
29. Berroeta L, Man I, Goudie DR, Whatley SD, Elder GH, Ibbotson SH. Late presentation of erythropoietic protoporphyria: case report and genetic analysis of family members. Br J Dermatol. 2007;157(5):1030–1031. doi:10.1111/j.1365-2133.2007.08117.x [PubMed: 17711525]
30. Lecluse ALY, Kuck-Koot VCM, van Weelden H, et al. Erythropoietic protoporphyria without skin symptoms-you do not always see what they feel. Eur J Pediatr. 2008;167(6):703–706. doi:10.1007/s00431-007-0557-1 [PubMed: 17710435]
31. Gou EW, Balwani M, Bissell DM, et al. Pitfalls in erythrocyte protoporphyrin measurement for diagnosis and monitoring of protoporphyrias. Clin Chem. 2015;61(12):1453–1456. doi:10.1373/clinchem.2015.245456 [PubMed: 26482161]
32. Mizawa M, Makino T, Nakano H, Sawamura D, Shimizu T. Incomplete erythropoietic protoporphyria caused by a splice site modulator homozygous IVS3–48C polymorphism in the ferrochelatase gene. Br J Dermatol. 2016;174(1):172–175. doi:10.1111/bjd.14078 [PubMed: 26280465]
33. Schneider-Yin X, Mamet R, Minder EI, Schoenfeld N. Biochemical and molecular diagnosis of erythropoietic protoporphyria in an Ashkenazi Jewish family. J Inherit Metab Dis. 2008;31(SUPPL. 2):S363–7. doi:10.1007/s10545-008-0924-8 [PubMed: 18758989]
34. Smith CM, Jerkovic A, Puy H, et al. Red cells from ferrochelatase-deficient erythropoietic protoporphyria patients are resistant to growth of malarial parasites. Blood. 2015;125(3):534–541. doi:10.1182/blood-2014-04-567149 [PubMed: 25414439]
35. Barman-Aksoezen J, Girelli D, Aurizi C, et al. Disturbed iron metabolism in erythropoietic protoporphyria and association of GDF15 and gender with disease severity. J Inherit Metab Dis. 2017;40(3):433–441. doi:10.1007/s10545-017-0017-7 [PubMed: 28185024]
36. Bossi K, Lee J, Schmeltzer P, et al. Homeostasis of iron and hepcidin in erythropoietic protoporphyria. Eur J Clin Invest. 2015;45(10):1032–1041. doi:10.1111/eci.12503 [PubMed: 26199063]
37. Phillips J, Farrell C, Wang Y, et al. Strong correlation of ferrochelatase enzymatic activity with Mitoferrin-1 mRNA in lymphoblasts of patients with protoporphyria. Mol Genet Metab. 2018;128:391–395. doi:10.1016/j.ymgme.2018.10.005 [PubMed: 30391163]
38. Paradkar PN, Zumbrennen KB, Paw BH, Ward DM, Kaplan J. Regulation of Mitochondrial Iron Import through Differential Turnover of Mitoferrin 1 and Mitoferrin 2. Mol Cell Biol. 2009;29(4):1007–1016. doi:10.1128/mcb.01685-08 [PubMed: 19075006]
39. Mizawa M, Makino T, Nakano H, Sawamura D, Shimizu T. Erythropoietic protoporphyria in a Japanese population. Acta Derm Venereol. 2019;99(7):634–639. doi:10.2340/00015555-3184 [PubMed: 30938825]

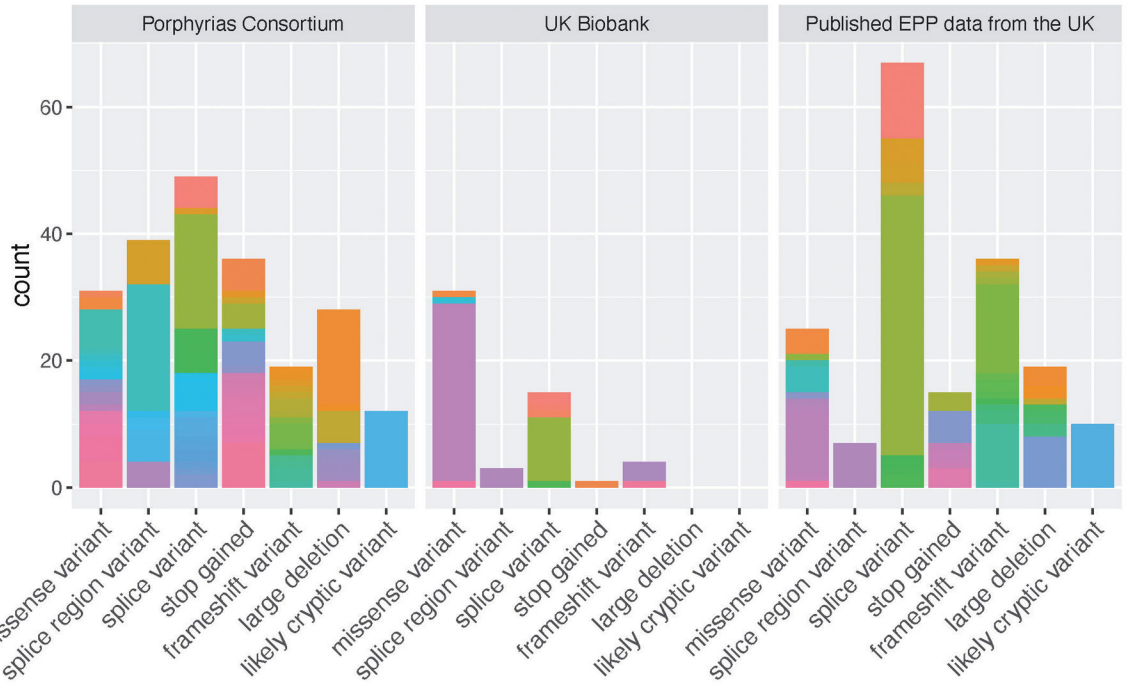
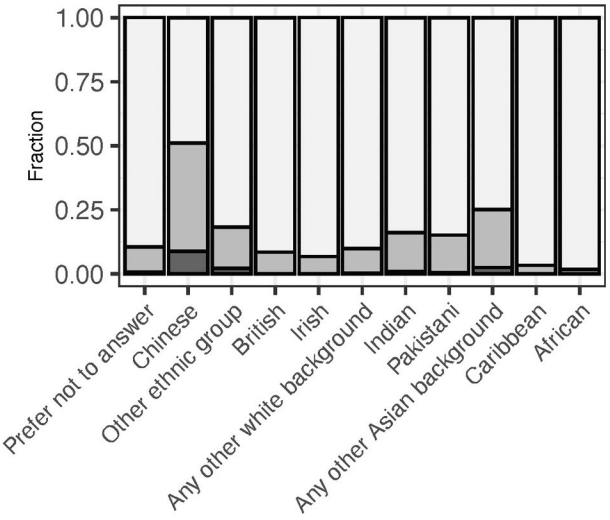


Figure 1. Pathogenic FECH variants in the UK Biobank and the Porphyrias Consortium dataset, alongside published EPP patient data from the UK.

The *FECH* variants in the UK Biobank exome data and the Porphyrias Consortium datasets are displayed alongside published EPP epidemiologic data from the UK. Variants are displayed according to their predicted consequence, regardless of whether or not a more important consequence has also been determined *in vitro*. Large deletions are not reported in the UK Biobank. Each color depicts a different variant. No variant is duplicated between the seven predicted consequences. Both the c.315–48T>C allele and variants that are only pathogenic when inherited with another rare variant have been excluded from the above figure.

The Porphyrias Consortium dataset subset possessing both c.315–48T>C and one rare variant, n=213; the same subset in the published UK EPP patient dataset, p=179; UK Biobank subset of those with 1 rare pathogenic variant, n=54.



Ethnicity	Total Count	c.315-48T>C MAF
Prefer not to answer	1,583	0.052
Chinese	1,504	0.255
Other ethnic group	4,356	0.091
British	431,038	0.043
Irish	12,756	0.034
Any other white background	15,818	0.049
Indian	5,716	0.080
Pakistani	1,748	0.076
Any other Asian background	1,747	0.126
Caribbean	4,297	0.017
African	3,204	0.009

Figure 2. The c.315–48T>C allele among major ethnicities in the UK Biobank.

The fractions of individuals with 0, 1, or 2 c.315–48T>C variants in the UK Biobank are depicted according to the major documented ethnicities. The numbers of individuals with each ethnicity are listed along with the minor allele frequency. Each ethnicity category with n>1000 was included. These ethnicities are listed exactly as they were in the UK Biobank dataset. MAF, minor allele frequency.

Table 1.

The frequency of pathogenic *FECH* variants that result in EPP when in *trans* of c.315–48T>C

The count and frequency of pathogenic *FECH* variants that result in EPP when in *trans* of c.315–48T>C were determined in participants in the exome sequences and in the SNP array of UK Biobank (UKB) dataset. These frequencies are displayed alongside the frequencies of these *FECH* variants in European individuals included in gnomAD, as well as in the SNP arrays of the Partners Biobank and the Women's Genome Health Study (WGHS) datasets.

Variant	European gnomAD MAF	Partners Biobank MAF	WGHS MAF	UKB MAF	UKB exomes variant count	UKB SNPs variant count
c.40delG p.(Ala14Profs*59)	2.3E-5	NA	NA	3.0E-5	3	NA
c.67+1G>T	0	NA	NA	1.0E-5	1	NA
c.314+2T>G	4.4E-5	1.4E-5	NA	1.0E-4	10	106
c.315–48T>C	4.9E-02	5.1E-02	NA	4.6E-02	4,463	43,418
c.343C>T p.(Arg115*)	0	NA	NA	1.0E-5	1	NA
c.418G>A p.(Gly140Arg)	8.8E-6	NA	NA	1.0E-5	1	NA
c.463+1G>C	0	NA	NA	2.0E-5	2	NA
c.479A>G p.(Tyr160Cys)	0	NA	NA	1.0E-5	1	NA
c.490C>T p.(Arg164Trp)	8.8E-6	NA	NA	1.0E-5	1	16
c.599–2A>G	6.5E-5	NA	NA	1.0E-5	1	NA
c.599–1G>A	0	NA	NA	1.0E-5	1	NA
c.1001C>T p.(Pro334Leu)	1.0E-4	1.4E-4	2.4E-4	2.8E-4	28	150
c.1077G>A	0	NA	NA	3.0E-5	3	NA
c.1115dup p.(Asn372Lysfs*8)	0	NA	NA	1.0E-5	1	NA
c.1137+3A>G	8.8E-6	NA	NA	0	0	7
Pathogenic variants that result in EPP when in *trans* of c.315–48T>C	2.6E-4	1.5E-4	2.4E-4	5.4E-4	54	279
Total Count	63,369	36,422	22,618	49,960	49,960	500,953

[a] MAF, minor allele frequency; SNP, single nucleotide polymorphism.

Table 2.

The prevalence of EPP and the frequency of pathogenic *FECH* variants among unrelated individuals of European descent.

Among unrelated individuals of European descent, the total count and frequency of identified pathogenic *FECH* variants that result in EPP when in *trans* of c.315–48T>C were determined in in the UK Biobank (UKB) dataset, as well as in the Partners Biobank and the Women's Genome Health Study (WGHS) datasets. Calculated estimates of additional pathogenic variants, including large deletions, likely cryptic variants (i.e. intronic variant or missed large deletions), and private missense or splice region variants, are provided. These estimates are based on the relative counts of these variants versus frameshift, splice, and nonsense variants in the Porphyrias Consortium dataset, as well as in a previously published EPP data from the UK. The calculated EPP prevalence based on the above frequencies is provided.

	Partners Biobank MAF	WGHS MAF	UKB exome MAF/ prevalence	UKB exomes count	UKB SNPs count
Pathogenic variants that result in EPP when in *trans* of c.315–48T>C	1.9E-4	2.4E-4	5.9E-4	46	222
Total calculated pathogenic[b] *FECH* variants			7.9E-5	6.15	
Calculated large deletions			4.5E-5	3.49	
Calculated likely cryptic variants			2.0E-5	1.57	
Calculated private missense/splice region variants			1.4E-5	1.09	
Calculated total pathogenic[b] variants			6.7E-4	52.15	
c.315–48T>C	4.1E-2	NA	4.4E-2	3371	33264
Observed count/prevalence of compound heterozygotes for a pathogenic[b] *FECH* variant and c.315–48T>C			1.0E-4	4	12
Count/prevalence of EPP not including calculate variants			5.2E-5	2.0	
Count/prevalence of EPP including calculated variants			5.9E-5	2.3	
Total count of unrelated individuals of European descent	25,696	22,618	38,841	38,841	379,387

[a] MAF, minor allele frequency; SNP, single nucleotide polymorphism.

[b] Pathogenic missense variants that have only been observed to cause EPP in *trans* of another rare variant and not c.315–48T>C were excluded from the pathogenic group in this analysis.

Table 3.

EPP-related clinical and laboratory findings according to *FECH* variant category.

Relevant EPP-related laboratory changes and associated diagnoses are listed according to the major categories of *FECH* variants. UK Biobank (UKB) and Partners Biobank analyses were restricted to unrelated individuals. Linear regression models (evaluating the linear effect for the rightmost column and categorical effects for the remaining columns) were used for mean corpuscular volume (MCV), hemoglobin (Hgb), and alanine aminotransferase (ALT); and logistic regression models were used for Hgb<12.5 status and for the diagnosis of anemia, gallstones, and liver disease. The top 10 genetic principal components were adjusted for in each regression model. Columns 1–6 correspond to mutually exclusive categories with no overlapping individuals. P-values for columns 2–6 were calculated using column 1 (No *FECH* Variants) as the reference group. Inverse variance weighted (IVW) meta-analysis of the top 3 datasets was performed.

		No FECH variants	1 c.315–48T>C variant	2 c.315–48T>C variants	Pathogenic[b] variant	Pathogenic[b] variant and c.315–48T>C	ICD10 Code Diagnosis of EPP	Linear Effect for c.315–48T>C
UK Biobank European descent subset 1	Count	346627	31788	737	222	13	3	
	MCV (fL)	91.13	91.04***	90.70**	90.87	87.92**	86.28	P=3.84e-5***
	Hgb (g/dL)	14.19	14.19	14.08*	14.14	14.01	13.02	P=0.18
	ALT (IU/L)	23.66	23.66	23.31	26.57	26.57	17.94	P=0.87
	Hgb <12.5 Prevalence	0.068	0.069	0.096**	0.072	0.077	0.333	
	Anemia Diagnosis Prevalence	0.024	0.025	0.031	0.041	0	0	
	Gallstones Diagnosis Prevalence	0.039	0.041	0.052	0.041	0.077	0	
	Liver Ds Diagnosis Prevalence	0.013	0.013	0.007	0.027	0	0	
UK Biobank South Asian	Count	5729	1034	61	1			
	MCV (fL)	89.29	89.13	88.62	92.52			P=0.30
	Hgb (g/dL)	14.21	14.16	14.17	14.84			P=0.33
	Hgb <12.5 Prevalence	0.161	0.173	0.180	0.000			
Partners Biobank	Count	23582	2042	39	9	1	23	
	MCV (fL)	90.24	90.29	89.24	91.17	87.99	89.67	P=0.97
	Hgb (g/dL)	12.92	12.93	12.85	13.58	9.84	12.72	P=0.85
	ALT (IU/L)	30.99	32.10	32.85	26.84	60.65	30.49	P=0.28
	Hgb <12.5 Prevalence	0.340	0.346	0.410	0.222	1.000	0.348	
Meta-analysis of above	MCV (fL)		P=0.00088***	P=0.0034**	P=0.46			P= 4.19e-5***
	Hgb (g/dL)		P=0.47	P=0.017*	P=0.68			P=0.16
	Hgb< 12.5 prevalence		P=0.44	P=0.00195 **	P=0.67			P= 0.98

		No FECH variants	1 c.315–48T>C variant	2 c.315–48T>C variants	Pathogenic[b] variant	Pathogenic[b] variant and c.315–48T>C	ICD10 Code Diagnosis of EPP	Linear Effect for c.315–48T>C
UK Biobank European 2	Count	58316	5198	112	44			
	MCV (fL)	91.054	90.869 **	91.154	90.853			p=0.0072**
	Hgb (g/dL)	14.126	14.109	14.087	13.820			p=0.34
	Hgb <12.5 Prevalence	0.072	0.072	0.080	0.091			
Porphyrias Consortium	Count					172		
	MCV (fL)					83.33		
	Hgb (g/dL)					13.04		
	ALT (IU/L)					32.71		
	Hgb <12.5 Prevalence					0.30		

a. UK Biobank unrelated individuals of European descent subset #1, n=379,390; UK Biobank unrelated individuals of South Asian descent, n = 6,825; Partners Biobank unrelated individuals of European descent, n=25,696; UK Biobank unrelated individuals of European descent subset #2, n= 63,670; Porphyrias Consortium participants with both one copy of c.315–48T>C and hematologic data, n=172.

b. Pathogenic missense variants that have only been observed to cause EPP in *trans* of another rare variant and not c.315–48T>C were excluded from the pathogenic group in this analysis.

c. * P<0.05, ** P< 0.01, ***P<0.001. All omitted p-values are listed in Supplemental Table 3.

d. MCV in femtoliter (SI unit fL, conventional unit μm^3, conversion factor 1), Hgb in gram per deciliter (conventional unit g/dL, SI unit g/L, conversion factor 10), and ALT in international unit per liter (conventional unit IU/L, SI unit ukat/L, conversion factor 0.0166).