HHS Public Access

Author manuscript

Mol Genet Metab. Author manuscript; available in PMC 2024 April 01.

Published in final edited form as:

Mol Genet Metab. 2023 April ; 138(4): 107542. doi:10.1016/j.ymgme.2023.107542.

Racial and ethnic diversity of classic and clinical variant galactosemia in the United States

Nichole M. Stettner[1], **David J. Cutler**[2], **Judith L. Fridovich-Keil**[2]

[1]Emory College of Arts and Sciences, Emory University, Atlanta, Georgia USA

[2]Department of Human Genetics, Emory University School of Medicine, Atlanta, Georgia USA

Abstract

Classic and clinical variant galactosemia (CG/CVG) are allelic, autosomal recessive disorders that result from deficiency of galactose-1-P uridylyltransferase (GALT). CG/CVG has been reported globally among patients of diverse ancestries, but most large studies of outcomes have included, almost exclusively, patients categorized as White or Caucasian. As a first step to explore whether the cohorts studied are representative of the CG/CVG population at large, we sought to define the racial and ethnic makeup of CG/CVG newborns in a diverse population with essentially universal newborn screening (NBS) for galactosemia: the United States (US). First, we estimated the predicted racial and ethnic distribution of CG/CVG by combining the reported demographics of US newborns from 2016–2018 with predicted homozygosity or compound heterozygosity of pathogenic, or likely pathogenic, *GALT* alleles from the relevant ancestral groups. Incorporating some simplifying assumptions, we predicted that of US newborns diagnosed with CG/CVG, 65% should be White (non-Hispanic), 23% should be Black (non-Hispanic), 10% should be Hispanic, and 2% should be Asian (non-Hispanic). Next, we calculated the observed racial and ethnic distribution of US newborns diagnosed with CG/CVG using available de-identified data from state NBS programs from 2016–2018. Of the 235 newborns in this cohort, 41 were categorized as other or unknown. Of the remaining 194, 66% were White (non-Hispanic or ethnicity unknown), 16% were Black (non-Hispanic or ethnicity unknown),15% were Hispanic, and 2% were Asian

Corresponding Author: Judith L. Fridovich-Keil, Department of Human Genetics, Emory University School of Medicine, Rm. 325.2 Whitehead Bldg., 615 Michael St., Atlanta, GA USA 30322, TEL 404-727-3924, jfridov@emory.edu.

Author contributions:

Nichole Stettner participated in study design, data analysis, and manuscript editing. David Cutler participated in study design, data gathering, data analysis, and manuscript editing. Judith Fridovich-Keil initiated the project and participated in study design, data gathering, data analysis, manuscript writing, and manuscript editing.

Conflict of Interest

Nichole Stettner, David Cutler, and Judith Fridovich-Keil all declare that they have no conflict of interest with this work.

Human Subjects

This work was carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans.

Informed consent was not required for this study, which did not include any individually identifiable human subject data or protected health information.

Animal Rights:

This article does not contain any studies with animal subjects performed by the any of the authors.

(non-Hispanic or ethnicity unknown). This observed distribution was statistically indistinguishable from the predicted distribution. To the limits of our study, these data confirm the racial and ethnic diversity of newborns with CG/CVG in the US, demonstrate an approach for estimating CG/CVG racial and ethnic diversity in other populations, and raise the troubling possibility that current understanding of long-term outcomes in CG/CVG may be skewed by ascertainment bias of the cohorts studied.

Keywords

classic galactosemia; clinical variant galactosemia; diversity; newborn screening

1. Introduction

Classic galactosemia (CG) and clinical variant galactosemia (CVG) are rare, autosomal recessive disorders that result from profound, or near-profound, deficiency of galactose-1-phosphate uridylyltransferase (GALT), respectively [1]. Patients with CG/CVG demonstrate impaired ability to metabolize galactose, leading infants to accumulate extremely high levels of galactose, galactose-1P, and galactitol in blood and tissues following dietary exposure to milk, which contains abundant galactose in the form of lactose. Even following dietary restriction of galactose, which is the current standard of care, most patients with CG/CVG demonstrate detectably elevated metabolites, ostensibly reflecting continued exposure to endogenously produced galactose [2–4]. Because infants with CG/CVG are typically born looking healthy but are at risk for rapidly progressing acute sequelae that can be deadly within days to weeks of exposure to breastmilk or dairy milk formula, many countries now include CG/CVG in their population newborn screen (NBS) [5].

Dietary restriction of galactose in the first days of life, often prompted by an out-of-range NBS result, prevents or resolves the potentially lethal acute sequelae of CG/CVG [1]. However, despite early detection and rigorous lifelong dietary intervention, more than half of infants with CG/CVG grow to experience a broad range of long-term complications by mid-childhood; these can include growth delay and speech, cognitive, and motor disabilities, among other problems [6, 7]. More than 80% of girls and young women also experience primary ovarian insufficiency (POI) [8]. While striking, these outcome prevalence numbers derive from studies of cohorts described, almost exclusively, as White or Caucasian [6–10], raising the question of whether they accurately reflect the broader population of patients.

Specifically, prior studies have reported the presence of CG/CVG in populations of European, African, and Asian ancestry, among others [11]}, and disease prevalence varies widely by ancestral group. For example, among the Romani people in Ireland, about 1/430 births are affected by CG [12]. Among the non-Romani people in Ireland, prevalence is about 1/34,000 [12]. In the Netherlands, prevalence is about 1/52,800 [13]. In Africa, prevalence is estimated at 1/14,500 [14]. In the United States (USA), prevalence has been reported at close to 1/50,000 screened births [5, 15]. The lowest prevalence of CG/CVG (1/400,000) has been reported among Asian populations [16].

One goal of the current study was to describe and understand any variation in prevalence of CG/CVG among individuals of differing ancestry living in a racially and ethnically diverse country with essentially universal NBS for galactosemia: the United States. To do so, we first predicted the expected frequency of CG/CVG from a combination of publicly available allele frequency data (https://gnomad.broadinstitute.org), stratified by ancestry, and combined this with birth certificate data listing racial and ethnic makeup of the US newborn population. We then tested the accuracy of our calculations by comparing our predicted rates of CG/CVG occurrence to the rates observed in available NBS data from the same time period.

2. Materials and Methods

2.1 Estimating the frequency of homozygosity or compound heterozygosity for pathogenic or likely pathogenic *GALT* variants in US newborns

To predict the likely racial and ethnic distribution of babies born in the US with CG/CVG, we first assembled a list of 137 predicted pathogenic or likely pathogenic variants of *GALT* from the full list of variants reported in *Gnomad*. Next, we noted the frequency, (p_i), of each pathogenic or likely pathogenic variant in human populations that mapped to the racial and ethnic groups reported in US NBS records. Specifically, these included European (non-Finnish) that we mapped to “White”, African/African-American that we mapped to “Black”, Latino that we mapped to “Hispanic”, and South or East Asian (which were averaged and mapped to “Asian”). Assuming both parents in a family derived from the same ancestry and Hardy-Weinberg proportions within populations, we calculated the expected frequency of homozygosity ($\sum p_i^2$) or compound heterozygosity ($\sum 2p_i\, p_j$) of pathogenic and likely pathogenic *GALT* variants, stratified by ancestral population, to estimate the prevalence of infants in each group with GC/CVG. Finally, with the expected prevalence of CG/CVG for each ancestral population in hand, we predicted the number of newborns with CG/CVG by scaling these numbers to newborn counts stratified by reported ancestry [17], predicting a combined total of 310 CG/CVG newborns in the United States between 2016 and 2018.

2.2. Calculating the racial and ethnic distribution of newborns identified with CG/CVG using de-identified newborn screening results

To test the accuracy of our predictions, we reviewed de-identified NBS data collected during the same years (2016–2018) from a subset of states. This combined set included a total of 235 newborns identified with CG/CVG, of whom 194 were listed as White (non-Hispanic or ethnicity unknown), Black (non-Hispanic or ethnicity unknown), Hispanic, or Asian (non-Hispanic or ethnicity unknown) (Table 3). The remaining 41 (17%) were listed on their NBS card as race/ethnicity unknown or other; these babies were excluded from later calculations.

2.3. Statistical comparison between the predicted and observed racial and ethnic distribution of CG/CVG infants born in the US 2016–2018

To compare the expected to observed proportions of newborns within ancestry groups, we used a Fisher’s exact test with predicted versus observed counts of newborns with CG/CVG, for each of the four groups (White, Black, Hispanic, Asian).

3. Results

3.1. Predicted racial and ethnic diversity of babies diagnosed with classic or clinical variant galactosemia in the US 2016–2018

As a first step to predict the racial and ethnic distribution of US newborns with CG/CVG, we consulted a National Vital Statistics Report available from the US Centers for Disease Control and Prevention (CDC) that provided racial and ethnic descriptors for US babies born in the years 2016–2018 [17]. Data from years prior to 2016 were excluded because not all states included a racial designation for Asian babies, or ethnic designations of Hispanic versus non-Hispanic.

During the years 2016–2018, a total of 11,593,087 US newborns were reported, of whom 11,124,512 (96%) were listed on their birth certificates as a single race and either Hispanic or non-Hispanic ethnicity. The remaining 468,575 (4%) were listed as Other or Unknown (Table 2 and Supplemental Table 2).

Next, we identified pathogenic, or likely pathogenic, variants of *GALT* listed in the *Gnomad* data set (https://gnomad.broadinstitute.org) and calculated the frequencies of these alleles in the relevant ancestral populations (Supplemental Table 1). Because the vital statistics report and NBS records grouped East Asians and South Asians together, whereas *Gnomad* listed them separately, we created a single –averaged– allele frequency for these 2 groups, effectively assuming that each Asian ancestral group contributed equally to the US newborn population in 2016–2018). We also made the imperfect but simplifying assumption that both parents in a family would be from the same ancestral group.

Combining the available newborn population demographic data with estimated *GALT* allele frequency by corresponding ancestral group, we calculated the anticipated prevalence of CG/CVG genotypes among babies in each racial and ethnic category. Our calculations predicted that 310 babies born in the US between 2016–2018 should have been diagnosed with CG/CVG. Of these, 202 (65%) would be expected to be White (non-Hispanic), 72 (23%) would be expected to be Black (non-Hispanic), 31 (10%) would be expected to be Hispanic (any race), and 5 (2%) would be expected to be Asian (non-Hispanic) (Table 2). The overall prevalence of CG/CVG among US newborns was predicted from this calculation to be 1/35,886.

3.2. Observed racial and ethnic diversity of babies diagnosed with classic or clinical variant galactosemia in the US 2016–2018

To test our prediction, we used de-identified NBS data from 2016 (19 states), 2017 (21 states), and 2018 (24 states), kindly shared by Sari Edelman, MPH of the NewSTEPs program (https://www.newsteps.org) of the Association of Public Health Laboratories (https://www.aphl.org/Pages/default.aspx), to calculate the observed racial and ethnic diversity of US babies born with CG/CVG. This cohort included a total of 235 babies, or 76% of the 310 we predicted should have been born with CG/CVG in the entire US during the same period. The newborn populations of US states vary widely, for example in 2016 California reported 488,827 births while Vermont reported 5,756 [18], and because privacy restrictions prevented us from knowing which states were included in the cohorts

shared with us, we cannot know whether 235 is high, low, or on target for states contributing to these counts.

Of the 235 newborns for whom we received deidentified NBS data, 41 (17%) were categorized as race and ethnicity “other” or “unknown” and so were excluded from further calculations. To be clear, while 17% is clearly higher than the 4% reported from birth certificates as “other or unknown” for the same years [17], given that parents typically complete these fields in the birth certificate application, whereas a nurse or other healthcare provider typically completes the NBS form without input from the parent(s), it is perhaps not surprising that racial and ethnic designations are more frequently listed as “other or unknown” on NBS forms.

Of the remaining 194 newborns, 129 (66%) were listed as White (non-Hispanic or ethnicity unknown), 32 (16%) were listed as Black (non-Hispanic or ethnicity unknown), 29 (15%) were listed as Hispanic (of any race), and 4 (2%) were listed as Asian (non-Hispanic or ethnicity unknown) (Table 3). As presented in Table 3, none of the individual p-values were even nominally significant ($p < 0.05$ for all), suggesting that we have no statistically significant evidence for any deviation from expected proportions.

3.3. Racial and ethnic diversity of patients included in large studies of outcome in classic galactosemia

Finally, we wanted to know how our expected and observed racial and ethnic diversity of CG/CVG newborns in the US compared with study cohorts from 5 published reports each describing the long-term outcomes of patients with CG/CVG. Each of these 5 reports derived from a study cohort of greater than 100 patients. Of the 3 US studies in this set [7–9], the percentage of cases categorized as White or Caucasian ranged from 84.2% [8] to 93% [7] (Table 4) – all substantially higher than the 65% predicted for the newborn US CG/CVG population (Table 2). The 2 more recent US studies also listed 1.2% to 2.9% of participants as Asian or Pacific Islander, which is not far from the expected prevalence of 2% (Table 2). Of the 2 European studies in the set [6, 10], one [6] listed their cohort as 93.6% Caucasian, and the other [10] did not list race but described their 134 cases as born in the Federal Republic of Germany. Given the predominance of Caucasians in the German population at that time (https://en.wikipedia.org/wiki/Afro-Germans), we surmise that nearly all the 134 patients were likely White/Caucasian.

With regard to ethnicity, 1 of the US studies [9] listed 5% of their participants as Hispanic, while the other [8] listed 7.84%. These numbers were low, but not egregiously so, relative to the 10% predicted from allele frequencies and NBS data (Tables 2 and 3).

4. Discussion

The results presented here demonstrate 2 important points. First, in a racially and ethnically diverse population, CG/CVG is a racially and ethnically diverse disease. This is not a surprising result, but one that deserves attention. This result for CG/CVG is similar to that seen for other allelically heterogeneous disorders, such as phenylketonuria [19], and at odds with data for disorders such as Tay-Sachs [20] or sickle cell anemia [21] that show highly

skewed prevalence numbers among different human populations believed to result from founder effects, carrier selection, or other factors. This result for CG/CVG also stands in stark contrast to Duarte galactosemia (DG), where despite a diverse screened population, the vast majority of babies identified with DG are White/Caucasian [22, 23] because one of the *GALT* alleles required for a diagnosis of DG (D2) is found almost exclusively in people of European ancestry [24]. In CG/CVG, the pathogenic *GALT* alleles that lead to disease, and therefore also the disease itself, are found in peoples of many different ancestries (https://gnomad.broadinstitute.org).

It is also important to note that while the expected and observed racial and ethnic distributions of CG/CVG newborns in the US from 2016–2018 were not statistically distinguishable, they were not identical. Specifically, White (non-Hispanic) and Asian (non-Hispanic) babies were identified in almost exactly the expected proportions (65% expected versus 66% observed, and 2% expected versus 2% observed, respectively). In contrast, Hispanic (any race) babies were identified at a rate slightly higher than expected (10% expected and 15% observed), and Black (non-Hispanic) babies were identified at a rate slightly lower than expected (23% expected and 16% observed).

While these discrepancies are completely consistent with the expected variation in samples this small, because of the small sample size we also are not powered to rule out “small but real” differences at this scale (5–10% differences between expected and observed). Moreover, because our expected counts are based on all 50 states and our observed data represent only a subset of states, some difference between expected and observed should not be surprising. It is also possible the frequency of pathogenic or likely-pathogenic *GALT* alleles listed in *Gnomad* for a given ancestral group might not accurately represent the ancestry of the US group to which we matched it. Finally, it is very likely that some *GALT* variants currently listed as pathogenic/likely pathogenic may eventually be reclassified, while other variants that are truly pathogenic may not yet have been identified and labeled as such.

Another concern is that 41 of the 235 babies (17%) in our NBS set were excluded from further calculations because their race and/or ethnicity was listed as “other” or “unknown” by the person who completed the NBS form. If babies of different ancestries were disproportionately represented in this excluded group, it is possible this discrepancy contributed to, or masked, inter-ancestry differences.

Finally, our calculations assumed complete ascertainment of all US CG/CVG newborns by NBS, and that may not be true. Babies with CG/CVG of African ancestry are more likely than their counterparts of other ancestries to carry the S135L allele [14], which retains substantial residual GALT function [25–27]. It is therefore possible that babies of African ancestry might have been disproportionately missed by NBS, at least in some states, as has been documented for some cases [27]. Nevertheless, all caveats kept carefully in mind, the first major conclusion of this study is that the prevalence of newborns with CG/CVG is consistent with the observed frequency of pathogenic alleles across all ancestry groups examined.

The second important result of this study is that a substantial number of babies born in the US with CG/CVG (approximately 1 in 3) are from non-European backgrounds. In contrast, studies of long-term outcomes in CG/CVG cohort studies in both the US and Europe have been dramatically skewed (usually 90% or more) toward White/Caucasian patients. It is possible, or even likely, the published European studies contain patients roughly in proportion to their ancestry: European. However, this study strongly suggests that long-term studies in the US do not adequately reflect the ancestral diversity of US newborns with CG/CVG. For the US population, and for other countries with a substantial non-European population, we have little current evidence that the published results of long-term outcomes applies equally well to the roughly 1/3 of patients with different ancestry. A recent study documenting notably milder long-term outcomes among patients homozygous or compound heterozygous for S135L, a recognized hypomorphic allele that is most prevalent in patients of African [14, 28], Brazilian [29], and Portuguese [30] ancestry, clearly illustrates this point [27].

A further result reported here is that we predicted the prevalence of CG/CVG in the US as approximately 1/35,000 (Table 2), which is notably higher than the 1/50,000 estimated from US NBS results in 2014, or the 1/47,192 reported in 2020 from US NBS results from 2015–2017 [5, 15]. It is possible that some of the *GALT* alleles defined in *Gnomad* as pathogenic or likely pathogenic do not actually lead to newborn abnormality sufficient to be flagged by NBS, and/or that some of our other simplifying assumptions caused us to overestimate. However, it is also possible that actual prevalence is higher than 1 in 47–50,000, and that NBS misses some babies with CG/CVG. Of note, the total US newborn count in 2016–2018 was 11,593,087. If the prevalence were 1 in 47,192, there should have been a total of 246 newborns with CG/CVG in the whole US. Here we describe 235 newborns ascertained from 19–24 US states, depending on the year. This apparent discrepancy might be, at least in part, explained if the states missing from our data set are those with relatively low annual birth counts. Of note, this discrepancy cannot be attributed to some infants with the more common Duarte galactosemia [5] inadvertently being reported with NBS records for classic galactosemia, as that would have resulted in the NBS prevalence of CG/CVG exceeding prediction; this is the opposite of what we see.

Limitations:

This study makes important points, but the results must be interpreted in context, and mindful of many important limitations. First, NBS records listed Hispanic babies only by ethnicity, and not also by race. This means that Hispanic babies were missing from their corresponding racial group. If the racial categories of these Hispanic babies were asymmetrically distributed, that could have impacted our results.

Further, all 5 of the outcome studies we considered here (Table 4) included adults and older children, as well as babies. If population demographics in the US and Europe have changed over time, the racial and ethnic distribution of patients born in 2016–2018 may not accurately predict the racial and ethnic distribution of a mixed-age patient cohort at the time those outcome studies were performed. Another important limitation in comparing the racial and ethnic distribution of recent US newborns with CG/CVG to both the US

and European study cohorts listed in Table 4 is that all of those studies included patients born in more than one country. While confounding, this caveat is a reality for most rare disease studies. Whether any assembled cohort of patients is representative of any individual nation's population is therefore almost impossible to assess.

Other limitations of this study include the relatively small sample size (e.g., 235 newborns with CG/CVG) and the fact that these newborns came from an anonymous subset of US states. The simplifying assumptions embedded in our estimation of disease prevalence are also imperfect. Specifically, we assumed complete ascertainment of CG/CVG by NBS, accurate correspondence of ancestries between similarly labeled groups in *Gnomad*, US Vital Statistics, and US NBS records, and both parents of newborns being from the same ancestral group. None of these assumptions is likely to be completely true.

5. Conclusions:

Both population allele frequencies and newborn screening records confirm a racial and ethnic diversity among patients with classic and clinical variant galactosemia in the US that is not represented in most reported studies of long-term patient outcome.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We gratefully acknowledge the assistance of Ms. Sari Edelman, MPH of the NewSTEPs, Newborn Screening & Genetics team at APHL who provided us access to the de-identified NBS records used in this study.

Data Availability Statement:

All data for this manuscript are included in compiled form in Supplemental Tables.

7. References

[1]. Berry G, Classic Galactosemia and Clinical Variant Galactosemia, in: Adam MP ED, Mirzaa GM, et al. (Ed.), GeneReviews® [Internet], University of Washington, Seattle (WA), 2020.

[2]. Berry GT, Moate PJ, Reynolds RA, Yager CT, Ning C, Boston RC, Segal S, The rate of de novo galactose synthesis in patients with galactose-1-phosphate uridyltransferase deficiency Mol Genet Metab 81 (2004) 22–30. [PubMed: 14728988]

[3]. Berry GT, Nissim I, Lin ZP, Mazur AT, Gibson JB, Segal S, Endogenous Synthesis of Galactose in Normal Men and Patients with Hereditary Galactosemia Lancet 346 (1995) 1073–1074. [PubMed: 7564790]

[4]. Schadewaldt P, Kamalanathan L, Hammen HW, Kotzka J, Wendel U, Endogenous galactose formation in galactose-1-phosphate uridyltransferase deficiency Arch Physiol Biochem 120 (2014) 228–239. [PubMed: 25268296]

[5]. Pyhtila BM, Shaw KA, Neumann SE, Fridovich-Keil JL, Newborn screening for galactosemia in the United States: looking back, looking around, and looking ahead JIMD Rep 15 (2015) 79–93. [PubMed: 24718839]

[6]. Rubio-Gozalbo ME, Haskovic M, Bosch AM, Burnyte B, Coelho AI, Cassiman D, Couce ML, Dawson C, Demirbas D, Derks T, Eyskens F, Forga MT, Grunewald S, Haberle J, Hochuli M, Hubert A, Huidekoper HH, Janeiro P, Kotzka J, Knerr I, Labrune P, Landau YE, Langendonk

JG, Moslinger D, Muller-Wieland D, Murphy E, Ounap K, Ramadza D, Rivera IA, Scholl-Buergi S, Stepien KM, Thijs A, Tran C, Vara R, Visser G, Vos R, de Vries M, Waisbren SE, Welsink-Karssies MM, Wortmann SB, Gautschi M, Treacy EP, Berry GT, The natural history of classic galactosemia: lessons from the GalNet registry Orphanet J Rare Dis 14 (2019) 86.

[7]. Waggoner DD, Buist NR, Donnell GN, Long-term prognosis in galactosaemia: results of a survey of 350 cases J Inherit Metab Dis 13 (1990) 802–818. [PubMed: 1706789]

[8]. Frederick AB, Zinsli AM, Carlock G, Conneely K, Fridovich-Keil JL, Presentation, progression, and predictors of ovarian insufficiency in classic galactosemia J Inherit Metab Dis 41 (2018) 785–790. [PubMed: 29721917]

[9]. Frederick AB, Cutler DJ, Fridovich-Keil JL, Rigor of non-dairy galactose restriction in early childhood, measured by retrospective survey, does not associate with severity of five long-term outcomes quantified in 231 children and adults with classic galactosemia J Inherit Metab Dis 40 (2017) 813–821. [PubMed: 28695375]

[10]. Schweitzer S, Shin Y, Jakobs C, Brodehl J, Long-term outcome in 134 patients with galactosaemia Eur J Pediatr 152 (1993) 36–43. [PubMed: 8444204]

[11]. Coelho AI, Rubio-Gozalbo ME, Vicente JB, Rivera I, Sweet and sour: an update on classic galactosemia J Inherit Metab Dis 40 (2017) 325–342. [PubMed: 28281081]

[12]. Coss KP, Doran PP, Owoeye C, Codd MB, Hamid N, Mayne PD, Crushell E, Knerr I, Monavari AA, Treacy EP, Classical Galactosaemia in Ireland: incidence, complications and outcomes of treatment J Inherit Metab Dis 36 (2013) 21–27. [PubMed: 22870861]

[13]. Welling L, Boelen A, Derks TG, Schielen PC, de Vries M, Williams M, Wijburg FA, Bosch AM, Nine years of newborn screening for classical galactosemia in the Netherlands: Effectiveness of screening methods, and identification of patients with previously unreported phenotypes Mol Genet Metab 120 (2017) 223–228. [PubMed: 28065439]

[14]. Henderson H, Leisegang F, Brown R, Eley B, The clinical and molecular spectrum of galactosemia in patients from the Cape Town region of South Africa BMC Pediatr 2 (2002) 7. [PubMed: 12350230]

[15]. Sontag MK, Yusuf C, Grosse SD, Edelman S, Miller JI, McKasson S, Kellar-Guenther Y, Gaffney M, Hinton CF, Cuthbert C, Singh S, Ojodu J, Shapira SK, Infants with Congenital Disorders Identified Through Newborn Screening - United States, 2015–2017 MMWR Morb Mortal Wkly Rep 69 (2020) 1265–1268. [PubMed: 32915168]

[16]. Cheung K, Tang N, Hsiao K, Law L, Wong W, Ng P, Pang C, Applegarth D, Fok T, Hjelm N, Classical galactosaemia in Chinese: A case report and review of disease incidence Journal of Paediatrics and Child Health 35 (1999) 399–400.

[17]. Martin J, Hamilton B, Osterman M, Driscoll A, Births: Final data for 2018, National Vital Statistics Reports, National Center for Health Statistics (U.S.), Hyattsville, MD, 2019.

[18]. Martin J, Hamilton B, Osterman M, Driscoll A, Drake P, Title : Births : final data for 2016, National vital statistics reports, National Center for Health Statistics (U.S.), 2018.

[19]. Foreman PK, Margulis AV, Alexander K, Shediac R, Calingaert B, Harding A, Pladevall-Vila M, Landis S, Birth prevalence of phenylalanine hydroxylase deficiency: a systematic literature review and meta-analysis Orphanet J Rare Dis 16 (2021) 253.

[20]. Toro C, Shirvan L, Tifft C, HEXA Disorders, in: Adam M, Everman D, Mirzaa G, e. al. (Eds.), GeneReviews® [Internet], University of Washington, Seattle, Seattle (WA), 1999.

[21]. Bender M, Carlberg K, Sickle Cell Disease, in: Adam M, Everman D, Mirzaa G, e. al. (Eds.), GeneReviews® [Internet], University of Washington, Seattle, Seattle (WA), 2003.

[22]. Carlock G, Fischer ST, Lynch ME, Potter NL, Coles CD, Epstein MP, Mulle JG, Kable JA, Barrett CE, Edwards SM, Wilson E, Fridovich-Keil JL, Developmental Outcomes in Duarte Galactosemia Pediatrics 143 (2019) e20182516. [PubMed: 30593450]

[23]. Fridovich-Keil JL, Carlock G, Coles CD, Lynch ME, Millians MN, Potter NL, Powell K, Richards P, Singh R, Wittenauer A, Developmental outcomes of children with Duarte galactosemia: exploring the bases of an apparent contradiction in the literature Genet Med 21 (2019) 2683–2685. [PubMed: 31160755]

[24]. Carney AE, Sanders RD, Garza KR, McGaha LA, Bean LJ, Coffee BW, Thomas JW, Cutler DJ, Kurtkaya NL, Fridovich-Keil JL, Origins, distribution and expression of the Duarte-2 (D2) allele

of galactose-1-phosphate uridylyltransferase Hum Mol Genet 18 (2009) 1624–1632. [PubMed: 19224951]

[25]. Fridovich-Keil JL, Langley SD, Mazur LA, Lennon JC, Dembure PP, Elsas JL, 2nd, Identification and functional analysis of three distinct mutations in the human galactose-1-phosphate uridyltransferase gene associated with galactosemia in a single family Am J Hum Genet 56 (1995) 640–646. [PubMed: 7887417]

[26]. Wells L, Fridovich-Keil JL, Biochemical characterization of the S135L allele of galactose-1-phosphate uridylyltransferase associated with galactosemia J. Inher. Metab. Dis. 20 (1997) 633–642. [PubMed: 9323558]

[27]. Katler QS, Stepien KM, Paull N, Patel S, Adams M, Balci MC, Berry GT, Bosch AM, DeLa OA, Demirbas D, Edman J, Ficicioglu C, Goff M, Hacker S, Knerr I, Lancaster K, Li H, Mendelsohn BA, Nichols B, de Rezende Pinto WBV, Rocha JC, Rubio-Gozalbo ME, Saad-Naguib M, Scholl-Buergi S, Searcy S, de Souza PVS, Wittenauer A, Fridovich-Keil JL, A multinational study of acute and long-term outcomes of Type 1 galactosemia patients who carry the S135L (c.404C > T) variant of GALT J Inherit Metab Dis (2022).

[28]. Lai K, Langley SD, Singh RH, Dembure PP, Hjelm LN, Elsas LJ, A prevalent mutation for galactosemia among black Americans Journal of Pediatrics 128 (1996) 89–95. [PubMed: 8551426]

[29]. Garcia DF, Camelo JS Jr., Molfetta GA, Turcato M, Souza CF, Porta G, Steiner CE, Silva WA Jr., Clinical profile and molecular characterization of Galactosemia in Brazil: identification of seven novel mutations BMC Med Genet 17 (2016) 39. [PubMed: 27176039]

[30]. Coelho AI, Ramos R, Gaspar A, Costa C, Oliveira A, Diogo L, Garcia P, Paiva S, Martins E, Teles EL, Rodrigues E, Cardoso MT, Ferreira E, Sequeira S, Leite M, Silva MJ, de Almeida IT, Vicente JB, Rivera I, A frequent splicing mutation and novel missense mutations color the updated mutational spectrum of classic galactosemia in Portugal J Inherit Metab Dis 37 (2014) 43–52. [PubMed: 23749220]

Highlights

- *GALT* variant frequencies in relevant ancestral groups predict the diversity of infants born with CG/CVG in a population
- Newborn screening records confirm the racial and ethnic diversity of US newborns with CG/CVG
- Prior large studies of outcome in CG/CVG do not accurately reflect the diversity of patients anticipated in the US

Table 1:

Combined allele frequencies and derived prevalence of homozygosity or compound heterozygosity for *GALT* variants considered pathogenic or likely pathogenic in relevant populations described in *Gnomad*

Population	Allele Frequency of Pathogenic or Likely Pathogenic *GALT* variants	Predicted Newborn Prevalence of CG/CVG	Predicted Prevalence of CG/CVG
Worldwide	0.0053	2.79E-05	1/35,823
African/African-American	0.0066	4.31E-05	1/23,211
Asian, averaged	0.0034	6.37E-06	1/156,972
East Asian	(0.0019)	(3.71E-06)	(1/269,797)
South Asian	(0.0048)	(2.27E-05)	(1/44,148)
European (non-Finnish)	0.0058	3.37E-05	1/29,660
Latino	0.0034	1.16E-05	1/86,257

Table 2:

Predicted US CG/CVG births 2016 – 2018 by race and ethnicity

Group	US newborns in this group (2016–2018)*	predicted prevalence of CG/CVG in this group (from Table 1)	# CG/CVG US newborns predicted in this group (2016–2018)
White (non-Hispanic)	6,005,206 (52%)	1/29,660	202 (65%)
Hispanic (any race)	2,703,421 (23%)	1/86,257	31 (10%)
Black (non-Hispanic)	1,671,366 (14%)	1/23,211	72 (23%)
Asian (non-Hispanic)	744,519 (6%)	1/156,972	5 (2%)
Other or Unknown	468,575 (4%)	-----	-----
Total (groups listed above)	11,593,087 (100%)	1/35,886**	310 (100%)

*Data from Supplemental Table 2. Percentages were rounded to whole numbers.

**Calculated from the data in this table.

Table 3:

Demographics of US newborns identified by NBS with classic galactosemia in 2016–2018 included in this study

Group	# CG/CVG identified by NBS 2016–2018	% of CG/CVG newborns (excluding those listed as other or unknown)	Fisher's exact test p-value (observed category versus the corresponding prediction from Table 2)
White (non-Hispanic or unknown)	129	66%	p=0.7733
Black (non-Hispanic or unknown)	32	16%	p=0.0715
Hispanic (any race)	29	15%	p=0.1194
Asian (non-Hispanic or unknown)	4	2%	p=0.7386
Other or unknown	41	excluded	---
Total	235	194 (100%)	---

Table 4:

Racial and ethnic diversity of study cohorts from 5 published studies of outcome in CG/CVG each including at least 100 cases

Report	Racial makeup of cases in this study cohort
Waggoner et al (1990) [7] 350 cases, 79% USA and 21% Europe	93% Caucasian 7% Black 1 mixed race
Schweitzer et al (1993) [10] 134 cases (Federal Republic of Germany)	Race not listed but from national demographics assumed to be almost exclusively White/ Caucasian
Frederick et al (2017) [9] 322 cases, predominantly USA	81.1% White (non-Hispanic) 5% White (Hispanic) 1.2% Asian or Pacific Islander 12.7% Mixed or unknown
Frederick et al (2018) [8] 102 cases, predominantly USA	76.4% White (non-Hispanic) 7.84% White (Hispanic) 2.9% Asian or Pacific Islander 12.74% Mixed or unknown
Rubio-Gozalbo et al (2019) [6] 509 cases, predominantly Europe	93.6% Caucasian 6.4% other