

## Phenotypes of undiagnosed adults with actionable *OTC* and *GLA* variants

Jessica I. Gold,<sup>1</sup> Sarina Madhavan,<sup>2,3</sup> Joseph Park,<sup>4</sup> Hana Zouk,<sup>2,5,6</sup> Emma Perez,<sup>7,8</sup> Alanna Strong,<sup>1,9,10</sup> Theodore G. Drivas,<sup>11</sup> Amel Karaa,<sup>12,13</sup> Marc Yudkoff,<sup>1,9</sup> Daniel Rader,<sup>4,11,14</sup> Regeneron Genetics Center, Penn Medicine BioBank, Robert C. Green,<sup>2,8,15,16</sup> and Nina B. Gold<sup>12,13,17,\*</sup>

### Summary

Inherited metabolic disorders (IMDs) are variably expressive, complicating identification of affected individuals. A genotype-first approach can identify individuals at risk for morbidity and mortality from undiagnosed IMDs and can lead to protocols that improve clinical detection, counseling, and management. Using data from 57,340 participants in two hospital biobanks, we assessed the frequency and phenotypes of individuals with pathogenic/likely pathogenic variants (PLPVs) in two IMD genes: *GLA*, associated with Fabry disease, and *OTC*, associated with ornithine transcarbamylase deficiency. Approximately 1 in 19,100 participants harbored an undiagnosed PLPV in *GLA* or *OTC*. We identified three individuals (2 male, 1 female) with PLPVs in *GLA*, all of whom were undiagnosed, and three individuals (3 female) with PLPVs in *OTC*, two of whom were undiagnosed. All three individuals with PLPVs in *GLA* (100%) had symptoms suggestive of mild Fabry disease, and one individual (14.2%) had an ischemic stroke at age 33, likely indicating the presence of classic disease. No individuals with PLPVs in *OTC* had documented hyperammonemia despite exposure to catabolic states, but all (100%) had chronic symptoms suggestive of attenuated disease, including mood disorders and migraines. Our findings suggest that *GLA* and *OTC* variants identified via a genotype-first approach are of high penetrance and that population screening of these genes can be used to facilitate stepwise phenotyping and appropriate care.

Genetic testing is a powerful diagnostic tool in the evaluation of individuals with symptoms suggestive of rare disease, however, this “phenotype-first” approach may overlook individuals with misdiagnoses or subclinical disease, leading to underdiagnosis, symptom progression, and inappropriate recurrence risk counseling.<sup>1</sup> In contrast, “genotype-first” methods, in which unbiased genomic studies are carried out through biobanks or other population cohorts, followed by subsequent phenotype delineation, have revealed a higher frequency and a wider range of phenotypic heterogeneity than expected for many monogenic disorders.<sup>2,3</sup> For actionable disorders, such as inherited metabolic disorders (IMDs), it is important to identify individuals who could benefit from appropriate surveillance and management to prevent unnecessary morbidity and mortality. Conversely, identifying and describing the clinical features of undiagnosed individuals with pathogenic/likely pathogenic variants (PLPVs) can allow for more accurate counseling regarding phenotype variability in IMDs and can inform protocols to prevent the inappropriate application of costly or risky therapeutics in individuals with minimal signs of disease. Although individuals with IMDs have been previously identified in unselected co-

horts, to our knowledge, no efforts to ascertain the phenotypes of such individuals has been undertaken.

The genes designated for secondary findings (SFs) by the American College of Medical Genetics and Genomics (ACMG) represent a clinical genotype-first approach in which individuals with PLPVs in medically actionable genes may be identified during an evaluation process for unrelated symptoms. Variants in the 56–59 genes originally selected for SFs<sup>4</sup> have consistently been found in 1%–3% of individuals.<sup>5,6</sup> The current list (ACMG SF v.3.1) now includes 78 genes, including four genes associated with IMDs: *BTD*, associated with biotinidase deficiency (MIM: 609019); *GLA*, associated with Fabry disease (MIM: 300644); *OTC*, associated with ornithine transcarbamylase (OTC) deficiency (MIM: 300461); and *GAA*, associated with Pompe disease (MIM: 606800).<sup>7</sup> The frequency and phenotypes of unselected individuals in biobanks with PLPVs in genes associated with these IMDs have not been systematically evaluated.

In this study, we used a genotype-first approach to assess the prevalence and expressivity of two X-linked IMDs, Fabry disease and OTC deficiency, in two hospital biobanks. Fabry disease is commonly associated with neurologic, cardiac,

<sup>1</sup>Division of Human Genetics, Department of Pediatrics, Children’s Hospital of Philadelphia; Philadelphia, PA, USA; <sup>2</sup>Harvard Medical School, Boston, MA, USA; <sup>3</sup>Harvard Business School, Cambridge, MA, USA; <sup>4</sup>Department of Genetics, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA; <sup>5</sup>Laboratory for Molecular Medicine, Mass General Brigham Personalized Medicine, Cambridge MA, USA; <sup>6</sup>Department of Pathology, Massachusetts General Hospital, Boston, MA, USA; <sup>7</sup>Mass General Brigham Personalized Medicine, Cambridge MA, USA; <sup>8</sup>Department of Medicine, Brigham and Women’s Hospital, Boston, MA, USA; <sup>9</sup>Division of Human Genetics, Children’s Hospital of Philadelphia, Philadelphia, PA, USA; <sup>10</sup>Center for Applied Genomics, Children’s Hospital of Philadelphia, Philadelphia, PA, USA; <sup>11</sup>Division of Translational Medicine and Human Genetics, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA; <sup>12</sup>Massachusetts General Hospital for Children, Boston, MA, USA; <sup>13</sup>Department of Pediatrics, Harvard Medical School, Boston, MA, USA; <sup>14</sup>Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; <sup>15</sup>Broad Institute, Boston, MA, USA; <sup>16</sup>Ariadne Labs, Boston, MA, USA

<sup>17</sup>Lead contact

\*Correspondence: [ngold@mgh.harvard.edu](mailto:ngold@mgh.harvard.edu)

<https://doi.org/10.1016/j.xhgg.2023.100226>.

© 2023 The Author(s). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



and renal disease and has previously been estimated to occur in 1 in 40,000 to 1 in 117,000 individuals,<sup>8</sup> although a recent genotype-first analysis in the UK Biobank revealed more than 1 in 10,000 participants with PLPVs in *GLA*.<sup>9</sup> OTC deficiency, estimated to occur in 1 in 14,000 to 1 in 77,000 individuals<sup>10</sup> and found in 1 in approximately 22,000 unselected individuals in the eMERGE study,<sup>1</sup> leads to variably expressive hyperammonemia, encephalopathy, neuropsychiatric symptoms, and hepatic dysfunction, particularly in the setting of malnutrition, intercurrent illness, or systemic steroid usage. X-linked disorders such as OTC deficiency and Fabry disease are ideal candidates for study given that heterozygous and hemizygous PLPVs in these genes can cause symptoms and do not require parental sequencing or other strategies for variant phasing. The identification of individuals with PLPVs in *GLA* and *OTC* in hospital biobanks can inform the prognosis and potential role for therapy for individuals receiving SFs in these genes.

The Mass General Brigham Biobank (MGBB) is a biorepository within the Mass General Brigham (MGB) academic medical center and is linked to electronic medical records (EMRs). Between July 1, 2010, and March 31, 2021, the MGBB enrolled 124,391 individuals.<sup>11</sup> A subset of 13,340 samples underwent exome sequencing at the Clinical Research Sequencing Platform (CRSP) of the Broad Institute of Harvard and MIT using a custom capture library from TWIST Biosciences with sequencing on the Illumina NovaSeq using 150 bp paired reads, as previously described.<sup>11</sup> Variants were filtered to a list of 59 genes comprising ACMG SF v.2.0,<sup>12</sup> which included *GLA* and *OTC*. Variant annotation, additional filtration, and classification are described in detail by Blout Zawatsky et al.<sup>11</sup>

The Penn Medicine Biobank (PMBB) is an EMR-linked biorepository at the University of Pennsylvania. Since its creation in 2013, the PMBB has enrolled 174,712 participants and has completed exome sequencing on approximately 44,000 participants using an Illumina NovaSeq platform, recently detailed by Verma et al.<sup>13</sup> PMBB exome sequencing data were subsetted to include only the *GLA* and *OTC* gene loci and were subsequently annotated using the Ensembl Variant Effect Predictor (VEP; v.102) with the LOFTEE (v.0.3) and dbNSFP (v.4.2) plugins. We excluded participants with variants in allele balance <40% in females and <80% in males.

Informed consent for broad use of samples and data, including but not limited to exome sequencing, phenotype assessment, and publication of results, was obtained at the time of enrollment for both MGBB and PMBB participants. EMR review for this study was approved by the Institutional Review Board (IRB) of both Massachusetts General Hospital and the University of Pennsylvania. Participants provided informed consent at the time of biobank enrollment.

Phenotyping was completed through EMR review in both biobanks. All documentation in the EMRs, including progress and consult notes, operative reports, and reports from diagnostic procedures, were reviewed to identify features known to be associated with either Fabry disease or OTC deficiency. Data from the EMR were collected using the REDCap

electronic data capture tool. Criteria for the REDCap survey were based on physical manifestations described in Batshaw et al.<sup>14</sup> and in Saudubray<sup>15</sup> (Tables S3 and S4). Demographic information was also ascertained (Table S5).

Three individuals (all from MGBB) had a PLPV in *GLA* (Figure 1). Two individuals had one X chromosome (chromosomal males) and one had two X chromosomes (chromosomal females) (Table S6). None had a diagnosis or known family history of Fabry disease.

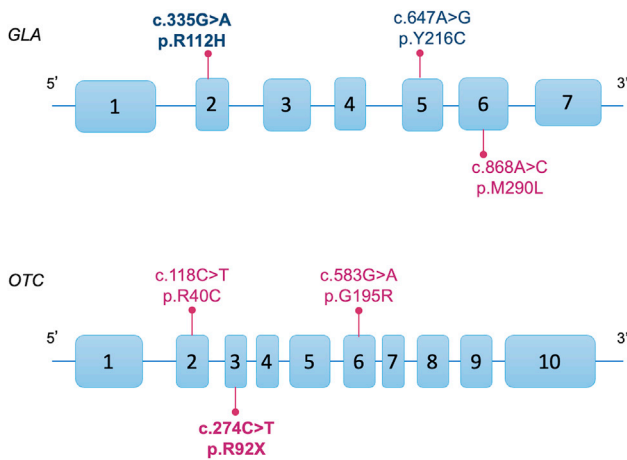
Retrospective phenotyping suggested a possible family history of Fabry disease in one individual whose maternal aunt (M5) had a stroke at unknown age. All individuals had previously had at least one encounter with a medical specialist that evaluated a system affected in Fabry disease, including cardiology (2/3, 66.7%), ophthalmology (1/3, 33.3%), neurology (1/3, 33.3%), or dermatology (1/3, 33.3%). No individuals had an assessment by nephrology prior to the receipt of genomic results.

At least one potential sign or symptoms consistent with Fabry disease was present in all identified individuals (Figure 2B). One individual (M5) had an ischemic stroke at 33 years of age. Another individual (M4) had concentric left ventricular hypertrophy and an abnormal electrocardiogram (first-degree atrioventricular block, right bundle branch block). The only female individual (F1) endorsed subjective symptoms of Fabry disease including acroparesthesia, vertigo, and heat intolerance, but had no objective signs of disease on an ophthalmologic exam and transthoracic echocardiogram.

Three individuals (one from MGBB, two from PMBB), all chromosomal females, had a PLPV in *OTC*. Two of these three individuals were previously undiagnosed. One chromosomal female (F2) had previously received a molecular diagnosis of OTC deficiency due to a known history of disease in her daughter. No other individuals had a documented family history that was suggestive of OTC deficiency, and, aside from F1, none had received prior clinical diagnoses.

All individuals with PLPVs in *OTC* had at least one clinical exposure that could precipitate hyperammonemia, including pregnancy, prolonged fasting, or major infection; but none had documentation of acute hyperammonemia, a protracted episode of altered mental status, or seizures. Imaging of the liver in one individual (F2) did not reveal any OTC-related findings such as hepatic inflammation or fibrosis. However, all individuals with *OTC* variants (100%) had symptoms of attenuated disease. The most common clinical features were mood disorders (3/3, 100%) and migraines (2/3, 66.7%) (Figure 2A).

A genotype-first approach can identify the frequency and variable expressivity of IMDs in unselected biobank participants, thereby better informing the prognosis and utility of treatment in individuals who receive PLPVs in genes related to these conditions outside of diagnostic testing, such as by genomic newborn screening or as a clinical SF. In this study, we found that approximately 1 in 19,000 adults in two hospital biobanks harbored



**Figure 1. Pathogenic and likely pathogenic variants in *OTC* and *GLA* identified in hospital biobank participants**

Gene schematic of *OTC* (top) and *GLA* (bottom) showing pathogenic or likely pathogenic variants (PLPVs) identified in the Mass General Brigham Biobank and Penn Medicine Biobank. Pathogenic variants are bolded. PLPVs identified in individuals with one X chromosome are in blue, and PLPVs identified in individuals with two X chromosomes are in purple.

previously unrecognized PLPVs in *GLA* or *OTC*. This parallels the rate of PLPVs in *OTC* found in the eMERGE study. For *GLA*, our rate is lower than that found in the UK Biobank study or newborn screening for Fabry disease in Taiwan, possibly due to the difference in enrolled populations or variant curation criteria.<sup>9,16</sup> Phenotypes in individuals in this study with PLPVs ranged from very mildly symptomatic to classic disease, with the majority demonstrating signs of attenuated disease.

Our findings suggest that PLPVs in IMD genes may serve as markers of disease risk much like other genes on the ACMG SF list, such as those associated with hereditary cancer predisposition or heart disease, rather than as straightforward diagnostic markers. Although most individuals had signs of attenuated disease, one individual (M2) with a PLPV in *GLA* experienced an ischemic stroke at age 33, possibly as a direct result of classic Fabry disease. The identification of a PLPV in an IMD gene may be used to prompt disease surveillance such as a targeted history and physical exam, biochemical laboratory tests, and imaging, the results of which can be used in concert to determine the need for further therapy. Additionally, this information can be used to initiate cascade familial screening and recurrence risk counseling.

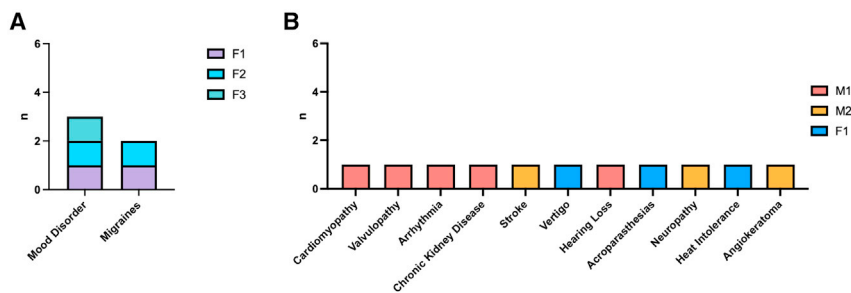
This study highlights the importance of medically actionable return of genomic results (gRoR) in research. Variants identified in participants from the MGBB underwent orthogonal confirmation via Sanger sequencing, which we suggest is an essential measure prior to the return of genotype-first results from research. All four of the individuals in this study from the MGBB were then offered gRoR, and three accepted, thereby facilitating appropriate evaluations and therapy. In two of three such individuals, biochemical testing was consistent with the genetic variant, substantiating the molecular diagnosis. For some individuals, such information

can be lifesaving. Additionally, genetic information may be useful in the management of common, complex disease symptoms such as mood disorders, migraines, and cardiomyopathy. Although such symptoms are widespread in the adult population, in rare cases, they can be due to an attenuated form of a genetic disorder.<sup>16–20</sup> Future studies can determine the frequency of IMD-related PLPVs in biobank participants with common diagnoses, such as mood disorders, to explore the utility of genomic testing in this population.

This study also highlighted the importance of assessing the allelic fractions of variants identified in biobank participants. Initial analyses demonstrated a higher proportion of individuals with pathogenic variants in *GLA* and *OTC*. In several instances, these variants were found to have low allelic fractions, however, suggesting that they may be due to post-zygotic mosaicism or sequencing artifact. As such, we must take caution when using biobank data to assess variant prevalence and penetrance by first ensuring that the allele balance of each variant is consistent with the appropriate mechanism of inheritance and disease.

This study has several limitations. First, hospital biobanks may not represent the general population. Both the MGBB and the PMBB are limited in their representation of participants from diverse backgrounds. Individuals with complex medical needs or developmental disabilities may have been less likely to consent to biobank protocols. Data available for EMR review are also limited among those individuals with limited interaction with the healthcare system. The forms used to capture participant information did not distinguish between a lack of information in EMRs and the absence of a symptom. It is therefore possible that some individuals might report more symptoms if they were asked targeted disease-related questions, suggesting that return of results and recalling participants for deeper phenotyping is an important next step in studying the penetrance of IMDs. Additionally, several of the symptoms of attenuated IMDs are common, complex disorders, such as mood disorders and migraines, that may be caused by a variety of factors. Comparison of rates of these conditions among individuals with PLPVs and control participants, or statistical methods such as phenome-wide association studies, may be useful to determine which symptoms are most strongly linked to IMD genes in the future.

Taken together, our findings demonstrate that unselected individuals in hospital-based biobanks with PLPVs in genes associated with IMDs are at high risk for disease symptoms. Although only a minority of individuals with PLPVs had classic symptoms of disease, all had signs of at least attenuated disease and would benefit from appropriate surveillance and possibly therapy. Although the expressivity of the PLPV varies, the overall positive predictive rate of genomic screening for these genes would be expected to be high. Over time, data from unselected individuals in biobanks will be essential to determine the variable expressivity of PLPVs in genes associated with IMDs and other monogenic disorders, thereby improving the counseling of individuals who receive unanticipated PLPVs in such genes.



**Figure 2. Phenotypes of individuals with pathogenic and likely pathogenic variants in *OTC* and *GLA* identified via a genotype-first approach in hospital biobanks**

(A and B) Phenotypes of individuals with pathogenic or likely pathogenic variants in *OTC* (A) or *GLA* (B). For individuals with variants in *OTC*, there was no documentation of the following signs or symptoms: prenatal or postnatal complications, psychosis, history of altered mental status, cognitive impairment, cyclic vomiting, seizures, protein intolerance, self-restricted diet, malignancies of the liver (including hepatocellular carcinoma),

Reye syndrome, hyperammonemia, or elevated transaminases. For *GLA*, there was no documentation of the following signs or symptoms of Fabry disease: exercise intolerance, aortopathy corneal whorls, proteinuria, end-stage renal disease (including history of dialysis or renal transplantation), heart failure, or hypohidrosis.

## Data and code availability

The data supporting the current study have not been deposited in a public repository because data from hospital biobank participants is not public. Deidentified data are available upon request through Dr. Nina Gold.

## Supplemental information

Supplemental information can be found online at <https://doi.org/10.1016/j.xhgg.2023.100226>.

## Acknowledgments

See [Table S1](#) for complete author information for the Regeneron Genetics Center. See [Table S2](#) for complete author information for the PMBB. We acknowledge the PMBB for providing data and thank the patient participants of Penn Medicine who consented to participate in this research program. We would also like to thank the PMBB team and the Regeneron Genetics Center for providing genetic variant data for analysis. The PMBB is approved under IRB protocol #813913 and is supported by Perelman School of Medicine at University of Pennsylvania, a gift from the Smilow family, and the National Center for Advancing Translational Sciences of the National Institutes of Health under CTSA award number UL1TR001878. We acknowledge the Mass General Brigham Biobank for providing samples, genomic data, and health information data. The contribution of Dr. Marc Yudkoff was supported by the US Army Medical Research Acquisition Activity, 808 Schreider Street, Fort Detrick, MD 21702-5014, under award no. W81XWH2210590. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the Department of Defense. J.G. received funding from NIH T32 GM008638. M.Y. received funding from DOD W81XWH2210590. R.C.G. received funding from NIH HD077671 and TR003201. N.B.G. received funding from NIH TR003201 and K08 HG012811.

## Author contributions

Conceptualization, J.I.G., S.M., H.Z., A.S., D.R., R.C.G., and N.B.G.; data curation, J.I.G., S.M., J.P., H.Z., E.P., T.G.D., A.K., and N.B.G.; funding acquisition, R.C.G. and N.B.G.; investigation, J.I.G., S.M., J.P., H.Z., E.P., T.G.D., A.K., M.Y., D.R., and N.B.G.; visualization, J.I.G. and N.B.G.; writing – original draft, J.I.G.

and N.B.G.; writing – review & editing, J.I.G., S.M., J.P., H.Z., E.P., A.S., T.G.D., A.K., M.Y., D.R., R.C.G., and N.B.G.

## Declaration of interests

E.P. is a paid consultant for Allelica, Inc. R.C.G. has received compensation for advising the following companies, Allelica, Atria, Fabric, Genome Web, Genomic Life, Verily, and VinBigData, and is co-founder of Genome Medical and Nurture Genomics. N.B.G. is a paid consultant for RCG Consulting.

Received: March 20, 2023

Accepted: July 21, 2023

## Web resources

Online Mendelian Inheritance in Man. 1966–2023, <https://omim.org/>.

## References

1. eMERGE Clinical Annotation Working Group (2020). Frequency of genomic secondary findings among 21,915 eMERGE network participants. *Genet. Med.* 22, 1470–1477.
2. Wenger, B.M., Patel, N., Lui, M., Moscati, A., Do, R., Stewart, D.R., Tartaglia, M., Muiño-Mosquera, L., De Backer, J., Kontorovich, A.R., and Gelb, B.D. (2021). A genotype-first approach to exploring Mendelian cardiovascular traits with clear external manifestations. *Genet. Med.* 23, 94–102.
3. Abul-Husn, N.S., Manickam, K., Jones, L.K., Wright, E.A., Hartzel, D.N., Gonzaga-Jauregui, C., O'Dushlaine, C., Leader, J.B., Lester Kirchner, H., Lindbuchler, D.M., et al. (2016). Genetic identification of familial hypercholesterolemia within a single U.S. health care system. *Science* 354, aaf7000. <https://doi.org/10.1126/science.aaf7000>.
4. Green, R.C., Berg, J.S., Grody, W.W., Kalia, S.S., Korf, B.R., Martin, C.L., McGuire, A.L., Nussbaum, R.L., O'Daniel, J.M., Ormond, K.E., et al. (2013). ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet. Med.* 15, 565–574.
5. Natarajan, P., Gold, N.B., Bick, A.G., McLaughlin, H., Kraft, P., Rehm, H.L., Peloso, G.M., Wilson, J.G., Correa, A., Seidman, J.G., et al. (2016). Aggregate penetrance of genomic variants for actionable disorders in European and African Americans. *Sci. Transl. Med.* 8, 364ra151.

6. Amendola, L.M., Dorschner, M.O., Robertson, P.D., Salama, J.S., Hart, R., Shirts, B.H., Murray, M.L., Tokita, M.J., Gallego, C.J., Kim, D.S., et al. (2015). Actionable exomic incidental findings in 6503 participants: challenges of variant classification. *Genome Res.* *25*, 305–315.
7. Miller, D.T., Lee, K., Abul-Husn, N.S., Amendola, L.M., Brothers, K., Chung, W.K., Gollob, M.H., Gordon, A.S., Harrison, S.M., Hershberger, R.E., et al. (2022). ACMG SF v3.1 list for reporting of secondary findings in clinical exome and genome sequencing: A policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet. Med.* *24*, 1407–1414.
8. Mallett, A., Kearey, P., Cameron, A., Healy, H., Denaro, C., Thomas, M., Lee, V.W., Stark, S., Fuller, M., and Hoy, W.E. (2020). The Ckd. Qld fabRy Epidemiology (aCQuiRE) study protocol: identifying the prevalence of Fabry disease amongst patients with kidney disease in Queensland, Australia. *BMC Nephrol.* *21*, 58.
9. Gilchrist, M., Casanova, F., Tyrrell, J.S., Cannon, S., Wood, A.R., Fife, N., Young, K., Oram, R.A., and Weedon, M.N. (2023). Prevalence of Fabry disease-causing variants in the UK Biobank. *J. Med. Genet.* *60*, 391–396. <https://doi.org/10.1136/jmg-2022-108523>.
10. Caldovic, L., Abdikarim, I., Narain, S., Tuchman, M., and Morizono, H. (2015). Genotype-Phenotype Correlations in Ornithine Transcarbamylase Deficiency: A Mutation Update. *J Genet Genomics* *42*, 181–194.
11. Blout Zawatsky, C.L., Shah, N., Machini, K., Perez, E., Christensen, K.D., Zouk, H., Steeves, M., Koch, C., Uveges, M., Shea, J., et al. (2021). Returning actionable genomic results in a research biobank: Analytic validity, clinical implementation, and resource utilization. *Am. J. Hum. Genet.* *108*, 2224–2237. <https://doi.org/10.1016/j.ajhg.2021.10.005>.
12. Kalia, S.S., Adelman, K., Bale, S.J., Chung, W.K., Eng, C., Evans, J.P., Herman, G.E., Hufnagel, S.B., Klein, T.E., Korf, B.R., et al. (2017). Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet. Med.* *19*, 249–255.
13. Verma, A., Damrauer, S.M., Naseer, N., Weaver, J., Kripke, C.M., Guare, L., Sirugo, G., Kember, R.L., Drivas, T.G., Dudek, S.M., et al. (2022). The Penn Medicine BioBank: Towards a Genomics-Enabled Learning Healthcare System to Accelerate Precision Medicine in a Diverse Population. *J. Personalized Med.* *12*, 1974. <https://doi.org/10.3390/jpm12121974>.
14. Batshaw, M.L., Tuchman, M., Summar, M., Seminara, J.; and Members of the Urea Cycle Disorders Consortium (2014). Members of the Urea Cycle Disorders Consortium. A longitudinal study of urea cycle disorders. *Mol. Genet. Metabol.* *113*, 127–130.
15. Inborn Metabolic Diseases (2022). Springer Berlin Heidelberg <https://link.springer.com/book/10.1007/978-3-662-63123-2>.
16. Hsu, T.R., and Niu, D.M. (2018). Fabry disease: Review and experience during newborn screening. *Trends Cardiovasc. Med.* *28*, 274–281.
17. Bireley, W.R., Van Hove, J.L.K., Gallagher, R.C., and Fenton, L.Z. (2012). Urea cycle disorders: brain MRI and neurological outcome. *Pediatr. Radiol.* *42*, 455–462.
18. Hentzen, D., Pelet, A., Feldman, D., Rabier, D., Berthelot, J., and Munnich, A. (1991). Fatal hyperammonemia resulting from a C-to-T mutation at a MspI site of the ornithine transcarbamylase gene. *Hum. Genet.* *88*, 153–156.
19. García-Pérez, M.A., Climent, C., Briones, P., Vilaseca, M.A., Rodés, M., and Rubio, V. (1997). Missense mutations in codon 225 of ornithine transcarbamylase (OTC) result in decreased amounts of OTC protein: a hypothesis on the molecular mechanism of the OTC deficiency. *J. Inherit. Metab. Dis.* *20*, 769–777.
20. Serrano, M., Martins, C., Pérez-Dueñas, B., Gómez-López, L., Murgui, E., Fons, C., García-Cazorla, A., Artuch, R., Jara, F., Aranz, J.A., et al. (2010). Neuropsychiatric manifestations in late-onset urea cycle disorder patients. *J. Child Neurol.* *25*, 352–358.

**HGGA, Volume 4**

**Supplemental information**

**Phenotypes of undiagnosed adults**

**with actionable *OTC* and *GLA* variants**

**Jessica I. Gold, Sarina Madhavan, Joseph Park, Hana Zouk, Emma Perez, Alanna Strong, Theodore G. Drivas, Amel Karaa, Marc Yudkoff, Daniel Rader, Regeneron Genetics Center, Penn Medicine BioBank, Robert C. Green, and Nina B. Gold**

## Supplemental materials

### Table of contents

**Table S1.** Regeneron Genetics Center author information.

**Table S2.** Penn Medicine Biobank author information.

**Table S3.** RedCAP data capture form for Fabry disease.

**Table S4.** RedCAP data capture form for OTC deficiency.

**Table S5.** Table of participant demographics.

**Table S6.** American College of Medical Genetics variant classification criteria for *GLA* and *OTC* variants found in biobank participants.

**Table S1.** Regeneron Genetics Center author information.

Regeneron Genetics Center Banner Author List and Contribution Statements

**RGC Management and Leadership Team**

Goncalo Abecasis, PhD , Aris Baras, M.D. , Michael Cantor, M.D. , Giovanni Coppola, M.D. , Andrew Deubler , Aris Economides, Ph.D. , Luca A. Lotta, M.D., Ph.D. , John D. Overton, Ph.D. , Jeffrey G. Reid, Ph.D. , Katherine Siminovitch, M.D. , Alan Shuldiner, M.D.

**Sequencing and Lab Operations**

Christina Beechert , Caitlin Forsythe, M.S. , Erin D. Fuller , Zhenhua Gu, M.S. , Michael Lattari , Alexander Lopez, M.S. , John D. Overton, Ph.D. , Maria Sotiropoulos Padilla, M.S. , Manasi Pradhan, M.S. , Kia Manoochehri, B.S. , Thomas D. Schleicher, M.S. , Louis Widom , Sarah E. Wolf, M.S. , Ricardo H. Ulloa, B.S.

**Clinical Informatics**

Amelia Averitt, Ph.D. , Nilanjana Banerjee, Ph.D. , Michael Cantor, M.D. , Dadong Li, Ph.D. , Sameer Malhotra, M.D. , Deepika Sharma, MHI , Jeffrey Staples , Ph.D.

**Genome Informatics**

Xiaodong Bai, Ph.D. , Suganthi Balasubramanian, Ph.D. , Suying Bao, Ph.D. , Boris Boutkov, Ph.D. , Siying Chen, Ph.D. , Gisu Eom, B.S. , Lukas Habegger, Ph.D. , Alicia Hawes, B.S. , Shareef Khalid , Olga Krasheninina, M.S. , Rouel Lanche, B.S. , Adam J. Mansfield, B.A. , Evan K. Maxwell, Ph.D. , George Mitra, B.A. , Mona Nafde, M.S. , Sean O'Keeffe, Ph.D. , Max Orelus, B.B.A. , Razvan Panea, Ph.D. , Tommy Polanco, B.A. , Ayesha Rasool, M.S. , Jeffrey G. Reid, Ph.D. , William Salerno, Ph.D. , Jeffrey C. Staples, Ph.D. , Kathie Sun, Ph.D.

**Analytical Genomics and Data Science**

Goncalo Abecasis, D.Phil. , Joshua Backman, Ph.D. , Amy Damask, Ph.D. , Lee Dobbyn, Ph.D. , Manuel Allen Revez Ferreira, Ph.D. , Arkopravo Ghosh, M.S. , Christopher Gillies, Ph.D. , Lauren Gurski, B.S. , Eric Jorgenson, Ph.D. , Hyun Min Kang, Ph.D. , Michael Kessler, Ph.D. , Jack Kosmicki, Ph.D. , Alexander Li , Ph.D. , Nan Lin, Ph.D. , Daren Liu, M.S. , Adam Locke, Ph.D. , Jonathan Marchini, Ph.D. , Anthony Marcketta, M.S. , Joelle Mbatchou, Ph.D. , Arden Moscati, Ph.D. , Charles Paulding, Ph.D. , Carlo Sidore, Ph.D. , Eli Stahl, Ph.D. , Kyoko Watanabe, Ph.D. , Bin Ye, Ph.D. , Blair Zhang, Ph.D. , Andrey Ziyatdinov, Ph.D.

**Therapeutic Area Genetics**

Ariane Ayer, B.S. , Aysegul Guvenek, Ph.D. , George Hindy, Ph.D. , Giovanni Coppola, M.D. , Jan Freudenberg, M.D. , Jonas Bovijn M.D. , Katherine Siminovitch, M.D. , Kavita Praveen, Ph.D. , Luca A. Lotta, M.D. , Manav Kapoor, Ph.D. , Mary Haas, Ph.D. , Moeen Riaz , Ph.D. , Niek Verweij, Ph.D. , Olukayode Sosina, Ph.D. , Parsa Akbari, Ph.D. , Priyanka Nakka, Ph.D. , Sahar Gelfman, Ph.D. , Sujit Gokhale, B.E. , Tanima De, Ph.D. , Veera Rajagopal, Ph.D. , Alan Shuldiner, M.D. , Bin Ye, Ph.D. , Gannie Tzoneva, Ph.D. , Juan Rodriguez-Flores, Ph.D.

**Research Program Management & Strategic Initiatives**

Esteban Chen, M.S. , Marcus B. Jones, Ph.D. , Michelle G. LeBlanc, Ph.D. , Jason Mighty, Ph.D. , Lyndon J. Mitnaul, Ph.D. , Nirupama Nishtala, Ph.D. , Nadia Rana, Ph.D. , Jaimee Hernandez



**Table S2.** Penn Medicine Biobank author information.

## **Penn Medicine BioBank Banner Author List and Contribution Statements**

### **PMBB Leadership Team**

Daniel J. Rader, M.D., Marylyn D. Ritchie, Ph.D., Michael D. Feldman M.D.

Contribution: All authors contributed to securing funding, study design and oversight. All authors reviewed the final version of the manuscript.

### **Patient Recruitment and Regulatory Oversight**

JoEllen Weaver, Nawar Naseer, Ph.D., M.P.H., Afiya Poindexter, Ashlei Brock, Khadijah Hu-Sain, Yi-An Ko

Contributions: JW manages patient recruitment and regulatory oversight of study. NN manages participant engagement, assists with regulatory oversight, and researcher access. AP, AB, KH, YK perform recruitment and enrollment of study participants.

### **Lab Operations**

JoEllen Weaver, Meghan Livingstone, Fred Vadivieso, Ashley Kloter, Stephanie DerOhannessian, Teo Tran, Linda Morrel, Ned Haubein, Joseph Dunn

Contribution: JW, ML, FV, SD conduct oversight of lab operations. ML, FV, AK, SD, TT, LM perform sample processing. NH, JD are responsible for sample tracking and the laboratory information management system.

### **Clinical Informatics**

Anurag Verma, Ph.D., Colleen Morse, M.S., Marjorie Risman, M.S., Renae Judy, B.S.

Contribution: All authors contributed to the development and validation of clinical phenotypes used to identify study subjects and (when applicable) controls.

### **Genome Informatics**

Anurag Verma Ph.D., Shefali S. Verma, Ph.D., Yuki Bradford, M.S., Scott Dudek, M.S., Theodore Drivas, M.D., PH.D.

Contribution: A.V., S.S.V. are responsible for the analysis, design, and infrastructure needed to quality control genotype and exome data. Y.B. performs the analysis. T.D. and A.V. provides variant and gene annotations and their functional interpretation of variants.

# Fabry Chart Review

---

Record ID

---

---

Variant

---

---

MRN

---

---

Sex

- Male  
 Female

---

Age

---

---

Race

- American Indian or Alaskan Native  
 Asian Indian  
 Black or African American  
 East Asian  
 Native Hawaiian or Pacific Islander  
 Other  
 White

---

Ethnicity

- Hispanic  
 Not Hispanic

---

Diagnosed with Fabry?

- Yes  
 No

---

Age at diagnosis

---

---

Method of Dx

---

---

Seeing genetics?

- Yes  
 No

---

Receiving treatment for Fabry?

- Yes  
 No

---

What treatment?

---

---

Problem list

---

---

Family history

---

---

Current medications

---

**Specialists seen**

	Yes	No
Ophtho	<input type="radio"/>	<input type="radio"/>
Renal	<input type="radio"/>	<input type="radio"/>
Neuro	<input type="radio"/>	<input type="radio"/>
Cardiology	<input type="radio"/>	<input type="radio"/>
Dermatology	<input type="radio"/>	<input type="radio"/>

---

Last ophtho A/P

---

---

Last renal A/P

---

---

Last neuro A/P

---

---

Last cardiology A/P

---

---

Last derm a/p

---

---

Other Specialists seen

---

**Presence of Fabry symptoms**

	Yes	No
distal extremity pain/discomfort	<input type="radio"/>	<input type="radio"/>
Neuropathic pain	<input type="radio"/>	<input type="radio"/>
Exercise Intolerance	<input type="radio"/>	<input type="radio"/>
Cold Intolerance	<input type="radio"/>	<input type="radio"/>
Heat Intolerance	<input type="radio"/>	<input type="radio"/>
Hypohydrosis	<input type="radio"/>	<input type="radio"/>
Corneal problems	<input type="radio"/>	<input type="radio"/>
Proteinuria	<input type="radio"/>	<input type="radio"/>

Chronic Kidney disease	<input type="radio"/>	<input type="radio"/>
s/p renal transplant	<input type="radio"/>	<input type="radio"/>
requiring dialysis	<input type="radio"/>	<input type="radio"/>
GI symptoms	<input type="radio"/>	<input type="radio"/>
Angiokeratomas	<input type="radio"/>	<input type="radio"/>
Cardiomyopathy	<input type="radio"/>	<input type="radio"/>
Aortic disease	<input type="radio"/>	<input type="radio"/>
Cardiac valve disease	<input type="radio"/>	<input type="radio"/>
Arrhythmia	<input type="radio"/>	<input type="radio"/>
Heart failure	<input type="radio"/>	<input type="radio"/>
Myocardial infarction	<input type="radio"/>	<input type="radio"/>
s/p stent placement or CABG	<input type="radio"/>	<input type="radio"/>
s/p ICD or pacemaker placement	<input type="radio"/>	<input type="radio"/>
s/p heart transplant or LVAD placement	<input type="radio"/>	<input type="radio"/>
Stroke	<input type="radio"/>	<input type="radio"/>
Vertigo	<input type="radio"/>	<input type="radio"/>
TIA	<input type="radio"/>	<input type="radio"/>
hearing loss	<input type="radio"/>	<input type="radio"/>

---

Cardiac Testing

- EKG
- Echocardiogram
- Stress Test
- Cardiac MRI
- CTA chest
- Cardiac biopsy
- Troponin
- BNP
- Lipid panel
- Other

---

Results of last EKG

\_\_\_\_\_

---

Results of last echocardiogram

\_\_\_\_\_

---

Results of last stress test

\_\_\_\_\_

---

Results of last cardiac MRI

\_\_\_\_\_

---

Results of last CTA chest

\_\_\_\_\_

---

Troponin results

---

---

BNP levels (maximum and most recent)

---

---

Last lipid panel results

---

---

Cardiac Biopsy

---

---

Other cardiac results: what test and results

---

---

Renal testing

- UA
- Urine protein
- Urine microalbumin
- BMP
- Cystatin C
- Renal Ultrasound
- Renal Biopsy
- Other

---

UA results

---

---

Urine protein/creatinine ratio

---

---

Urine microalbumin results

---

---

Last serum Cr

---

---

Max serum Cr

---

---

Current GFR

---

---

Cystatin C results

---

---

Renal Ultrasound results

---

---

Renal Biopsy results

---

---

Other renal testing: what and results?

---

---

Neurological Testing

- CT Brain
- MRI Brain
- MRA Head/Neck
- Carotid Ultrasound
- EMG/NCV
- Other Ultrasound (I.e. carpal tunnel)
- Other

---

CT head results

---

---

MRI Brain results

---

---

MRA head/neck results

---

---

Carotid Ultrasound results

---

---

EMG/NCV results

---

---

Other ultrasound: what ultrasound and results

---

---

Other neuro tests: what test and results

---

---

Other comments

---

# Otc Deficiency

Record ID

\_\_\_\_\_

Record ID

\_\_\_\_\_

## Basic Demographic Info

Variant

\_\_\_\_\_

MRN

\_\_\_\_\_

Biological sex

- Male  
 Female

Age

\_\_\_\_\_

Race

- American Indian or Alaskan Native  
 Asian Indian  
 Black or African American  
 East Asian  
 Native Hawaiian or Pacific Islander  
 White  
 Other  
 Unknown

Ethnicity

- Hispanic  
 Not Hispanic  
 Unknown

Diagnosed with OTC?

- Yes  
 No

Age at diagnosis

\_\_\_\_\_

Followed by medical genetics?

- Yes  
 No

Problem list

\_\_\_\_\_

Current medications

\_\_\_\_\_

**Family History**

Family history

---

**Family history suggestive of OTC deficiency**

	Yes	No
Known OTC deficiency	<input type="radio"/>	<input type="radio"/>
Urea cycle disorders	<input type="radio"/>	<input type="radio"/>
Unexplained newborn male death	<input type="radio"/>	<input type="radio"/>
Cerebral palsy NOS	<input type="radio"/>	<input type="radio"/>
Migraines / Headaches	<input type="radio"/>	<input type="radio"/>
Psychiatric conditions	<input type="radio"/>	<input type="radio"/>
Recurrent vomiting	<input type="radio"/>	<input type="radio"/>
Other decompensation during hospitalization / stress	<input type="radio"/>	<input type="radio"/>

**Pregnancy History**

Ever pregnant?  Yes  No

Living children. List as (#M, #F)

---

**Other pregnancy history suggestive of OTC deficiency**

	Yes	No
Spontaneous abortion	<input type="radio"/>	<input type="radio"/>
Had newborn male child who died	<input type="radio"/>	<input type="radio"/>
Had male child with sepsis in first week of life	<input type="radio"/>	<input type="radio"/>
Had male child with failure to thrive in first week of life	<input type="radio"/>	<input type="radio"/>
Had male child with somnolence in first week of life	<input type="radio"/>	<input type="radio"/>
Had a male child with unexplained tachypnea in first week of life	<input type="radio"/>	<input type="radio"/>
elective abortion	<input type="radio"/>	<input type="radio"/>

Other pregnancy history

---



**Specialists seen**

	Yes	No
Psychiatry	<input type="radio"/>	<input type="radio"/>
Neurology	<input type="radio"/>	<input type="radio"/>
OB/GYN	<input type="radio"/>	<input type="radio"/>
GI	<input type="radio"/>	<input type="radio"/>
Nutrition	<input type="radio"/>	<input type="radio"/>

Last Neurology A/P \_\_\_\_\_

Last Psychiatry A/P \_\_\_\_\_

Last OB/GYN A/P \_\_\_\_\_

Last GI A/P \_\_\_\_\_

Last Nutrition A/P \_\_\_\_\_

**Presence of OTC deficiency symptoms**

	Yes	No
Delirium	<input type="radio"/>	<input type="radio"/>
Encephalopathy or Altered Mental Status	<input type="radio"/>	<input type="radio"/>
Psychosis or erratic behavior NOS	<input type="radio"/>	<input type="radio"/>
Anxiety/Depression	<input type="radio"/>	<input type="radio"/>
Recurrent vomiting	<input type="radio"/>	<input type="radio"/>
Headaches / Migraine headaches	<input type="radio"/>	<input type="radio"/>
Seizure	<input type="radio"/>	<input type="radio"/>
Allergy to protein / protein intolerance	<input type="radio"/>	<input type="radio"/>
HCC	<input type="radio"/>	<input type="radio"/>
Reye-like syndrome	<input type="radio"/>	<input type="radio"/>
Executive function deficits	<input type="radio"/>	<input type="radio"/>
Mild cognitive impairment	<input type="radio"/>	<input type="radio"/>

Self-restriction of protein / Vegetarian diet / Avoidance of milk, red meat, eggs, high-protein foods (list specific food restrictions) \_\_\_\_\_

Headache description (From most recent PCP A/P Problem list)

\_\_\_\_\_

**Presence of OTC deficiency symptoms during precipitating events or catabolic stress**

	Altered Mental Status	Vomiting	Headache
Pregnancy and/or Delivery	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Systemic corticosteroid administration	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Infection	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Trauma (e.g. motor vehicle collision)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cancer therapy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Meat/protein ingestion	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Labs/Imaging**

Neurologic testing

- EEG
- Head CT
- Head MRI
- Other neuropsych testing

Results of any abnormal EEG

\_\_\_\_\_

Head CT Results

\_\_\_\_\_

Head MRI Results

\_\_\_\_\_

What other neuropsych testing and results

\_\_\_\_\_

Other imaging

- Abdominal US
- CT A/P
- Other Liver imaging

Results of last abdominal U/S

\_\_\_\_\_

Results of last CT A/P (Liver Findings and Impression)

\_\_\_\_\_

Other liver imaging: what test and results

\_\_\_\_\_

---

Other Labs

- Ammonia (serum)
- Blood gas
- LFTs
- PT/PTT
- Urine studies

---

Serum ammonia levels with dates

\_\_\_\_\_

---

ABG showing respiratory alkalosis?

- Yes
- No

---

Last AST/ALT

\_\_\_\_\_

---

Last PT/PTT

\_\_\_\_\_

---

Elevated urine orotic acid

---

Other comments

\_\_\_\_\_

**Table S5.** Table of participant demographics.

	<i>OTC</i> PLPV	<i>GLA</i> PLPV
Total	3	3
Female, n (%)	3 (100)	1 (33.3)
Age, Mean $\pm$ SD	51 (11.8)	54.3 (17.6)
Race/Ethnicity, n (%)		
Asian	0 (0)	0 (0)
Black	1 (33.3)	0 (0)
Other	0 (0)	0 (0)
White, Hispanic	0 (0)	0 (0)
White, Non-Hispanic	2 (66.7)	3 (100)

**Table S6.** American College of Medical Genetics variant classification criteria for *GLA* and *OTC* variants found in biobank participants.

Gene	DNA	Protein	Classification	Classification Criteria				Allelic fraction
				Very Strong	Strong	Moderate	Supporting	
<i>GLA</i>	c.335G>A	p.Arg112His	Pathogenic		PS3, PS4	PM2, PM5	PP1, PP3	1.00 (male)
<i>GLA</i>	c.647A>G	p.Tyr216Cys	Likely Pathogenic		PS4	PM3	PP1, PP3	1.00 (male)
<i>GLA</i>	c.868A>C	p.Met290Leu	Likely Pathogenic		PS4	PM2	PM2, PP1, PP4	0.56 (female)
<i>OTC</i>	c.118C>T	p.Arg40Cys	Likely pathogenic		PS3	PM2, PM5	PP!	0.50 (female)
<i>OTC</i>	c.274C>T	p.Arg92*	Pathogenic	PVS1	PS3	PM2	PP3	0.48 (female)
<i>OTC</i>	c.583G>A	p.Gly195Arg	Pathogenic		PS3	PM2, PM5	PP3, PP5	0.65 (female)