Received: 26 July 2021 | Accepted: 2 August 2021

DOI: 10.1002/ajh.26323

CORRESPONDENCE

The International Hemoglobinopathy Research Network (INHERENT): An international initiative to study the role of genetic modifiers in hemoglobinopathies

To the Editor:

Hemoglobinopathies, including sickle cell disease (SCD) and thalassemia syndromes, represent the commonest monogenic diseases in the world. Although their pathogenesis is established, the diverse clinical manifestations and the varying degree of severity are less understood and are thought to be governed, in part, by genetic modifiers. Previous studies have demonstrated the role of genetic modifiers in different hemoglobinopathy phenotypes, with co-inheritance of α-thalassemia and higher levels of fetal hemoglobin (HbF) being the best characterized disease modifiers. Several genome-wide analyses have identified three major quantitative trait loci modulating HbF levels: a promoter variant on *HBG2* (XmnI-rs7482144), the *HBS1L-MYB* intergenic region and *BCL11A*, which together explain up to 50% of the genetic variation affecting HbF.[1,2] More recently, studies have identified genetic modifiers associated with laboratory and clinical markers of disease complications.[3,4] However, few of these modifiers have reached a level of clinical utility.

Importantly, most association studies in SCD have been restricted to patients that have not received disease-modifying therapies. Given that the influence of genetic disease modifiers may change with treatment, identification of genetic modifiers requires high-quality clinical, laboratory and treatment data to allow accurate genotype/phenotype correlation. Furthermore, with the emergence of novel targeted therapies for hemoglobinopathies, such as gene therapy, genetic modifiers can facilitate patient stratification and, also, influence the response to these treatments.

Currently, the ITHANET portal[5] manually curates around 800 disease-modifying variants in over 420 genomic locations. However, most of these variants have not been validated with confirmatory or large-scale studies, and across diverse ethnic populations. In addition, data from different studies are not frequently reproducible and their possible effect size remains unknown.[6] Most importantly, with most studies having a sample size of less than 2000 patients, it is not possible to identify genetic modifiers with high confidence. As a result, the translation of these results into clinical practice has been limited. There are currently very few established polygenic risk scores related to disease complications, severity, or response to treatment, that can be used as an evidence base to stratify disease and offer patient specific treatment regimens in hemoglobinopathy patients.[7]

A validated standard for data collection and phenotypic definitions is crucial for the accurate comparison and pooling of data. The recently developed Sickle Cell Disease Ontology[8] represents a positive step towards disease-specific standardization that can facilitate integration of datasets in the field. In parallel, other ongoing initiatives, such as patient registries by RADeep and SPARCO, are working on standardization of clinical data collection for hemoglobinopathies using well-established international standards, such as the Human Phenotype Ontology.[9] Despite these efforts, a common understanding and discussion among different initiatives is necessary to allow integration of data for large-scale clinical and genomic studies. Furthermore, there is a limited amount of high-throughput or genome-wide data available for further research, despite several genetic studies in the field. A large, international disease-specific data repository, compliant with the FAIR data principles,[10] would revolutionize research in the field of hemoglobinopathies towards evidence-based approaches that utilize data science and artificial intelligence.

In 2020, nine existing international or regional consortia recognized the need for global synergies to address the above challenges and they agreed to create the International Hemoglobinopathy Research Network (INHERENT) as an umbrella network focused on the study of genetic modifiers of hemoglobinopathies. Specifically, INHERENT brings together the following consortia:

1. ITHANET (https://ithanet.eu/),
2. Rare Anemia Disorders European Epidemiological Platform (RADeep; https://www.radeepnetwork.eu/),
3. African Research and Innovative Initiative for Sickle Cell Education (ARISE; https://www.ariseinitiative.org/),
4. Sickle Pan-African Research Consortium (SPARCO),
5. Sickle Africa Data Coordinating Center (SADaCC; https://sickleinafrica.org/),
6. Réseau d'Etude de la Drépanocytose en Afrique Centrale (REDAC; https://redacnetwork.org/),
7. Human Variome Project Global Globin Network (HVP GGN; https://www.humanvariomeproject.org/gg2020),
8. International Health Repository (IHR),
9. ClinGen Hemoglobinopathy Variant Curation Expert Panel (Hemoglobinopathy VCEP; https://clinicalgenome.org/affiliation/50052/)

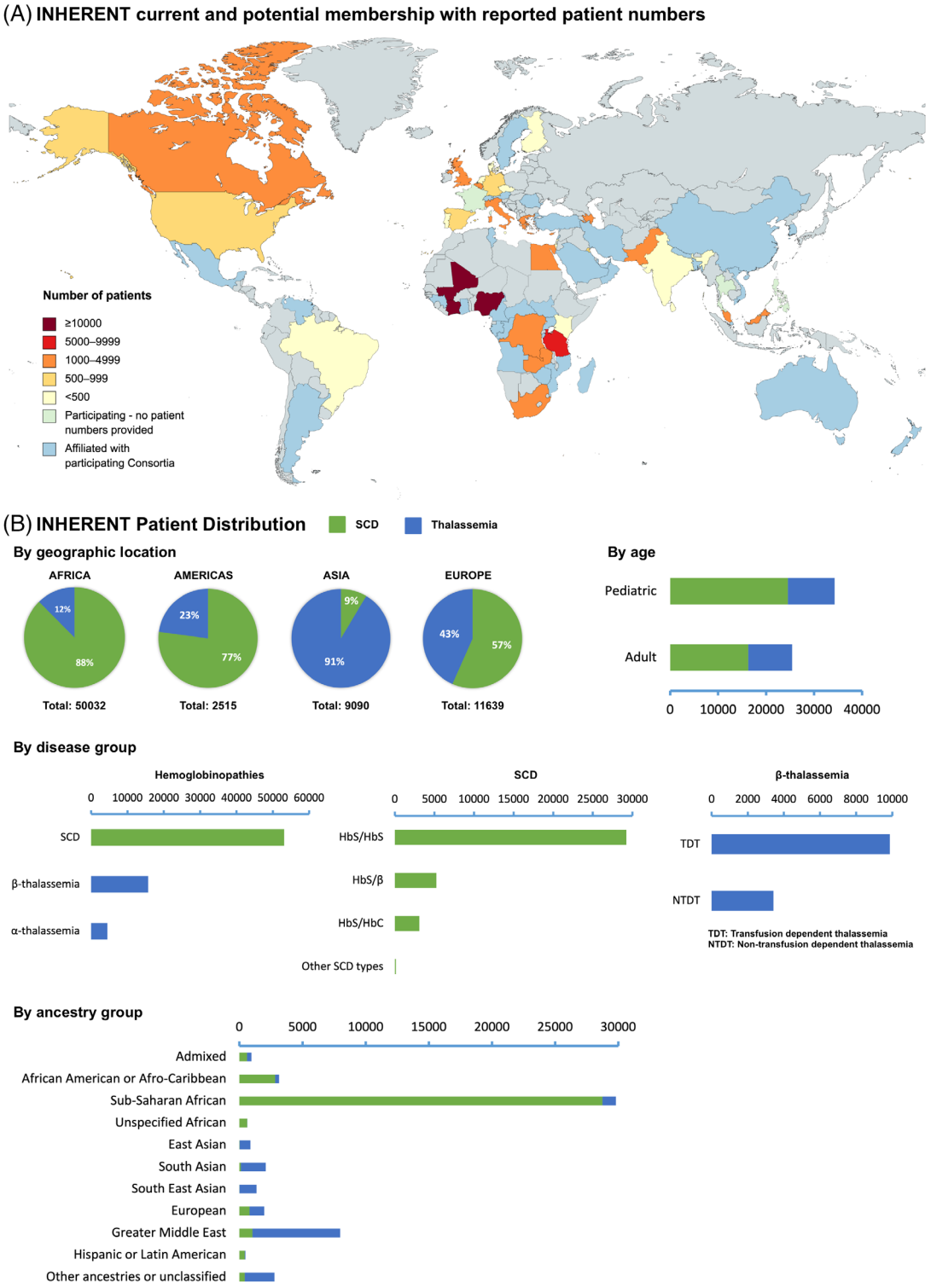

FIGURE 1 An overview of the INHERENT and the reported data by its current membership. (A) geographic distribution of the current and and potential membership of INHERENT, illustrating the reported number of patients per country. (B) Number of patients reported by INHERENT members, distributed by geographic location, age, disease group, and self-reported ancestry group

INHERENT is also endorsed by the European Reference Network on rare hematological diseases, ERN-EuroBloodNet.

The primary aim of INHERENT is to study the role of genetic modifiers in hemoglobinopathies through a large-scale, multi-ethnic genome-wide association study (GWAS). Therefore, INHERENT will address challenges of previous studies related to small sample sizes and low statistical power, while promoting participation of diverse populations worldwide. Specifically, INHERENT aims to: (a) discover new genetic modifiers of hemoglobinopathies, (b) validate previously reported genetic modifiers, (c) pool and analyze existing genomic data,

(d) standardize phenotypic descriptions using established standards, aligned with international recommendations, (e) develop a case report form (CRF) to efficiently gather sufficient high-quality data across countries accounting for different resource capabilities, and (f) develop a research resource of disease-specific data generated in INHERENT, including genomic, phenotypic, and functional data.

All INHERENT members agreed on the participation criteria, which are aimed at being inclusive and straightforward. Hence, participation in INHERENT is open for any group that can submit a minimum of 30 DNA samples with their core phenotypic description. Additional members that can significantly contribute to specific network activities, such as bioinformatics and biostatistical analyses, data management, genotyping, and regulatory/ethical issues, have also been invited to join the network. As a result, the current INHERENT membership is both international and interdisciplinary and includes over 160 experts from 89 organizations, spanning 36 countries worldwide (Table S1). Notably, based on the current membership of participating consortia, the projected membership of INHERENT can reach 73 countries, as shown in Figure 1A.

The target sample size of INHERENT is at least 30 000 individuals with hemoglobinopathies from diverse ancestries and geographic locations. Both pediatric and adult patients will be enrolled to study disease complications across all ages. To investigate the feasibility of the network's primary goal to perform a large-scale, multi-ethnic GWAS, we performed a survey among all centers that participate in INHERENT. The survey requested the number of patients per center and their distribution per disease group, age, and ancestry group. A total of 81 participating centers (91% of all members) responded to the survey, with the results illustrated in Figure 1. Through its current membership, INHERENT has the potential to enroll over 73 200 individuals with hemoglobinopathies. Specifically, 53 100 people with SCD have been reported by the participating centers, with 54.8% being homozygous S/S, while other major disease groups (S/β and S/C) are being well-represented. In addition, around 15 600 people with β-thalassemia have been reported, with over 62% of them being transfusion-dependent, and 4400 people with α-thalassemia. Basic genetic data are available for fewer than 45% of the patients, highlighting the need for genomic research in the field. Importantly, the diverse spectrum of ancestry groups for different disease groups clearly reflects the existing knowledge about the global hemoglobinopathy epidemiology, thus further supporting the multi-ethnic nature of the project. Some of the requested information was not readily available in the centers, and for large centers without an existing electronic patient registry the responses were approximations. Nevertheless, the results of the survey clearly support that the minimum target sample size set by INHERENT is realistic, highlighting the network's potential to be the largest international network on hemoglobinopathies.

INHERENT is in a unique position to address limitations of previous GWAS studies in hemoglobinopathies by increasing the sample size and achieving sufficient power to detect associations at GWAS significant levels. The Figure S1 illustrates different hypothetical, yet realistic, scenarios for the analyses to be performed to estimate the power under different assumptions. Panels A, B, and C refer to the evaluation of binary phenotypes using logistic regression with a P value of 5×10^{-8}. Based on the survey results, the assumed total sample size of 10 000 is realistic for both SCD and β-thalassemia, while the case rate of 0.3 is applicable for several hemoglobinopathy complications, such as the prevalence of stroke[11] in SCD, and osteoporosis in β-thalassemia.[12] We show that we have the statistical power to detect associations down to a minor allele frequency (MAF) of 0.05, with an Odds Ratio of 1.5, while detection of associations for less common disease complications is also possible. Similarly for the analyses using survival endpoints (Figure S1, panels D–F), we have sufficient power to detect associations even with 5000 samples, assuming an event rate of 30%, hazard ratio of 1.5, and MAF greater than 0.05.

INHERENT is coordinated by a Steering and Data Access Committee, which provides direction to the network and includes the working group (WG) chairs and one representative from each of the participating consortia. To achieve the main objectives, the activities of the network have been divided into five WGs: clinical, genotyping, data management and analysis, ethics, and knowledge translation.

In conclusion, by bringing together existing consortia and additional partners throughout the world, INHERENT avoids duplication of efforts and is focused on integration and consolidation of evidence as well as the generation and analysis of novel and large datasets. The increased sample size and the diversity in the studied populations can lead to novel discoveries and translational impact in the field of hemoglobinopathies. Moreover, the interdisciplinary and international nature of the network can result in synergistic research studies that promote innovative use of the collected and generated data as well as novel methodological approaches.

ACKNOWLEDGMENT

The coordination of INHERENT, the ITHANET portal and the ClinGen Hemoglobinopathy VCEP are co-funded by the European Regional Development Fund and the Republic of Cyprus through the Research and Innovation Foundation (Project: EXCELLENCE/1216/256). The ARISE project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement No. 824021. INHERENT is endorsed by the European Reference Network on Rare Hematological Diseases (ERN-EuroBloodNet) "ERN-2016 – Framework Partnership Agreement 2017–2021." FPA 739541.

CONFLICT OF INTEREST

Kevin H.M. Kuo – Grants/Research Support: NHLBI/NIH: 1R33HL147845, Thalassemia Foundation Canada, Peter Munk Cardiac Centre, University of Toronto, Cincinnati Children's Hospital Medical Center, Canadian Hematology Society, Pfizer; Consultancy: Agios, Alexion, Apellis, Aruvant, Bluebirdbio, Celgene/BMS, Novartis, Pfizer; Chair of Data Safety Monitoring Board: Bioverativ/Sanofi (ZFN gene editing trial for sickle cell disease NCT03653247); Research collaboration: Phoenicia Biosciences; Lead investigator: BENEFiTS (Phoenicia), AG348-C-017 (Agios); Site-PI (in alphabetical order): ACE-536-B-THAL-001, ACE-536-LTFU (Celgene/BMS); AG348-C-010, AG348-C-018, AG-348-C-020, (Agios); GBT440-031, GBT440-034

(Global Blood Therapeutics); ARU-1801 (CCMHC, Aruvant); APL2-302 (Apellis); CSEG101A2301 (Novartis); CTX001-131 (Vertex); Co-investigator for: APL2-307 (Apellis), CTX001-111, CTX001-121 (Vertex).

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Petros Kountouris[1,2], Coralea Stephanou[1], Natasha Archer[3], Fedele Bonifazi[4], Viviana Giannuzzi[4], Kevin H. M. Kuo[5], Aurelio Maggio[6], Julie Makani[7], María del Mar Mañú-Pereira[8], Kyriaki Michailidou[1,2], Siana Nkya[7,9], Obiageli E. Nnodu[10], Sara Trompeter[11,12], Léon Tshilolo[13], Ambroise Wonkam[14], Bin Alwi Zilfalil[15], Baba P. D. Inusa[16], Marina Kleanthous[1,2] on behalf of the International Hemoglobinopathy Research Network (INHERENT)

[1]*The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus*
[2]*Cyprus School of Molecular Medicine, Nicosia, Cyprus*
[3]*Pediatric Hematology/Oncology, Dana-Farber/Boston Children's Cancer and Blood Disorders Center, Boston, Massachusetts, USA*
[4]*Fondazione per la Ricerca Farmacologica Gianni Benzi Onlus, Bari, Italy*
[5]*Division of Hematology, Department of Medicine, University of Toronto, Toronto, Ontario, Canada*
[6]*A.O. Villa Sofia-Cervello, Palermo, Italy*
[7]*Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania*
[8]*Translational Research in Child and Adolescent Cancer. Vall d'Hebron Institut de recerca, Barcelona, Spain*
[9]*Dar es Salaam University College of Education, Dar es Salaam, Tanzania*
[10]*Department of Haematology and Blood Transfusion, University of Abuja, Abuja, Nigeria*
[11]*University College London Hospitals NHS Trust, London, UK*
[12]*NHS Blood and Transplant, London, UK*
[13]*Centre Hospitalier Monkole and CEFA-Institut de Recherche Biomédicale, Kinshasa, Democratic Republic of Congo*
[14]*University of Cape Town, Cape Town, South Africa*
[15]*Universiti Sains Malaysia, Health Campus, Kelantan, Malaysia*
[16]*Paediatric Haematology, Evelina London Children's Hospital, Guy's & St. Thomas' Hospital NHS Foundation Trust, London, UK*

Correspondence
Petros Kountouris, The Cyprus Institute of Neurology and Genetics, 6 International Airport Avenue, 1683 Nicosia, Cyprus.
Email: petrosk@cing.ac.cy

The full membership of INHERENT is provided in the Table S1.

ORCID

Petros Kountouris https://orcid.org/0000-0003-2681-4355
Aurelio Maggio https://orcid.org/0000-0002-9601-900X
Léon Tshilolo https://orcid.org/0000-0002-6318-0227
Baba P. D. Inusa https://orcid.org/0000-0003-2643-765X

REFERENCES

1. Thein SL, Menzel S, Lathrop M, Garner C. Control of fetal hemoglobin: new insights emerging from genomics and clinical implications. *Hum Mol Genet*. 2009;18:R216-R223.
2. Galarneau G, Palmer CD, Sankaran VG, Orkin SH, Hirschhorn JN, Lettre G. Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation. *Nat Genet*. 2010;42:1049-1051.
3. Flanagan JM, Frohlich DM, Howard TA, et al. Genetic predictors for stroke in children with sickle cell anemia. *Blood*. 2011;117:6681-6684.
4. Jacob SA, Novelli EM, Isenberg JS, et al. Thrombospondin-1 gene polymorphism is associated with estimated pulmonary artery pressure in patients with sickle cell anemia. *Am J Hematol*. 2017;92:E31-E34.
5. Kountouris P, Lederer CW, Fanis P, Feleki X, Old J, Kleanthous M. IthaGenes: an interactive database for haemoglobin variations and epidemiology. *PLoS One*. 2014;9(7):e103020. https://doi.org/10.1371/journal.pone.0103020
6. Stephanou C, Tamana S, Minaidou A, Papasavva P, Kleanthous M, Kountouris P. Genetic modifiers at the crossroads of personalised medicine for haemoglobinopathies. *J Clin Med*. 2019;8:1927. https://doi.org/10.3390/jcm8111927
7. Rampersaud E, Kang G, Palmer LE, et al. A polygenic score for acute vaso-occlusive pain in pediatric sickle cell disease. *Blood Adv*. 2021;5:2839-2851.
8. Sickle Cell Disease Ontology Working Group. The Sickle Cell Disease Ontology: enabling universal sickle cell-based knowledge representation. *Database*. 2019;2019:baz118. https://doi.org/10.1093/database/baz118
9. Köhler S, Gargano M, Matentzoglu N, et al. The human phenotype ontology in 2021. *Nucleic Acids Res*. 2021;49:D1207-D1217.
10. Wilkinson MD, Dumontier M, Aalbersberg IJJ, et al. The FAIR guiding principles for scientific data management and stewardship. *Sci Data*. 2016;3:160018.
11. Ohene-Frempong K, Weiner SJ, Sleeper LA, et al. Cerebrovascular accidents in sickle cell disease: rates and risk factors. *Blood*. 1998;91:288-294.
12. Taher AT, Musallam KM, Karimi M, et al. Overview on practices in thalassemia intermedia management aiming for lowering complication rates across a region of endemicity: the OPTIMAL CARE study. *Blood*. 2010;115:1886-1892.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.