medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Novel risk loci for COVID-19 hospitalization among admixed American populations

Silvia Diz-de Almeida,[1,2,135] Raquel Cruz,[1,2,135] Andre D. Luchessi,[3] José M. Lorenzo-Salazar,[4] Miguel López de Heredia,[2] Inés Quintela,[5] Rafaela González-Montelongo,[4] Vivian N. Silbiger,[3] Marta Sevilla Porras,[2,6] Jair Antonio Tenorio Castaño,[2,6,7] Julian Nevado,[2,6,7] Jose María Aguado,[8,9,10,11] Carlos Aguilar,[12] Sergio Aguilera-Albesa,[13,14] Virginia Almadana,[15] Berta Almoguera,[16,2] Nuria Alvarez,[17] Álvaro Andreu-Bernabeu,[18,10] Eunate Arana-Arri,[19,20] Celso Arango,[18,21,10] María J. Arranz,[22] Maria-Jesus Artiga,[23] Raúl C. Baptista-Rosas,[24,25,26] María Barreda- Sánchez,[27,28] Moncef Belhassen-Garcia,[29,30] Joao F. Bezerra,[31] Marcos A.C. Bezerra,[32] Lucía Boix-Palop,[33] María Brion,[34,35] Ramón Brugada,[36,37,35,38] Matilde Bustos,[39] Enrique J. Calderón,[39,40,41] Cristina Carbonell,[42,30] Luis Castano,[19,43,2,44,45] Jose E. Castelao,[46] Rosa Conde-Vicente,[47] M. Lourdes Cordero-Lorenzana,[48] Jose L. Cortes-Sanchez,[49,50,] Marta Corton,[16,2] M. Teresa Darnaude,[51] Alba De Martino-Rodríguez,[52,53] Victor del Campo-Pérez,[54] Aranzazu Diaz de Bustamante,[51] Elena Domínguez-Garrido,[55] Rocío Eirós,[56] María Carmen Fariñas,[57,58,59] María J. Fernandez-Nestosa [60] Uxía Fernández-Robelo,[61] Amanda Fernández-Rodríguez,[62,11] Tania Fernández-Villa,[63,41] Manuela Gago-Domínguez,[5,64] Belén Gil-Fournier,[65] Javier Gómez-Arrue,[52,53] Beatriz González Álvarez,[52,53] Fernan Gonzalez Bernaldo de Quirós,[66] Anna González-Neira,[17] Javier González-Peñas,[18,10,21] Juan F. Gutiérrez-Bautista,[67] María José Herrero,[68,69]Antonio Herrero-Gonzalez,[70] María A. Jimenez-Sousa,[62,11] María Claudia Lattig,[71,72] Anabel Liger Borja,[73] Rosario Lopez-Rodriguez,[16,2,74] Esther Mancebo,[75,76] Caridad Martín-López,[73] Vicente Martín,[63 41] Oscar Martinez-Nieto,[77,72] Iciar Martinez-Lopez,[78,79] Michel F. Martinez-Resendez,[49] Ángel Martinez-Perez,[80] Juliana F. Mazzeu,[81,82,83] Eleuterio Merayo Macías,[84] Pablo Minguez,[16,2] Victor Moreno Cuerda,[85,86] Silviene F. Oliveira,[87,88,83,89] Eva Ortega-Paino,[23] Mara Parellada,[18,21,10] Estela Paz-Artal,[75,76,90] Ney P.C. Santos,[91] Patricia Pérez-Matute,[92] Patricia Perez,[93] M. Elena Pérez-Tomás,[27] Teresa Perucho,[94] Mel·lina Pinsach-Abuin,[36,35] Guillermo Pita,[17] Ericka N. Pompa-Mera,[95,96] Gloria L. Porras-Hurtado,[97] Aurora Pujol,[98,2,99] Soraya Ramiro León,[65] Salvador Resino,[62,11] Marianne R. Fernandes,[91,100] Emilio Rodríguez-Ruiz,[101,64] Fernando Rodriguez-Artalejo,[102,103,41,104] José A. Rodriguez-Garcia,[105] Francisco Ruiz-Cabello,[67,106,107] Javier Ruiz-Hornillos,[108,109,110] Pablo Ryan,[111,112,113,11] José Manuel Soria,[80] Juan Carlos Souto,[114] Eduardo Tamayo,[115,116,11] Alvaro Tamayo-Velasco,[117] Juan Carlos Taracido-Fernandez,[70] Alejandro Teper,[118] Lilian Torres-Tobar,[119] Miguel Urioste,[120] Juan Valencia-Ramos,[121] Zuleima Yáñez,[122] Ruth Zarate,[123] Itziar de Rojas,[124,125] Agustín Ruiz,[124,125] Pascual Sánchez,[126] Luis Miguel Real,[127] SCOURGE Cohort Group, Encarna Guillen-Navarro,[27,128,129,130] Carmen Ayuso,[16,2] Esteban Parra, [131] José A. Riancho,[57,58,59,2] Augusto Rojas-Martinez,[132] Carlos Flores,[4,133,134,136,*] Pablo Lapunzina,[2,6,7,136] Ángel Carracedo,[1,2,5,64,136,**]

Affiliations

[1], Centro Singular de Investigación en Medicina Molecular y Enfermedades Crónicas (CIMUS), Universidade de Santiago de Compostela, Santiago de Compostela, Spain

[2], CIBERER, ISCIII, Madrid, Spain

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[3], Universidade Federal do Rio Grande do Norte, Departamento de Analises Clinicas e Toxicologicas, Natal, Brazil

[4], Genomics Division, Instituto Tecnológico y de Energías Renovables, Santa Cruz de Tenerife, Spain

[5], Fundación Pública Galega de Medicina Xenómica, Sistema Galego de Saúde (SERGAS) Santiago de Compostela, Spain

[6], Instituto de Genética Médica y Molecular (INGEMM), Hospital Universitario La Paz-IDIPAZ, Madrid, Spain

[7], ERN-ITHACA-European Reference Network

[8], Unit of Infectious Diseases, Hospital Universitario 12 de Octubre, Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas12), Madrid, Spain

[9], Spanish Network for Research in Infectious Diseases (REIPI RD16/0016/0002), Instituto de Salud Carlos III, Madrid, Spain

[10], School of Medicine, Universidad Complutense, Madrid, Spain

[11], CIBERINFEC, ISCIII, Madrid, Spain

[12], Hospital General Santa Bárbara de Soria, Soria, Spain

[13], Pediatric Neurology Unit, Department of Pediatrics, Navarra Health Service Hospital, Pamplona, Spain

[14], Navarra Health Service, NavarraBioMed Research Group, Pamplona, Spain

[15], Hospital Universitario Virgen Macarena, Neumología, Seville, Spain

[16], Department of Genetics & Genomics, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[17], Spanish National Cancer Research Centre, Human Genotyping-CEGEN Unit, Madrid, Spain

[18], Department of Child and Adolescent Psychiatry, Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón (IiSGM), Madrid, Spain

[19], Biocruces Bizkai HRI, Bizkaia, Spain

[20], Cruces University Hospital, Osakidetza, Bizkaia, Spain

[21], Centre for Biomedical Network Research on Mental Health (CIBERSAM), Instituto de Salud Carlos III, Madrid, Spain

[22], Fundació Docència I Recerca Mutua Terrassa, Barcelona, Spain

[23], Spanish National Cancer Research Center, CNIO Biobank, Madrid, Spain

[24], Hospital General de Occidente, Zapopan Jalisco, Mexico

[25], Centro Universitario de Tonalá, Universidad de Guadalajara, Tonalá Jalisco, Mexico

[26], Centro de Investigación Multidisciplinario en Salud, Universidad de Guadalajara, Tonalá idisJalisco, Mexico

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[27], Instituto Murciano de Investigación Biosanitaria (IMIB-Arrixaca), Murcia, Spain

[28], Universidad Católica San Antonio de Murcia (UCAM), Murcia, Spain

[29], Hospital Universitario de Salamanca-IBSAL, Servicio de Medicina Interna-Unidad de Enfermedades Infecciosas, Salamanca, Spain

[30], Universidad de Salamanca, Salamanca, Spain

[31], Escola Tecnica de Saúde, Laboratorio de Vigilancia Molecular Aplicada, Brazil

[32], Federal University of Pernambuco, Genetics Postgraduate Program, Recife, PE, Brazil

[33], Hospital Universitario Mutua Terrassa, Barcelona, Spain

[34], Instituto de Investigación Sanitaria de Santiago (IDIS), Xenética Cardiovascular, Santiago de Compostela, Spain

[35], CIBERCV, ISCIII, Madrid, Spain

[36], Cardiovascular Genetics Center, Institut d'Investigació Biomèdica Girona (IDIBGI), Girona, Spain

[37], Medical Science Department, School of Medicine, University of Girona, Girona, Spain

[38], Hospital Josep Trueta, Cardiology Service, Girona, Spain

[39], Institute of Biomedicine of Seville (IBiS), Consejo Superior de Investigaciones Científicas (CSIC)- University of Seville- Virgen del Rocio University Hospital, Seville, Spain

[40], Departamento de Medicina, Hospital Universitario Virgen del Rocío, Universidad de Sevilla, Seville, Spain

[41], CIBERESP, ISCIII, Madrid, Spain

[42], Hospital Universitario de Salamanca-IBSAL, Servicio de Medicina Interna, Salamanca, Spain

[43], Osakidetza, Cruces University Hospital, Bizkaia, Spain

[44], Centre for Biomedical Network Research on Diabetes and Metabolic Associated Diseases (CIBERDEM), Instituto de Salud Carlos III, Madrid, Spain

[45], University of Pais Vasco, UPV/EHU, Bizkaia, Spain

[46], Oncology and Genetics Unit, Instituto de Investigacion Sanitaria Galicia Sur, Xerencia de Xestion Integrada de Vigo-Servizo Galego de Saúde, Vigo, Spain

[47], Hospital Universitario Río Hortega, Valladolid, Spain

[48], Servicio de Medicina intensiva, Complejo Hospitalario Universitario de A Coruña (CHUAC), Sistema Galego de Saúde (SERGAS), A Coruña, Spain

[49], Tecnológico de Monterrey, Monterrey, Mexico

[50], Otto von Guericke University, Departament of Microgravity and Translational Regenerative Medicine, Magdeburg, Germany

[51], Hospital Universitario Mostoles, Unidad de Genética, Madrid, Spain

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[52], Instituto Aragonés de Ciencias de la Salud (IACS), Zaragoza, Spain

[53], Instituto Investigación Sanitaria Aragón (IIS-Aragon), Zaragoza, Spain

[54], Preventive Medicine Department, Instituto de Investigacion Sanitaria Galicia Sur, Xerencia de Xestion Integrada de Vigo-Servizo Galego de Saúde, Vigo, Spain

[55], Unidad Diagnóstico Molecular. Fundación Rioja Salud, La Rioja, Spain

[56], Hospital Universitario de Salamanca-IBSAL, Servicio de Cardiología, Salamanca, Spain

[57], IDIVAL, Cantabria, Spain

[58], Universidad de Cantabria, Cantabria, Spain

[59], Hospital U M Valdecilla, Cantabria, Spain

[60], Universidad Nacional de Asunción, Facultad de Politécnica, Paraguay

[61], Urgencias Hospitalarias, Complejo Hospitalario Universitario de A Coruña (CHUAC), Sistema Galego de Saúde (SERGAS), A Coruña, Spain

[62], Unidad de Infección Viral e Inmunidad, Centro Nacional de Microbiología (CNM), Instituto de Salud Carlos III (ISCIII), Madrid, Spain

[63], Grupo de Investigación en Interacciones Gen-Ambiente y Salud (GIIGAS) - Instituto de Biomedicina (IBIOMED), Universidad de León, León, Spain

[64], IDIS

[65], Hospital Universitario de Getafe, Servicio de Genética, Madrid, Spain

[66], Ministerio de Salud Ciudad de Buenos Aires, Buenos Aires, Argentina

[67], Hospital Universitario Virgen de las Nieves, Servicio de Análisis Clínicos e Inmunología, Granada, Spain

[68], IIS La Fe, Plataforma de Farmacogenética, Valencia, Spain

[69], Universidad de Valencia, Departamento de Farmacología, Valencia, Spain

[70], Data Analysis Department, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[71], Universidad de los Andes, Facultad de Ciencias, Bogotá, Colombia

[72], SIGEN Alianza Universidad de los Andes - Fundación Santa Fe de Bogotá, Bogotá, Colombia

[73], Hospital General de Segovia, Medicina Intensiva, Segovia, Spain

[74], Facultad de Farmacia, Universidad San Pablo-CEU, CEU Universities, Urbanización Montepríncipe, 28660 Boadilla del Monte, España

[75], Hospital Universitario 12 de Octubre, Department of Immunology, Madrid, Spain

[76], Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas12), Transplant Immunology and Immunodeficiencies Group, Madrid, Spain

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[77], Fundación Santa Fe de Bogota, Departamento Patologia y Laboratorios, Bogotá, Colombia

[78], Unidad de Genética y Genómica Islas Baleares, Islas Baleares, Spain

[79], Hospital Universitario Son Espases, Unidad de Diagnóstico Molecular y Genética Clínica, Islas Baleares, Spain

[80], Genomics of Complex Diseases Unit, Research Institute of Hospital de la Santa Creu i Sant Pau, IIB Sant Pau, Barcelona, Spain

[81], Universidade de Brasília, Faculdade de Medicina, Brazil

[82], Programa de Pós-Graduação em Ciências Médicas (UnB), Brazil

[83], Programa de Pós-Graduação em Ciencias da Saude (UnB), Brazil

[84], Hospital El Bierzo, Unidad Cuidados Intensivos, León, Spain

[85], Hospital Universitario Mostoles, Medicina Interna, Madrid, Spain

[86], Universidad Francisco de Vitoria, Madrid, Spain

[87], Departamento de Genética e Morfologia, Instituto de Ciências Biológicas, Universidade de Brasília, Brazil

[88], Programa de Pós-Graduação em Biologia Animal (UnB), Brazil

[89], Programa de Pós-Graduação Profissional em Ensino de Biologia (UnB), Brazil

[90], Universidad Complutense de Madrid, Department of Immunology, Ophthalmology and ENT, Madrid, Spain

[91], Universidade Federal do Pará, Núcleo de Pesquisas em Oncologia, Belém, Pará, Brazil

[92], Infectious Diseases, Microbiota and Metabolism Unit, CSIC Associated Unit, Center for Biomedical Research of La Rioja (CIBIR), Logroño, Spain

[93], Inditex, A Coruña, Spain

[94], GENYCA, Madrid, Spain

[95], Instituto Mexicano del Seguro Social (IMSS), Centro Médico Nacional Siglo XXI, Unidad de Investigación Médica en Enfermedades Infecciosas y Parasitarias, Mexico City, Mexico

[96], Instituto Mexicano del Seguro Social (IMSS), Centro Médico Nacional La Raza. Hospital de Infectología, Mexico City, Mexico

[97], Clinica Comfamiliar Risaralda, Pereira, Colombia

[98], Bellvitge Biomedical Research Institute (IDIBELL), Neurometabolic Diseases Laboratory, L'Hospitalet de Llobregat, Spain

[99], Catalan Institution of Research and Advanced Studies (ICREA), Barcelona, Spain

[100], Hospital Ophir Loyola, Departamento de Ensino e Pesquisa, Belém, Pará, Brazil

[101], Unidad de Cuidados Intensivos, Hospital Clínico Universitario de Santiago (CHUS), Sistema Galego de Saúde (SERGAS), Santiago de Compostela, Spain

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[102], Department of Preventive Medicine and Public Health, School of Medicine, Universidad Autónoma de Madrid, Madrid, Spain

[103], IdiPaz (Instituto de Investigación Sanitaria Hospital Universitario La Paz), Madrid, Spain

[104], IMDEA-Food Institute, CEI UAM+CSIC, Madrid, Spain

[105], Complejo Asistencial Universitario de León, León, Spain

[106], Instituto de Investigación Biosanitaria de Granada (ibs GRANADA), Granada, Spain

[107], Universidad de Granada, Departamento Bioquímica, Biología Molecular e Inmunología III, Granada, Spain

[108], Hospital Infanta Elena, Allergy Unit, Valdemoro, Madrid, Spain

[109], Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[110], Faculty of Medicine, Universidad Francisco de Vitoria, Madrid, Spain

[111], Hospital Universitario Infanta Leonor, Madrid, Spain

[112], Complutense University of Madrid, Madrid, Spain

[113], Gregorio Marañón Health Research Institute (IiSGM), Madrid, Spain

[114], Haemostasis and Thrombosis Unit, Hospital de la Santa Creu i Sant Pau, IIB Sant Pau, Barcelona, Spain

[115], Hospital Clinico Universitario de Valladolid, Servicio de Anestesiologia y Reanimación, Valladolid, Spain

[116], Universidad de Valladolid, Departamento de Cirugía, Valladolid, Spain

[117], Hospital Clinico Universitario de Valladolid, Servicio de Hematologia y Hemoterapia, Valladolid, Spain

[118], Hospital de Niños Ricardo Gutierrez, Buenos Aires, Argentina

[119], Fundación Universitaria de Ciencias de la Salud

[120], Spanish National Cancer Research Centre, Familial Cancer Clinical Unit, Madrid, Spain

[121], University Hospital of Burgos, Burgos, Spain

[122], Universidad Simón Bolívar, Facultad de Ciencias de la Salud, Barranquilla, Colombia

[123], Centro para el Desarrollo de la Investigación Científica, Paraguay

[124 183], Research Center and Memory clinic, ACE Alzheimer Center Barcelona, Universitat Internacional de Catalunya, Spain

[125 184], Centre for Biomedical Network Research on Neurodegenerative Diseases (CIBERNED), Instituto de Salud Carlos III, Madrid, Spain

[126 185], CIEN Foundation/Queen Sofia Foundation Alzheimer Center, Madrid, Spain

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[127] [186], Hospital Universitario de Valme, Unidad Clínica de Enfermedades Infecciosas y Microbiología, Sevilla, Spain

[128] [124], Sección Genética Médica - Servicio de Pediatría, Hospital Clínico Universitario Virgen de la Arrixaca, Servicio Murciano de Salud, Murcia, Spain

[129] [125], Departamento Cirugía, Pediatría, Obstetricia y Ginecología, Facultad de Medicina, Universidad de Murcia (UMU), Murcia, Spain

[130] [126], Grupo Clínico Vinculado, Centre for Biomedical Network Research on Rare Diseases (CIBERER), Instituto de Salud Carlos III, Madrid, Spain

[131] [127] Department of Anthropology, University of Toronto at Mississauga, Mississauga, Ontario, Canada.

[132] [128], Tecnologico de Monterrey, Escuela de Medicina y Ciencias de la Salud, Monterrey, Mexico

[133] [129], Research Unit, Hospital Universitario Nuestra Señora de Candelaria, Santa Cruz de Tenerife, Spain

[134] [130], Centre for Biomedical Network Research on Respiratory Diseases (CIBERES), Instituto de Salud Carlos III, Madrid, Spain

[135] [131], These authors contributed equally: Silvia Diz-de Almeida, Raquel Cruz

[136] [132], These authors contributed equally: Carlos Flores, Pablo Lapunzina, Ángel Carracedo

*Correspondence: cflores@ull.edu.es

**Correspondence: angel.carracedo@usc.es

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

The genetic basis of severe COVID-19 has been thoroughly studied and many genetic risk factors shared between populations have been identified. However, reduced sample sizes from non-European groups have limited the discovery of population-specific common risk loci. In this second study nested in the SCOURGE consortium, we have conducted the largest GWAS meta-analysis for COVID-19 hospitalization in admixed Americans, comprising a total of 4,702 hospitalized cases recruited by SCOURGE and other seven participating studies in the COVID-19 Host Genetic Initiative. We identified four genome-wide significant associations, two of which constitute novel loci and first discovered in Latin-American populations (*BAZ2B* and *DDIAS*). A trans-ethnic meta-analysis revealed another novel cross-population risk locus in *CREBBP*. Finally, we assessed the performance of a cross-ancestry polygenic risk score in the SCOURGE admixed American cohort.

Introduction

To date, more than 50 loci associated to COVID-19 susceptibility, hospitalization, and severity have been identified using genome-wide association studies (GWAS)[1,2]. The COVID-19 Host Genetics Initiative (HGI) has made significant efforts[3] to augment the power to identify disease loci by recruiting individuals from diverse populations and conducting a trans-ancestry meta-analysis. Despite this, the lack of genetic diversity and a focus on cases of European ancestries still predominate in the studies[4,5]. Besides, while trans-ancestry meta-analyses are a powerful approach for discovering shared genetic risk variants with similar effects across populations[6], they may fail to identify risk variants that have larger effects on particular underrepresented populations. Genetic disease risk has been shaped by the particular evolutionary history of populations and the environmental exposures[7]. Their action is particularly important for infectious

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

diseases due to the selective constrains that are imposed by the host-pathogen interactions[8,9]. Literature examples of this in COVID-19 severity includes a *DOCK2* gene variant in East Asians[10], and frequent loss of function variants in *IFNAR1* and *IFNAR2* genes in Polynesian and Inuit populations, respectively[11,12].

Including diverse populations in case-control GWAS studies with unrelated participants usually require a prior classification of individuals in genetically homogeneous groups, which are typically analysed separately to control the population stratification effects[13]. Populations with recent admixture impose an additional challenge to the GWAS due to their complex genetic diversity and linkage disequilibrium (LD) patterns, requiring the development of alternative approaches and a careful inspection of results to reduce the false positives due to population structure[7]. In fact, there are benefits in study power from modelling the admixed ancestries either locally, at regional scale in the chromosomes, or globally, across the genome, depending on factors such as the heterogeneity of the risk variant in frequencies or the effects among the ancestry strata[14]. Despite the development of novel methods specifically tailored for the analysis of admixed populations[15], the lack of a standardized analysis framework and the difficulties to confidently cluster the admixed individuals into particular genetic groups often leads to their exclusion from GWAS.

The Spanish Coalition to Unlock Research on Host Genetics on COVID-19 (SCOURGE) recruited COVID-19 patients between March and December 2020 from hospitals across Spain and from March 2020 to July 2021 in Latin-America (https://www.scourge-covid.org). A first GWAS of COVID-19 severity among Spanish patients of European descent revealed novel disease loci and explored age and sex varying effects of the genetic factors[16]. Here we present the findings of a GWAS meta-analysis in admixed American (AMR) populations, comprising individuals from the

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

SCOURGE Latin-American cohort and the HGI studies, which allowed to identify two novel severe COVID-19 loci, *BAZ2B* and *DDIAS*. Further analyses modelling the admixture from three genetic ancestral components and performing a trans-ethnic meta-analysis led to the identification of an additional risk locus near *CREBBP*. We finally assessed a cross-ancestry polygenic risk score model with variants associated with critical COVID-19.

Results

Meta-analysis of COVID-19 hospitalization in admixed Americans

Study cohorts

Within the SCOURGE consortium, we included 1,608 hospitalized cases and 1,887 controls (not hospitalized COVID-19 patients) from Latin-American countries and from recruitments of individuals of Latin-American descend conducted in Spain (Supplementary Table 1). Quality control details and estimation of global genetic inferred ancestry (GIA) (supplementary Figure 1) are described in Methods, whereas clinical and demographic characteristics of patients included in the analysis are shown in Table 1. Summary statistics from the SCOURGE cohort were obtained under a logistic mixed model with the SAIGE model (Methods). Another seven studies participating in the COVID-19 HGI consortium were included in the meta-analysis of COVID-19 hospitalization in admixed Americans (Figure 1).

GWAS meta-analysis

We performed a fixed-effects GWAS meta-analysis using the inverse of the variance as weights for the overlapping markers. The combined GWAS sample size consisted of 4,702 admixed AMR hospitalized cases and 68,573 controls.

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

This GWAS meta-analysis revealed genome-wide significant associations at four risk loci (Table 2, Figure 2), two of which (*BAZ2B* and *DDIAS*) were novel discoveries. Variants of these loci were prioritized by positional and expression quantitative trait loci (eQTL) mapping with FUMA, identifying four lead variants linked to other 310 variants and 31 genes (Supplementary Tables 2-4). A gene-based association test revealed a significant association in *BAZ2B* and in previously known COVID-19 risk loci: *LZTFL1*, *XCR1*, *FYCO1*, *CCR9*, and *IFNAR2* (Supplementary Table 5).

Located within the gene *BAZ2B*, the sentinel variant rs13003835 is an intronic variant associated with an increased risk of COVID-19 hospitalization (Odds Ratio [OR]=1.20, 95% Confidence Interval [CI]=1.12-1.27, $p=3.66x10^{-8}$). This association was not previously reported in any GWAS of COVID-19 published to date. Interestingly, rs13003835 did not reach significance (p=0.972) in the COVID-19 HGI trans-ancestry meta-analysis including the five population groups[1]. Based on our mapping strategy (see Methods), we also prioritized *PLA2R1*, *LY75*, *WDSUB1*, and *CD302* in this locus.

The other novel risk locus is led by the sentinel variant rs77599934, a rare intronic variant located in chromosome 11 within *DDIAS* and associated with risk of COVID-19 hospitalization (OR=2.27, 95%CI=1.70-3.04, $p=2.26x10^{-8}$). The *PRCP* gene was an additional prioritized gene at this locus.

We also observed a suggestive association with rs2601183 in chromosome 15, which is located between *ZNF774* and *IQGAP1* (allele-G OR=1.20, 95%CI=1.12-1.29, $p=6.11x10^{-8}$, see Supplementary Table 2), which has not yet been reported in any other GWAS of COVID-19 to date. This sentinel variant is in perfect LD (r^2=1) with rs601183, an eQTL of *ZNF774* in the lung.

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The GWAS meta-analysis also pinpointed two significant variants at known loci, *LZTFL1* and *FOXP4*. The SNP rs35731912 was previously associated with COVID-19 severity in EUR populations[17], and it was mapped to *LZTFL1*. As for rs2477820, while it is a novel risk variant within gene *FOXP4*, it has a moderate LD (r^2=0.295) with rs2496644, which has been linked to COVID-19 hospitalization[18]. This is consistent with the effects of LD in tag-SNPs when conducting GWAS in diverse populations.

Functional mapping of novel risk variants

Bayesian fine mapping

We performed different approaches to narrow down the prioritized loci to a set of most probable genes driving the associations. First, we computed credible sets at the 95% confidence for causal variants and annotated them with VEP and the V2G aggregate scoring (Supplementary Table 6, Supplementary Figure 3). The 95% confidence credible set from the region of chromosome 2 around rs13003835 included 76 variants. However, the approach was unable to converge allocating variants in a 95% confidence credible set for the region in chromosome 11.

Colocalization of eQTLs

To determine if the novel genetic risk loci were associated with gene expression in relevant tissues (whole blood, lung, lymphocytes, and oesophagus mucosa), we computed the posterior probabilities (PP) of colocalization for overlapping variants allocated to the 95% confidence credible set. We used the GTEx v8 tissues as the main expression dataset, although it is important to consider that the eQTL associations were carried out mainly on individuals of EUR ancestries. To confirm the colocalization in other ancestries, we also performed analyses on three expression datasets computed on admixed AMR, leveraging data from individuals with high African GIA, high Native-

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

American ancestry, and from a pooled cohort (Methods). Results are shown in the supplementary Table 7.

Five genes (*LY75*, *BAZ2B*, *CD302*, *WDSUB1*, and *PLA2R1*) were the candidates for eQTL colocalization in the associated region in chromosome 2. However, *LY75* emerged as the most likely causal gene for this locus since the colocalization in whole blood was supported with a PP for H4 (PPH4) of 0.941 and with robust results (supplementary Figure 4). Moreover, this also allowed to prioritize rs12692550 as the most probable causal variant for both traits at this locus with a PP_SNP_H4 of 0.74. Colocalization with gene expression data from admixed AMR validated this finding. *LY75* also had evidence of colocalization in lungs (PPH4=0.887) and the esophagus mucosa (PPH4=0.758). However, we could not prioritize a single causal variant in these two other tissues and sensitivity analyses revealed a weak support.

CD302 and *BAZ2B* were the second and third most likely genes that could drive the association, respectively, according to the colocalization evidence. *CD302* was the most probable according to the high AFR genetic ancestries dataset (supplementary Figure 5).

Despite the chromosome 11 region failing to colocalize with gene expression associations for any of the tissues, the lead variant rs77599934 is in moderate-to-strong LD (r^2=0.776) with rs60606421, which is an eQTL associated to a reduced expression of *DDIAS* in the lungs (supplementary Figure 6). The highest PPH4 for *DDIAS* was in the high AFR genetic ancestry expression dataset (0.71).

Transcriptome-wide association study (TWAS)

Five novel genes, namely *SLC25A37*, *SMARCC1*, *CAMP*, *TYW3*, and *S100A12* (supplementary Table 8) were found significantly associated in the cross-tissue TWAS.

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

To our knowledge, these genes have not been reported previously in any COVID-19 TWAS or GWAS analyses published to date. In the single tissue analyses, *ATP5O* and *CXCR6* were significantly associated in lungs, *CCR9* was significantly associated in whole blood, and *IFNAR2* and *SLC25A37* were associated in lymphocytes.

Likewise, we carried out the TWAS analyses using the models trained in the admixed populations. However, no significant gene-pairs were detected in this case. The 50 genes with the lowest p-values are shown in the supplementary Table 9.

Genetic architecture of COVID-19 hospitalization in AMR populations

Allele frequencies of rs13003835 and rs77599934 across ancestries

Neither rs13003835 (*BAZ2B*) or rs77599934 (*DDIAS*) were significantly associated in the COVID-19 HGI B2 cross-population or population-specific meta-analyses. Thus, we investigated their allele frequencies (AF) across populations and compared their effect sizes.

According to gnomAD v3.1.2, the T allele at rs13003835 (*BAZ2B*) has an AF of 43% in admixed AMR groups while AF is lower in the EUR populations (16%) and in the global sample (29%). Local ancestry inference (LAI) reported by gnomAD shows that within the Native-American component, the risk allele T is the major allele, whereas it is the minor allele within the African and European LAI components. These large differences in AF might be the reason underlying the association found in AMR populations. However, when comparing effect sizes between populations, we found that they were in opposite direction between SAS-AMR and EUR-AFR-EAS and that there was a large heterogeneity among them (Figure 3).

rs77599934 (*DDIAS*) had an AF of 1.1% for the G allele in the non-hospitalized controls (Table 2), in line with the recorded gnomAD AF of 1% in admixed AMR

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

groups. This variant has potential to be population-specific variant, given the allele frequencies in other population groups such as EUR (0% in Finnish, 0.025% in non-Finnish), EAS (0%) and SAS (0.042%) and its greater effect size over AFR populations (Figure 3). Examining the LAI, the G allele occurs at 1.1% frequency in the African component while it is almost absent in the Native-American and European. Due to its low MAF, rs77599934 was not analyzed in the COVID-19 HGI B2 cross-population meta-analysis and was only present in the HGI B2 AFR population-specific meta-analysis, precluding the comparison (Figure 3). For this reason, we retrieved the variant with the lowest p-value within a 50 kb region around rs77599934 in the COVID-19 HGI cross-population analysis to investigate if it was in moderate-to-strong LD with our sentinel variant. The variant with the smallest p-value was rs75684040 (OR=1.07, 95%CI=1.03-1.12, $p=1.84x10^{-3}$). Yet, LD calculations using the 1KGP phase 3 dataset indicated that rs77599934 and rs75684040 were poorly correlated (r^2=0.11).

Cross-population meta-analyses

We carried out two cross-ancestry inverse variance-weighted fixed-effects meta-analyses with the admixed AMR GWAS meta-analysis results to evaluate whether the discovered risk loci replicated when considering other population groups. In doing so, we also identified novel cross-population COVID-19 hospitalization risk loci.

First, we combined the SCOURGE Latin American GWAS results with the HGI B2 ALL analysis (supplementary Table 10). We refer to this analysis as the SC-HGI_{ALL} meta-analysis. Out of the 40 genome-wide significant loci associated with COVID-19 hospitalization in the last HGI release[1], this study replicated 39 and the association was stronger than in the original study in 29 of those (supplementary Table 11). However, the variant rs13003835 located in *BAZ2B* did not replicate (OR=1.00, 95%CI=0.98-1.03, p=0.644).

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

In this cross-ancestry meta-analysis, we replicated two associations that were not found in HGIv7 albeit they were sentinel variants in the latest GenOMICC meta-analysis[2]. We found an association at the *CASC20* locus led by the variant rs2876034 (OR=0.95, 95%CI=0.93-0.97, p=2.83x10^{-8}). This variant is in strong LD with the sentinel variant of that study (rs2326788, r^2=0.92), which was associated with critical COVID-19[2]. Besides, this meta-analysis identified the variant rs66833742 near *ZBTB7A* associated with COVID-19 hospitalization (OR=0.94, 95%CI=0.92-0.96, p=2.50x10^{-8}). Notably, rs66833742 or its perfect proxy rs67602344 (r^2=1) are also associated with upregulation of *ZBTB7A* in whole blood and in esophagus mucosa. This variant was previously associated with COVID-19 hospitalization[2].

In a second analysis, we also explored the associations across the defined admixed AMR, EUR, and AFR ancestral sources by combining through meta-analysis the SCOURGE Latin American GWAS results with the HGI studies in EUR, AFR, and admixed AMR, and excluding those from EAS and SAS (Supplementary Table 12). We refer to this as the SC-HGI_{3POP} meta-analysis. The association at rs13003835 (*BAZ2B*, OR=1.01, 95%CI=0.98-1.03, p=0.605) was not replicated and rs77599934 near *DDIAS* could not be assessed, although the association at the *ZBTB7A* locus was confirmed (rs66833742, OR=0.94, 95%CI=0.92-0.96, p=1.89x10^{-8}). The variant rs76564172 located near *CREBBP* also reached statistical significance (OR=1.31, 95% CI=1.25-1.38, p=9.64x10^{-9}). The sentinel variant of the region linked to *CREBBP* (in the trans-ancestry meta-analysis) was also subjected a Bayesian fine mapping (supplementary Table 6) and colocalization with eQTLs under the GTEx v8 MASHR models in lungs, esophagus mucosa, whole blood, and transformed lymphocytes. Eight variants were included in the credible set for the region in chromosome 16 (meta-analysis SC-HGI_{3POP}), although *CREBBP* did not colocalize in any of the tissues.

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Polygenic risk score models

Using the 49 variants associated with disease severity that are shared across populations according to the HGIv7, we constructed a polygenic risk score (PGS) model to assess its generalizability in the admixed AMR (Supplementary Table 13). First, we calculated the PGS for the SCOURGE Latin Americans and explored the association with COVID-19 hospitalization under a logistic regression model. The PGS model was associated with a 1.48-fold increase in COVID-19 hospitalization risk per every PGS standard deviation. It also contributed to explain a slightly larger variance (R2=1.07%) than the baseline model.

Subsequently, we divided the individuals into PGS deciles and percentiles to assess their risk stratification. The median percentile among controls was 40, while in cases it was 63. Those in the top PGS decile exhibited a 5.90-fold (95% CI=3.29-10.60, $p=2.79x10^{-9}$) greater risk compared to individuals in the lowest decile, whereas the effects for the rest of the comparisons were much milder.

We also examined the distribution of PGS scores across a 5-level severity scale to further determine if there was any correspondence between clinical severity and genetic risk. Median PGS scores were lower in the asymptomatic and mild groups, whereas higher median scores were observed in the moderate, severe, and critical patients (Figure 4). We fitted a multinomial model using the asymptomatic class as reference and calculated the OR for each category (Supplementary Table 13), observing that the disease genetic risk was similar among asymptomatic, mild, and moderate patients. Given that the PGS was built with variants associated with critical disease and/or hospitalization and that the categories severe and critical correspond to hospitalized patients, these results underscore the ability of cross-ancestry PGS for risk stratification even in an admixed population.

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Finally, we incorporated the novel lead SNPs from our AMR meta-analysis (rs13003835, rs2477820, and rs77599934) into the PGS model. Their inclusion in the model contributed to explain a larger variance (R2=1.74%) than the model without them. This result, however, should be taken with caution given the risk of overfitting due to the use of the same subjects both for the derivation and testing of the variants.

DISCUSSION

We have conducted the largest GWAS meta-analysis of COVID-19 hospitalization in admixed AMR to date. While the genetic risk basis discovered for COVID-19 is largely shared among populations, trans-ancestry meta-analyses on this disease have primarily included EUR samples. This dominance of GWAS in Europeans, and the subsequent bias in sample sizes, can mask population-specific genetic risks (i.e., variants that are monomorphic in some populations) or be less powered to detect risk variants having higher allele frequencies in population groups other than Europeans. In this sense, after combining data from admixed AMR patients, we found two risk loci which are first discovered in a GWAS of Latin-American populations. Interestingly, the sentinel variant rs77599934 in the *DDIAS* gene is a rare coding variant (~1% for allele G) with a large effect on COVID-19 hospitalization that is nearly monomorphic in most of the other populations. This has likely led to its exclusion from the cross-populations meta-analyses conducted to date, remaining undetectable.

Fine mapping of the region harbouring *DDIAS* did not reveal further information about which gene could be the more prone to be causal, or about the functional consequences of the risk variant. However, *DDIAS*, known as damage-induced apoptosis suppressor gene, is itself a plausible candidate gene. It has been linked to DNA damage repair

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

mechanisms: research showed that depletion of *DDIAS* led to an increase of ATM phosphorylation and the formation of p53-binding protein (53BP1) foci, a known biomarker of DNA double-strand breaks, suggesting a potential role in double-strand break repair[19]. Similarly, elevated levels of phosphorylated nuclear histone 2AXγ were detected after knocking down *DDIAS*, further emphasizing its role in DNA damage[20]. Interestingly, a study found that the infection by SARS-CoV-2 also triggered the phosphorylation of the ATM kinase and inhibited repair mechanisms, causing the accumulation of DNA damage[21]. This same study reported the activation of the pro-inflammatory pathway p38/MAPK by the virus, which was as well prompted after knocking-down *DDIAS*[20].

Regarding lung function, the role of *DDIAS* in lung cancer has been widely studied. It has been proposed as a potential biomarker for lung cancer after finding that it interacts with STAT3 in lung cancer cells, regulating IL-6[22,23] and thus mediating inflammatory processes. Furthermore, another study determined that its blockade inhibited lung cancer cell growth[20]. The sentinel variant was in strong LD with an eQTL that reduced gene expression of *DDIAS* in lung, and our findings suggest that *DDIAS* gene may be indeed involved in viral response. Hence, one reasonable hypothesis is that reduced expression of *DDIAS* could potentially facilitate SARS-CoV-2 infection through the downregulation of pathways involved in DNA repairment and inflammation. Another prioritized gene from this region was *PRCP*, an angiotensinase that has been linked to hypertension and for which a hypothesis on its role on COVID-19 progression has been raised[24,25].

The risk region found in chromosome 2 prioritized more than one gene. The lead variant rs13003835 is located within *BAZ2B*. *BAZ2B* encodes one of the regulatory subunits of the Imitation switch (ISWI) chromatin remodelers[26] constituting the BRF-1/BRF-5

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

complexes with SMARCA1 and SMARCA5, respectively, and the association signal colocalized with eQTLs in whole blood. The gene *LY75* (encoding the lymphocyte antigen 75) also colocalized with eQTLs in whole blood, esophagus mucosa, and lung tissues. Lymphocyte antigen 75 is involved in immune processes through antigen presentation in dendritic cells and endocytosis[27], and has been associated with inflammatory diseases, representing also a compelling candidate for the region. Increased expression of *LY75* has been detected within hours after the infection by SARS-CoV-2[28,29]. Lastly, the signal of *CD302* colocalized in individuals with high AFR ancestral admixture in whole blood. This gene is located in the vicinity of *LY75* and both conform the readthrough LY75-CD302. It is worth noting that differences in AF for this variant suggest that analyses in AMR populations might be more powered to detect the association, supporting the necessity of population-specific studies.

A third novel risk region was observed in chromosome 15, between the genes *IQGAP1* and *ZNF774*, although not reaching genome-wide significance.

Secondary analyses revealed five TWAS-associated genes, some of which have been already linked to severe COVID-19. In a comprehensive multi-tissue gene expression profiling study[30], decreased expression of *CAMP* and *S100A8*/*S100A9* genes in COVID-19 severe patients was observed, while another study detected the upregulation of *SCL25A37* among severe COVID-19 patients[31]. *SMARCC1* is a subunit of the SWI/SNF chromatin remodelling complex that has been identified as pro-viral for SARS-CoV-2 and other coronavirus strains through a genome-wide screen[32]. This complex is crucial for *ACE2* expression and the viral entry in the cell[33].

To explore the genetic architecture of the trait among admixed AMR populations, we performed two cross-ancestry meta-analyses including the SCOURGE Latin-American cohort GWAS findings. We found that the two novel risk variants did not associate with

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

COVID-19 hospitalization outside the population-specific meta-analysis, highlighting the importance of complementing trans-ancestry meta-analyses with group-specific analyses. Notably, this analysis did not replicate the association at the *DSTYK* locus, which was associated with severe COVID-19 in Brazilian individuals with higher European admixture[34]. This lack of replication supports the initial hypothesis of that study suggesting that the risk haplotype derived from European populations, as we have reduced the weight of this ancestral contribution in our study by excluding those individuals.

Moreover, these cross-ancestry meta-analyses pointed to three loci that were not genome-wide significant in the HGIv7 ALL meta-analysis: a novel locus at *CREBBP*, and two loci at *ZBTB7A* and *CASC20* that were reported in another meta-analysis. *CREBBP* and *ZBTB7A* achieved a stronger significance when considering only EUR, AFR, and admixed AMR GIA groups. According to a recent study, elevated levels of the *ZBTB7A* gene promote a quasi-homeostatic state between coronaviruses and host cells, preventing cell death by regulating oxidative stress pathways[35]. This gene is involved in several signalling pathways, such as B and T cell differentiation[36]. On a separate note, *CREBBP* encodes the CREB binding protein (CBP), involved in transcription activation, that is known to positively regulate the type I interferon response through virus-induced phosphorylation of IRF-3[37]. Besides, the CREBP/CBP interaction has been implicated in SARS-CoV-2 infection[38] via the cAMP/PKA pathway. In fact, cells with suppressed *CREBBP* gene expression exhibit reduced replication of the so called Delta and Omicron SARS-CoV-2 variants[38].

The cross-population PGS model effectively stratified individuals based on their genetic risk and demonstrated consistency with the clinical severity classification of the patients. The inclusion of the new variants in the PGS model slightly improved the

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

predictive value of the PGS. However, it is important to confirm this last finding in an external admixed AMR cohort to address potential overfitting arising from using the same individuals both for the discovery of the associations and for testing the model.

This study is subject to limitations, mostly concerning the sample recruitment and composition. The SCOURGE Latino-American sample size is small and the GWAS is underpowered. Another limitation is the difference in case-control recruitment across sampling regions that, yet controlled for, may reduce the ability to observe significant associations driven by different compositions of the populations. In this sense, the identified risk loci might not replicate in a cohort lacking any of the parental population sources from the three-way admixture. Likewise, we could not explicitly control for socio-environmental factors that could have affected COVID-19 spread and hospitalization rates, although genetic principal components are known to capture non-genetic factors. Finally, we must acknowledge the lack of a replication cohort. We have used all the available GWAS data for COVID-19 hospitalization in admixed AMR in this meta-analysis due to the low number of studies conducted. Therefore, we had no studies to replicate or validate the results. These concerns may be addressed in the future by including more AMR GWAS studies in the meta-analysis, both by involving diverse populations in study designs and by supporting research from countries in Latin-America.

This study provides novel insights into the genetic basis of COVID-19 severity, emphasizing the importance of considering host genetic factors through using non-European populations, especially of admixed sources. Such complementary efforts can pin down new variants and increase our knowledge on the host genetic factors of severe COVID-19.

Materials and methods

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

GWAS in Latin Americans from SCOURGE

The SCOURGE Latin American cohort

A total of 3,729 of COVID-19 positive cases were recruited across five countries from Latin America (Mexico, Brazil, Colombia, Paraguay, and Ecuador) by 13 participating centres (supplementary Table 1) from March 2020 to July 2021. In addition, we included 1,082 COVID-19 positive individuals recruited between March and December 2020 in Spain who either had evidence of origin from a Latin American country or showed inferred genetic admixture between AMR, EUR, and AFR (with < 0.05% SAS/EAS). These individuals were excluded from a previous SCOURGE study that focused on participants with European genetic ancestries[16]. We used hospitalization as a proxy for disease severity and defined as cases those COVID-19 positive patients that underwent hospitalization as a consequence of the infection and used as controls those that did not need hospitalization due to COVID-19.

Samples and data were collected with informed consent after the approval of the Ethics and Scientific Committees from the participating centres and by the Galician Ethics Committee Ref 2020/197. Recruitment of patients from IMSS (in Mexico, City), was approved by of the National Comitte of Clinical Research, from Instituto Mexicano del Seguro Social, Mexico (protocol R-2020-785-082).

Samples and data were processed following normalized procedures. The REDCap electronic data capture tool[39,40], hosted at Centro de Investigación Biomédica en Red (CIBER) from the Instituto de Salud Carlos III (ISCIII), was used to collect and manage demographic, epidemiological, and clinical variables. Subjects were diagnosed for COVID-19 based on quantitative PCR tests (79.3%), or according to clinical (2.2%) or laboratory procedures (antibody tests: 16.3%; other microbiological tests: 2.2%).

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

SNP array genotyping

Genomic DNA was obtained from peripheral blood and isolated using the Chemagic DNA Blood 100 kit (PerkinElmer Chemagen Technologies GmbH), following the manufacturer's recommendations.

Samples were genotyped with the Axiom Spain Biobank Array (Thermo Fisher Scientific) following the manufacturer's instructions in the Santiago de Compostela Node of the National Genotyping Center (CeGen-ISCIII; http://www.usc.es/cegen). This array contains probes for genotyping a total of 757,836 SNPs. Clustering and genotype calling were performed using the Axiom Analysis Suite v4.0.3.3 software.

Quality control steps and variant imputation

A quality control (QC) procedure using PLINK 1.9[41] was applied to both samples and the genotyped SNPs. We excluded variants with a minor allele frequency (MAF) <1%, a call rate <98%, and markers strongly deviating from Hardy-Weinberg equilibrium expectations ($p<1x10^{-6}$) with mid-p adjustment. We also explored the excess of heterozygosity to discard potential cross-sample contaminations. Samples missing >2% of the variants were filtered out. Subsequently, we kept the autosomal SNPs and removed high LD regions and conducted LD-pruning (windows of 1,000 SNPs, with step size of 80 and r^2 threshold of 0.1) to assess kinship and estimate the global ancestral proportions. Kinship was evaluated based on IBD values, removing one individual from each pair with PI_HAT>0.25 that showed a Z0, Z1, and Z2 coherent pattern (according to the theoretical expected values for each relatedness level). Genetic principal components (PCs) were calculated with PLINK with the subset of LD pruned variants.

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Genotypes were imputed with the TOPMed version r2 reference panel (GRCh38) using the TOPMed Imputation Server and variants with Rsq<0.3 or with MAF<1% were filtered out. A total of 4,348 individuals and 10,671,028 genetic variants were included in the analyses.

Genetic admixture estimation

Global genetic inferred ancestry (GIA), referred to the genetic similarity to the used reference individuals, was estimated with the ADMIXTURE[42] v1.3 software following a two-step procedure. First, we randomly sampled 79 European (EUR) and 79 African (AFR) samples from The 1000 Genomes Project (1KGP)[43] and merged them with the 79 Native American (AMR) samples from Mao et al.[44] keeping the biallelic SNPs. LD-pruned variants were selected from this merge using the same parameters as in the QC. We then run an unsupervised analysis with K=3 to redefine and homogenize the clusters and to compose a refined reference for the analyses, by applying a threshold of ≥95% of belonging to a particular cluster. As a result of this, 20 AFR, 18 EUR, and 38 AMR individuals were removed. The same LD-pruned variants data from the remaining individuals were merged with the SCOURGE Latin American cohort to perform a supervised clustering and estimated admixture proportions. A total of 471 samples from the SCOURGE cohort with >80% estimated European GIA were removed to reduce the weight of the European ancestral component, leaving a total of 3,512 admixed American (AMR) subjects for downstream analyses.

Association analysis

Results for the SCOURGE Latin Americans GWAS were obtained testing for COVID-19 hospitalization as a surrogate of severity. To accommodate the continuum of GIA in the cohort, we opted for a joint testing of all the individuals as a single study using a

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

mixed regression model, as this approach has demonstrated a greater power and to sufficiently control population structure[45]. The SCOURGE cohort consisted of 3,512 COVID-19 positive patients: cases (n=1,625) were defined as hospitalized COVID-19 patients and controls (n=1,887) as non-hospitalized COVID-19 positive patients.

Logistic mixed regression models were fitted using the SAIGEgds[46] package in R, which implements the two-step mixed SAIGE[47] model methodology and the SPA test. Baseline covariables included sex, age, and the first 10 PCs. To account for a potential heterogeneity in the recruitment and hospitalization criteria across the participating countries, we adjusted the models by groups of the recruitment areas classified in six categories: Brazil, Colombia, Ecuador, Mexico, Paraguay, and Spain. This dataset has not been used in any previously GWAS of COVID-19 published to date.

Meta-analysis of Latin-American populations

The results of the SCOURGE Latin American cohort were meta-analyzed with the AMR HGI-B2 data, conforming our primary analysis. Summary results from the HGI freeze 7 B2 analysis corresponding to the admixed AMR population were obtained from the public repository (April 8, 2022: https://www.covid19hg.org/results/r7/), summing up 3,077 cases and 66,686 controls from seven contributing studies. We selected the B2 phenotype definition because it offered more power and the presence of population controls not ascertained for COVID-19 does not have a drastic impact in the association results.

The meta-analysis was performed using an inverse-variance weighting method in METAL[48]. Average allele frequency was calculated and variants with low imputation quality (Rsq<0.3) were filtered out, leaving 10,121,172 variants for meta-analysis.

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Heterogeneity between studies was evaluated with the Cochran's-Q test. The inflation of results was assessed based on a genomic control (lambda).

Definition of the genetic risk loci and putative functional impact

Definition of lead variant and novel loci

To define the lead variants in the loci that were genome-wide significant, an LD-clumping was performed on the meta-analysis data using a threshold p-value<$5x10^{-8}$, clump distance=1500 kb, independence set at a threshold r^2=0.1 and used the SCOURGE cohort genotype data as LD reference panel. Independent loci were deemed as a novel finding if they met the following criteria: 1) p-value<$5x10^{-8}$ in the meta-analysis and p-value>$5x10^{-8}$ in the HGI B2 ALL meta-analysis or in the HGI B2 AMR and AFR and EUR analyses when considered by separate; 2) Cochran's Q-test for heterogeneity of effects is <$0.05/N_{loci}$, where N_{loci} is the number of independent variants with p<$5x10^{-8}$; and 3) the nearest gene has not been previously described in the latest HGIv7 update.

Annotation and initial mapping

Functional annotation was done with FUMA[49] for those variants with a p-value<$5x10^{-8}$ or in moderate-to-strong LD (r^2>0.6) with the lead variants, where the LD was calculated from the 1KGP AMR panel. Genetic risk loci were defined by collapsing LD-blocks within 250 kb. Then, genes, scaled CADD v1.4 scores, and RegulomeDB v1.1 scores were annotated for the resulting variants with ANNOVAR in FUMA[49]. Gene-based analysis was also performed using MAGMA[50] as implemented in FUMA, under the SNP-wide mean model using the 1KGP AMR reference panel. Significance was set at a threshold p<$2.66x10^{-6}$ (which assumes that variants can be mapped to a total of 18,817 genes).

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

FUMA allowed us to perform an initial gene mapping by two approaches: (1) positional mapping, which assigns variants to genes by physical distance using 10-kb windows; and (2) eQTL mapping based on GTEx v.8 data from whole blood, lungs, lymphocytes, and oesophagus mucosa tissues, establishing a False Discovery Rate (FDR) of 0.05 to declare significance for variant-gene pairs.

Subsequently, to assign the variants to the most likely gene driving the association, we refined the candidate genes by fine mapping the discovered regions and implementing functional mapping.

To conduct a Bayesian fine mapping, credible sets for the genetic loci considered novel findings were calculated on the results from each of the three meta-analyses to identify a subset of variants most likely containing the causal variant at 95% confidence level, assuming that there is a single causal variant and that it has been tested. We used *corrcoverage* (https://cran.rstudio.com/web/packages/corrcoverage/index.html) for R to calculate the posterior probabilities of the variant being causal for all variants with an r^2>0.1 with the leading SNP and within 1 Mb except for the novel variant in chromosome 19, for which we used a window of 0.5 Mb. Variants were added to the credible set until the sum of the posterior probabilities was ≥0.95. VEP (https://www.ensembl.org/info/docs/tools/vep/index.html) and the V2G aggregate scoring from Open Targets Genetics (https://genetics.opentargets.org) were used to annotate the biological function of the variants contained in the fine-mapped credible sets

Colocalization analysis

We also conducted colocalization analyses to identify the putative causal genes that could act through the regulation of gene expression. FUMA's eQTL mapping enabled

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

the identification of genes whose expression was associated with the variants in whole blood, lungs, lymphocytes, and oesophagus mucosa tissues. We combined this information with the VEP and V2G aggregate scoring to prioritize genes. For the fine-mapping regions, we included the variants within the calculated credible sets. In the cases where the fine mapping was unsuccessful, we considered variants within a 0.2 Mb window of the lead variant.

For each prioritized gene, we then run COLOC[51] to assess the evidence of colocalization between association signals and the eQTLs in each tissue, when at least one variant overlapped between them. COLOC estimates the posterior probability of two traits sharing the same causal variant in a locus. Prior probabilities of a variant being associated to COVID-19 phenotype (p1) and gene expression (p2) were set at $1x10^{-4}$, while pp2 was set at $1x10^{-6}$ as they are robust thresholds[52]. The posterior probability of colocalization (PP4) > 0.75 and a ratio PP4/PP3>3 were used as the criteria to support evidence of colocalization. Additionally, a threshold of PP4.SNP >0.5 was chosen for causal variant prioritization. In cases were colocalization of a single variant failed, we computed the 95% credible sets. The eQTL data was retrieved from GTEx v8 and only significant variant-gene pairs were considered in the analyses.

Colocalization in whole-blood was also performed using the recent published gene expression datasets derived from a cohort of African Americans, Puerto Ricans, and Mexican Americans (GALA II-SAGE)[53]. We used the results from the pooled cohort for the three discovered loci, and from the AFRHp5 (African genetic ancestry>50%) and IAMHp5 (Native American genetic ancestry>50%) cohorts for the risk loci in chromosomes 2 and 11. Results are shown in the Supplementary Table 10.

Sensitivity plots are shown in supplementary Figures 4 and 5.

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Transcription-wide association studies

Transcriptome-wide association studies (TWAS) were conducted using the pretrained prediction models with MASHR-computed effect sizes on GTEx v8 datasets[54,55]. Results from the Latin-American meta-analysis were harmonized and integrated with the prediction models through S-PrediXcan[56] for lungs, whole blood, lymphocytes and oesophagus mucosa tissues. Statistical significance was set at p-value<0.05 divided by the number of genes that were tested for each tissue. Subsequently, we leveraged results for all 49 tissues and run a multi-tissue TWAS to improve power for association, as demonstrated recently[57]. TWAS was also conducted with the MASHR models for whole-blood in the pooled admixed AMR from the GALA and SAGE studies[53].

Cross-population meta-analyses

We conducted two additional meta-analyses to investigate the ability of combining populations to replicate our discovered risk loci. This methodology enabled the comparison of effects and the significance of associations in the novel risk loci between the results from analyses that included or excluded other population groups.

The first meta-analysis comprised the five populations analysed within HGI (B2-ALL). Additionally, to evaluate the three GIA components within the SCOURGE Latin-American cohort[58], we conducted a meta-analysis of the admixed AMR, EUR, and AFR cohorts (B2). All summary statistics were retrieved from the HGI repository. We applied the same meta-analysis methodology and filters as in the admixed AMR meta-analysis. Novel variants from these meta-analyses were fine-mapped and colocalized with gene expression.

Cross-population Polygenic Risk Score

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

A polygenic risk score (PGS) for critical COVID-19 was derived combining the variants associated with hospitalization or disease severity that have been discovered to date. We curated a list of lead variants that were: 1) associated to either severe disease or hospitalization in the latest HGIv7 release[1] (using the hospitalization weights); or 2) associated to severe disease in the latest GenOMICC meta-analysis[2] that were not reported in the latest HGI release. A total of 49 markers were used in the PGS model (see supplementary Table 13) since two variants were absent from our study.

Scores were calculated and normalized for the SCOURGE Latin-American cohort with PLINK 1.9. This cross-ancestry PGS was used as a predictor for hospitalization (COVID-19 positive that were hospitalized *vs.* COVID-19 positive that did not necessitate hospital admission) by fitting a logistic regression model. Prediction accuracy for the PGS was assessed by performing 500 bootstrap resamples of the increase in the pseudo-R-squared. We also divided the sample in deciles and percentiles to assess risk stratification. The models were fit for the dependent variable adjusting for sex, age, the first 10 PCs, and the sampling region (in the Admixed AMR cohort) with and without the PGS, and the partial pseudo-R2 was computed and averaged among the resamples.

A clinical severity scale was used in a multinomial regression model to further evaluate the power of this cross-ancestry PGS for risk stratification. This severity strata were defined as follows: 0) asymptomatic; 1) mild, that is, with symptoms, but without pulmonary infiltrates or need of oxygen therapy; 2) moderate, that is, with pulmonary infiltrates affecting <50% of the lungs or need of supplemental oxygen therapy; 3) severe disease, that is with hospital admission and PaO_2<65 mmHg or SaO_2<90%, PaO_2/FiO_2<300, SaO_2/FiO_2<440, dyspnea, respiratory frequency≥22 bpm, and infiltrates affecting >50% of the lungs; and 4) critical disease, that is with an admission

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

to the ICU or need of mechanical ventilation (invasive or non-invasive). We also included the novel risk variants as predictors alongside the PRS to determine if they provided increased prediction ability.

Data availability

Summary statistics from the SCOURGE Latin-American GWAS will be available at https://github.com/CIBERER/Scourge-COVID19.

Funding

Instituto de Salud Carlos III (COV20_00622 to A.C., COV20/00792 to M.B., COV20_00181 to C.A., COV20_1144 to M.A.J.S. and A.F.R., PI20/00876 to C.F.); European Union (ERDF) 'A way of making Europe'. Fundación Amancio Ortega, Banco de Santander (to A.C.), Estrella de Levante S.A. and Colabora Mujer Association (to E.G.-N.) and Obra Social La Caixa (to R.B.); Agencia Estatal de Investigación (RTC-2017-6471-1 to C.F.), Cabildo Insular de Tenerife (CGIEU0000219140 'Apuestas científicas del ITER para colaborar en la lucha contra la COVID-19' to C.F.) and Fundación Canaria Instituto de Investigación Sanitaria de Canarias (PIFIISC20/57 to C.F.).

SD-DA was supported by a Xunta de Galicia predoctoral fellowship.

Author contributions

Study design: RC, AC, CF. Data collection: SCOURGE cohort group. Data analysis: SD-DA, RC, ADL, CF, JML-S. Interpretation: SD-DA, RC, ADL. Drafting of the manuscript: SD-DA, RC, ADL, CF, AR-M, AC. Critical revision of the manuscript:

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

SD-DA, RC, ADL, AC, CF, JAR, AR-M, PL. Approval of the final version of the publication: all co-authors.

Acknowledgements

The contribution of the Centro National de Genotipado (CEGEN), and Centro de Supercomputación de Galicia (CESGA) for funding this project by providing supercomputing infrastructures, is also acknowledged. Authors are also particularly grateful for the supply of material and the collaboration of patients, health professionals from participating centers and biobanks. Namely Biobanc-Mur, and biobancs of the Complexo Hospitalario Universitario de A Coruña, Complexo Hospitalario Universitario de Santiago, Hospital Clínico San Carlos, Hospital La Fe, Hospital Universitario Puerta de Hierro Majadahonda—Instituto de Investigación Sanitaria Puerta de Hierro—Segovia de Arana, Hospital Ramón y Cajal, IDIBGI, IdISBa, IIS Biocruces Bizkaia, IIS Galicia Sur. Also biobanks of the Sistema de Salud de Aragón, Sistema Sanitario Público de Andalucía, and Banco Nacional de ADN.

References

1. Initiative, T. C.-19 H. G. & Ganna, A. A second update on mapping the human genetic architecture of COVID-19. 2022.12.24.22283874 Preprint at https://doi.org/10.1101/2022.12.24.22283874 (2023).

2. GWAS and meta-analysis identifies 49 genetic variants underlying critical COVID-19 | Nature. https://www.nature.com/articles/s41586-023-06034-3.

3. Niemi, M. E. K. *et al.* Mapping the human genetic architecture of COVID-19. *Nature* **600**, 472–477 (2021).

4. Popejoy, A. B. & Fullerton, S. M. Genomics is failing on diversity. *Nature* **538**, 161–164 (2016).

5. Sirugo, G., Williams, S. M. & Tishkoff, S. A. The Missing Diversity in Human Genetic Studies. *Cell* **177**, 26–31 (2019).

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

6. Li, Y. R. & Keating, B. J. Trans-ethnic genome-wide association studies: advantages and challenges of mapping in diverse populations. *Genome Med.* **6**, 91 (2014).

7. Rosenberg, N. A. *et al.* Genome-wide association studies in diverse populations. *Nat. Rev. Genet.* **11**, 356–366 (2010).

8. Kwok, A. J., Mentzer, A. & Knight, J. C. Host genetics and infectious disease: new tools, insights and translational opportunities. *Nat. Rev. Genet.* **22**, 137–153 (2021).

9. Karlsson, E. K., Kwiatkowski, D. P. & Sabeti, P. C. Natural selection and infectious disease in human populations. *Nat. Rev. Genet.* **15**, 379–393 (2014).

10. Namkoong, H. *et al.* DOCK2 is involved in the host genetics and biology of severe COVID-19. *Nature* **609**, 754–760 (2022).

11. Bastard, P. *et al.* A loss-of-function IFNAR1 allele in Polynesia underlies severe viral diseases in homozygotes. *J. Exp. Med.* **219**, e20220028 (2022).

12. Duncan, C. J. A. *et al.* Life-threatening viral disease in a novel form of autosomal recessive IFNAR2 deficiency in the Arctic. *J. Exp. Med.* **219**, e20212427 (2022).

13. Peterson, R. E. *et al.* Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. *Cell* **179**, 589–603 (2019).

14. Mester, R. *et al.* Impact of cross-ancestry genetic architecture on GWAS in admixed populations. 2023.01.20.524946 Preprint at https://doi.org/10.1101/2023.01.20.524946 (2023).

15. Tractor uses local ancestry to enable the inclusion of admixed individuals in GWAS and to boost power | Nature Genetics. https://www.nature.com/articles/s41588-020-00766-y.

16. Cruz, R. *et al.* Novel genes and sex differences in COVID-19 severity. *Hum. Mol. Genet.* **31**, 3789–3806 (2022).

17. Degenhardt, F. *et al.* Detailed stratified GWAS analysis for severe COVID-19 in four European populations. *Hum. Mol. Genet.* **31**, 3945–3966 (2022).

18. Whole-genome sequencing reveals host factors underlying critical COVID-19 | Nature. https://www.nature.com/articles/s41586-022-04576-6.

19. Evolution-based screening enables genome-wide prioritization and discovery of DNA repair genes | PNAS. https://www.pnas.org/doi/full/10.1073/pnas.1906559116.

20. Human Noxin is an anti-apoptotic protein in response to DNA damage of A549 non-small cell lung carcinoma - Won - 2014 - International Journal of Cancer - Wiley Online Library. https://onlinelibrary.wiley.com/doi/10.1002/ijc.28600.

21. Gioia, U. *et al.* SARS-CoV-2 infection induces DNA damage, through CHK1 degradation and impaired 53BP1 recruitment, and cellular senescence. *Nat. Cell Biol.* **25**, 550–564 (2023).

22. Im, J.-Y. *et al.* DDIAS promotes STAT3 activation by preventing STAT3 recruitment to PTPRM in lung cancer cells. *Oncogenesis* **9**, 1–11 (2020).

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

23. Im, J.-Y., Kang, M.-J., Kim, B.-K. & Won, M. DDIAS, DNA damage-induced apoptosis suppressor, is a potential therapeutic target in cancer. *Exp. Mol. Med.* 1–7 (2023) doi:10.1038/s12276-023-00974-6.

24. Angeli, F. *et al.* The spike effect of acute respiratory syndrome coronavirus 2 and coronavirus disease 2019 vaccines on blood pressure. *Eur. J. Intern. Med.* **109**, 12–21 (2023).

25. Silva-Aguiar, R. P. *et al.* Role of the renin-angiotensin system in the development of severe COVID-19 in hypertensive patients. *Am. J. Physiol.-Lung Cell. Mol. Physiol.* **319**, L596–L602 (2020).

26. Li, Y. *et al.* The emerging role of ISWI chromatin remodeling complexes in cancer. *J. Exp. Clin. Cancer Res.* **40**, 346 (2021).

27. The Dendritic Cell Receptor for Endocytosis, Dec-205, Can Recycle and Enhance Antigen Presentation via Major Histocompatibility Complex Class II–Positive Lysosomal Compartments | Journal of Cell Biology | Rockefeller University Press. https://rupress.org/jcb/article/151/3/673/21295/The-Dendritic-Cell-Receptor-for-Endocytosis-Dec.

28. Sims, A. C. *et al.* Release of Severe Acute Respiratory Syndrome Coronavirus Nuclear Import Block Enhances Host Transcription in Human Lung Cells. *J. Virol.* **87**, 3885–3902 (2013).

29. A Network Integration Approach to Predict Conserved Regulators Related to Pathogenicity of Influenza and SARS-CoV Respiratory Viruses | PLOS ONE. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0069374.

30. Gómez-Carballa, A. *et al.* A multi-tissue study of immune gene expression profiling highlights the key role of the nasal epithelium in COVID-19 severity. *Environ. Res.* **210**, 112890 (2022).

31. Policard, M., Jain, S., Rego, S. & Dakshanamurthy, S. Immune characterization and profiles of SARS-CoV-2 infected patients reveals potential host therapeutic targets and SARS-CoV-2 oncogenesis mechanism. *Virus Res.* **301**, 198464 (2021).

32. Wei, J. *et al.* Genome-wide CRISPR Screens Reveal Host Factors Critical for SARS-CoV-2 Infection. *Cell* **184**, 76-91.e13 (2021).

33. Wei, J. *et al.* Pharmacological disruption of mSWI/SNF complex activity restricts SARS-CoV-2 infection. *Nat. Genet.* **55**, 471–483 (2023).

34. Pereira, A. C. *et al.* Genetic risk factors and COVID-19 severity in Brazil: results from BRACOVID study. *Hum. Mol. Genet.* **31**, 3021–3031 (2022).

35. Zhu, X. *et al.* ZBTB7A promotes virus-host homeostasis during human coronavirus 229E infection. *Cell Rep.* **41**, 111540 (2022).

36. Gupta, S. *et al.* Emerging role of ZBTB7A as an oncogenic driver and transcriptional repressor. *Cancer Lett.* **483**, 22–34 (2020).

37. Yoneyama, M. *et al.* Direct triggering of the type I interferon system by virus infection: activation of a transcription factor complex containing IRF-3 and CBP/p300. *EMBO J.* **17**, 1087–1095 (1998).

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

38. Yang, Q. *et al.* SARS-CoV-2 infection activates CREB/CBP in cellular cyclic AMP-dependent pathways. *J. Med. Virol.* **95**, e28383 (2023).

39. Harris, P. A. *et al.* Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* **42**, 377–381 (2009).

40. Harris, P. A. *et al.* The REDCap consortium: Building an international community of software platform partners. *J. Biomed. Inform.* **95**, 103208 (2019).

41. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

42. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).

43. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

44. Mao, X. *et al.* A Genomewide Admixture Mapping Panel for Hispanic/Latino Populations. *Am. J. Hum. Genet.* **80**, 1171–1178 (2007).

45. Wojcik, G. L. *et al.* Genetic analyses of diverse populations improves discovery for complex traits. *Nature* **570**, 514–518 (2019).

46. Zheng, X. & Davis, J. W. SAIGEgds—an efficient statistical tool for large-scale PheWAS with mixed models. *Bioinformatics* **37**, 728–730 (2021).

47. Zhou, W. *et al.* Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat. Genet.* **50**, 1335–1341 (2018).

48. METAL: fast and efficient meta-analysis of genomewide association scans | Bioinformatics | Oxford Academic. https://academic.oup.com/bioinformatics/article/26/17/2190/198154.

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

50. MAGMA: Generalized Gene-Set Analysis of GWAS Data | PLOS Computational Biology. https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004219.

51. Giambartolomei, C. *et al.* Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLOS Genet.* **10**, e1004383 (2014).

52. Wallace, C. Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. *PLOS Genet.* **16**, e1008720 (2020).

53. Kachuri, L. *et al.* Gene expression in African Americans, Puerto Ricans and Mexican Americans reveals ancestry-specific patterns of genetic architecture. *Nat. Genet.* **55**, 952–963 (2023).

54. Barbeira, A. N. *et al.* Exploiting the GTEx resources to decipher the mechanisms at GWAS loci. *Genome Biol.* **22**, 49 (2021).

55. Barbeira, A. N. *et al.* GWAS and GTEx QTL integration. (2019) doi:10.5281/zenodo.3518299.

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

56. Barbeira, A. N. *et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat. Commun.* **9**, 1825 (2018).

57. Barbeira, A. N. *et al.* Integrating predicted transcriptome from multiple tissues improves association detection. *PLOS Genet.* **15**, e1007889 (2019).

58. Genome-wide patterns of population structure and admixture among Hispanic/Latino populations | PNAS. https://www.pnas.org/doi/10.1073/pnas.0914618107?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed.

Table 1. Demographic characteristics of the SCOURGE Latin-American cohort.

Variable	Non Hospitalized N = 1,887	Hospitalized N = 1,608
Age – mean years ± SD	39.1 ± 11.9	54.1 ±14.5
Sex - N (%)		
Female (%)	1253 (66.4)	668 (41.5)
GIA* – % mean ±SD		
European	54.4 ±16.2	39.4 ± 20.7
African	15.3 ± 12.7	9.1 ± 11.6
Native American	30.3 ± 19.8	51.46 ± 26.5
Comorbidities - N (%)		
Vascular/endocrinological	488 (25.9)	873 (54.3)
Cardiac	60 (3.2)	150 (9.3)
Nervous	15 (0.8)	61 (3.8)
Digestive	14 (0.7)	33 (2.0)
Onco-hematological	21 (1.1)	48 (3.00)
Respiratory	76 (4.0)	118 (7.3)

*Global genetic inferred ancestry.

Table 2. Lead independent variants in the admixed AMR GWAS meta-analysis.

SNP rsID	***chr:pos***	***EA***	***NEA***	***OR (95% CI)***	***P-value***	***EAF cases***	***EAF controls***	***Nearest gene***

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

rs13003835	2:159407982	T	C	1.20 (1.12-1.27)	3.66E-08	0.563	0.429	*BAZ2B*
rs35731912	3:45848457	T	C	1.65 (1.47-1.85)	6.30E-17	0.087	0.056	*LZTFL1*
rs2477820	6:41535254	A	T	0.84 (0.79-0.89)	1.89E-08	0.453	0.517	*FOXP4-AS1*
rs77599934	11:82906875	G	A	2.27 (1.7-3.04)	2.26E-08	0.016	0.011	*DDIAS*

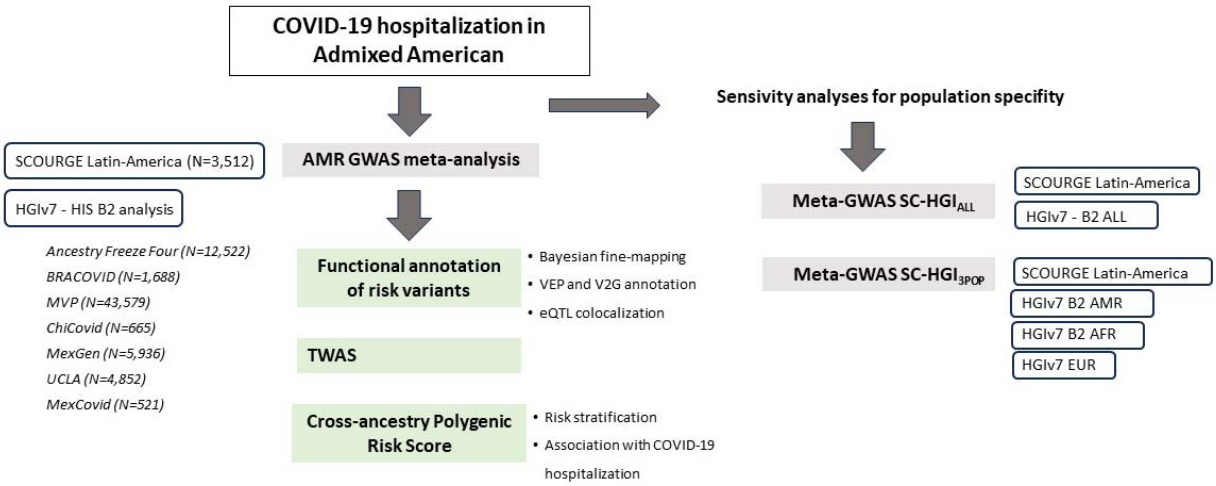
EA: effect allele; NEA: non-effect allele; EAF: effect allele frequency in the SCOURGE study.

Table 3. Novel variants in the SC-HGI$_{ALL}$ and SC-HGI$_{3POP}$ meta-analyses (with respect to HGIv7). Independent signals after LD clumping.

SNP rsID	***chr:pos***	***EA***	***NEA***	***OR (95% CI)***	***P-value***	***Nearest gene***	***Analysis***
rs76564172	16:3892266	T	G	1.31 (1.19-1.44)	9.64E-09	*CREBBP*	SC-HGI$_{3POP}$
rs66833742	19:4063488	T	C	0.94 (0.92-0.96)	1.89E-08	*ZBTB7A*	SC-HGI$_{3POP}$
rs66833742	19:4063488	T	C	0.94 (0.92-0.96)	2.50E-08	*ZBTB7A*	SC-HGI$_{ALL}$
rs2876034	20:6492834	A	T	0.95 (0.93-0.97)	2.83E-08	*CASC20*	SC-HGI$_{ALL}$

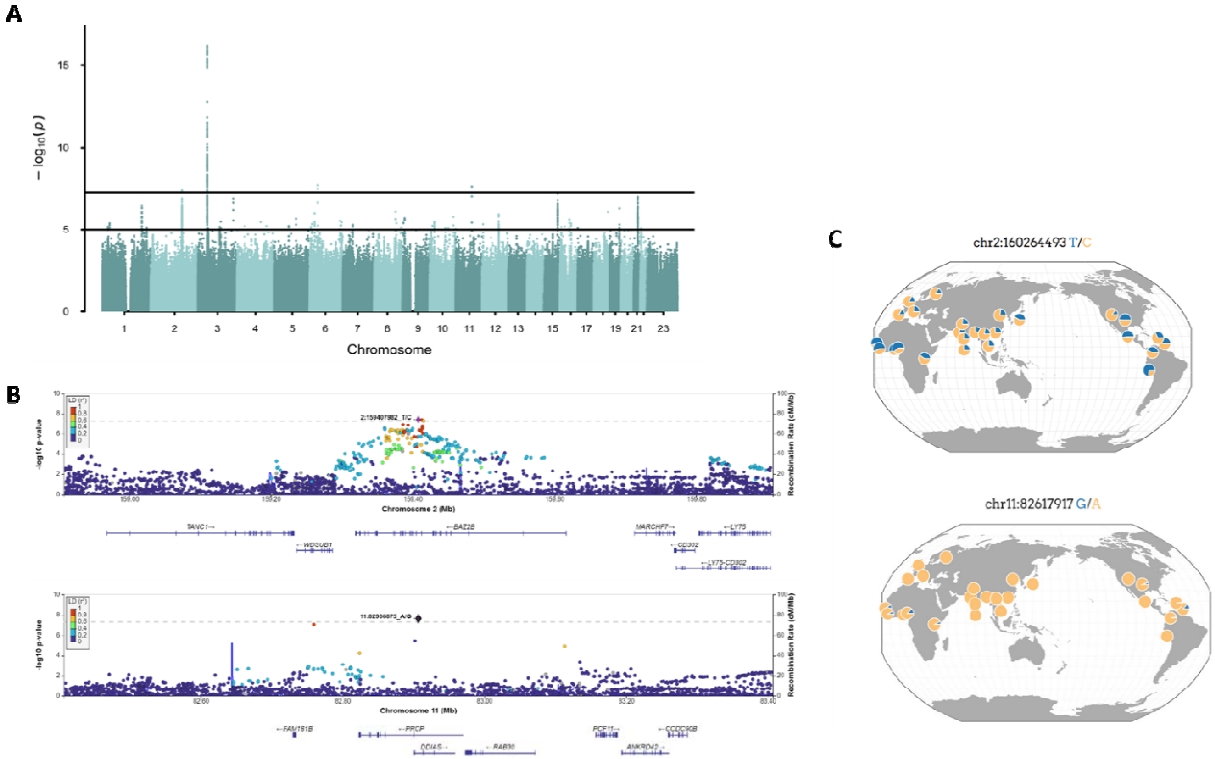
EA: effect allele; NEA: non-effect allele.

Figure 1. Flow chart of this study.



medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 2. A) Manhattan plot for the admixed AMR GWAS meta-analysis. Probability thresholds at $p=5x10^{-8}$ and $p=5x10^{-5}$ are indicated by the horizontal lines. Genome-wide significant associations with COVID-19 hospitalizations were found in chromosome 2 (within *BAZ2B*), chromosome 3 (within *LZTFL1*), chromosome 6 (within *FOXP4*), and chromosome 11 (within *DDIAS*). A Quantile-Quantile plot is shown in supplementary Figure 2. B) Regional association plots for rs1003835 at chromosome 2 and rs77599934 at chromosome 11; C) Allele frequency distribution across The 1000 Genomes Project populations for the lead variants rs1003835 and rs77599934.



medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 3. Forest plot showing effect sizes and the corresponding confidence intervals for the sentinel variants identified in the AMR meta-analysis across populations. All beta values with their corresponding CIs were retrieved from the B2 population-specific meta-analysis from the HGI v7 release, except for AMR, for which the beta value and IC from the HGI_{AMR}-SCOURGE meta-analysis is represented.

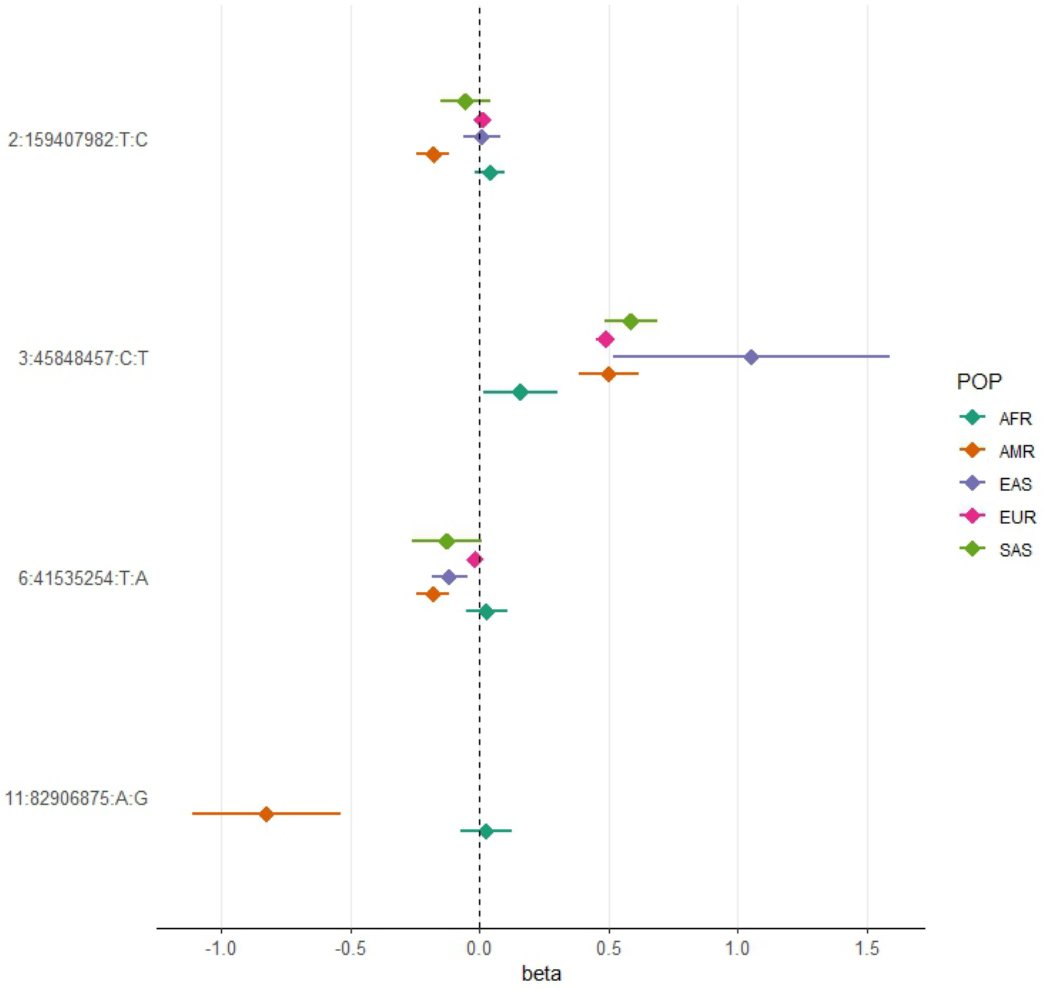
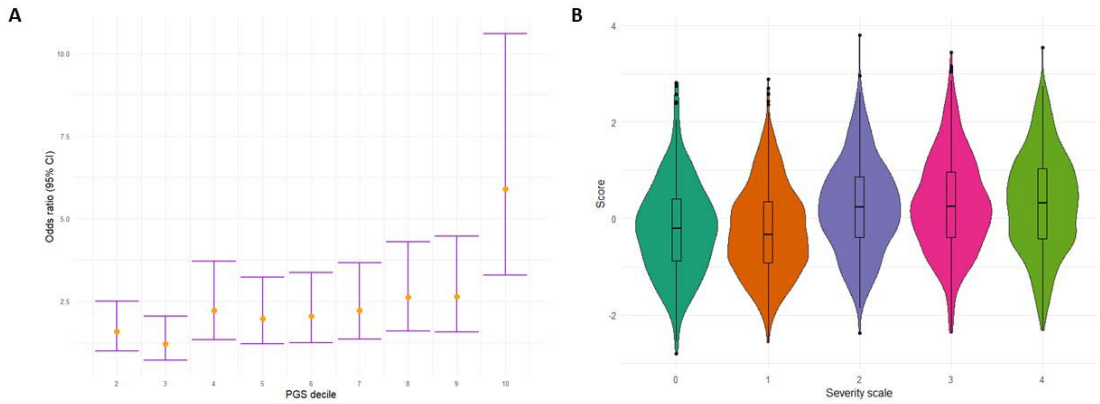


Figure 4. (A) Polygenic risk stratified by PGS deciles comparing each risk group against the lowest risk group (OR-95%CI); (B) Distribution of the PGS scores in each

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

of the severity scale classes (0-Asymptomatic, 1-Mild disease, 2-Moderate disease, 3-Severe disease, 4-Critical disease).



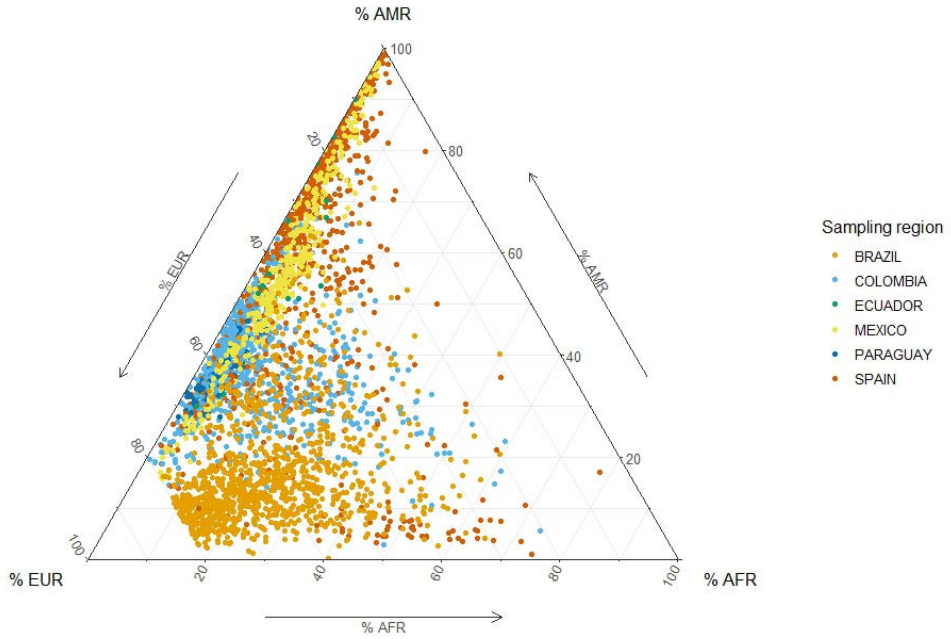
Supplementary Material for: Novel risk loci for COVID-19 hospitalization among admixed American populations

Supplementary Tables are provided in a separate excel file

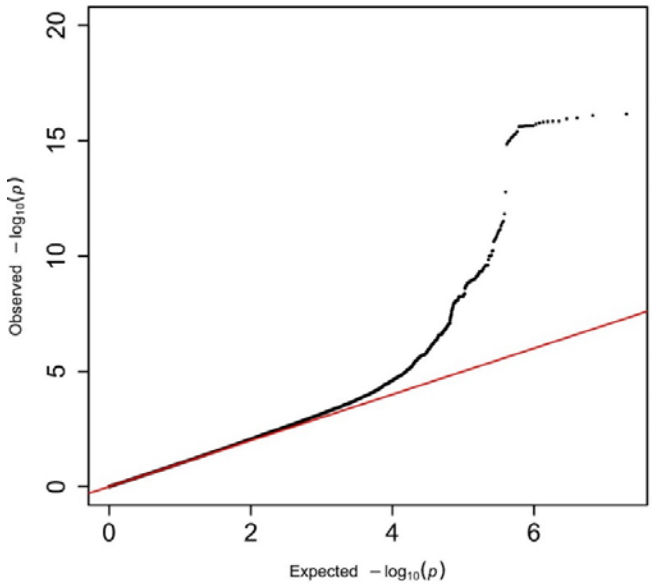
Supplementary figures

Supplementary Figure 1. Global Genetic Inferred Ancestry (GIA) composition in the SCOURGE Latin-American cohort. European (EUR), African (AFR) and Native American (AMR) GIA was derived with ADMIXTURE from a reference panel composed of Aymaran, Mayan, Nahuan, and Quechuan individuals of Native-American genetic ancestry and randomly selected samples from the EUR and AFR 1KGP populations. The colours represent the different geographical sampling regions from which the admixed American individuals from SCOURGE were recruited.

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



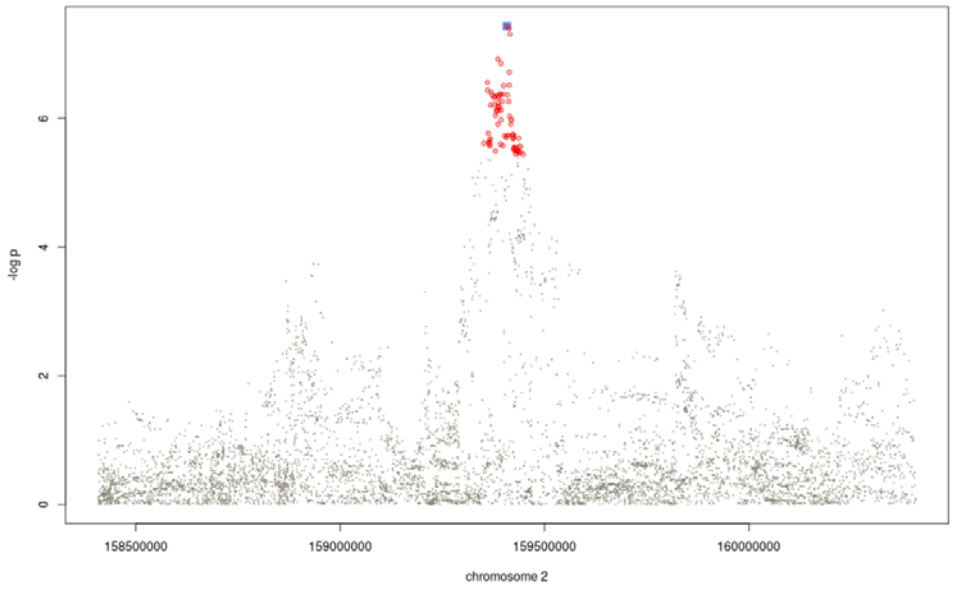
Supplementary Figure 2. Quantile-Quantile plot for the AMR GWAS meta-analysis. A lambda inflation factor of 1.015 was obtained.



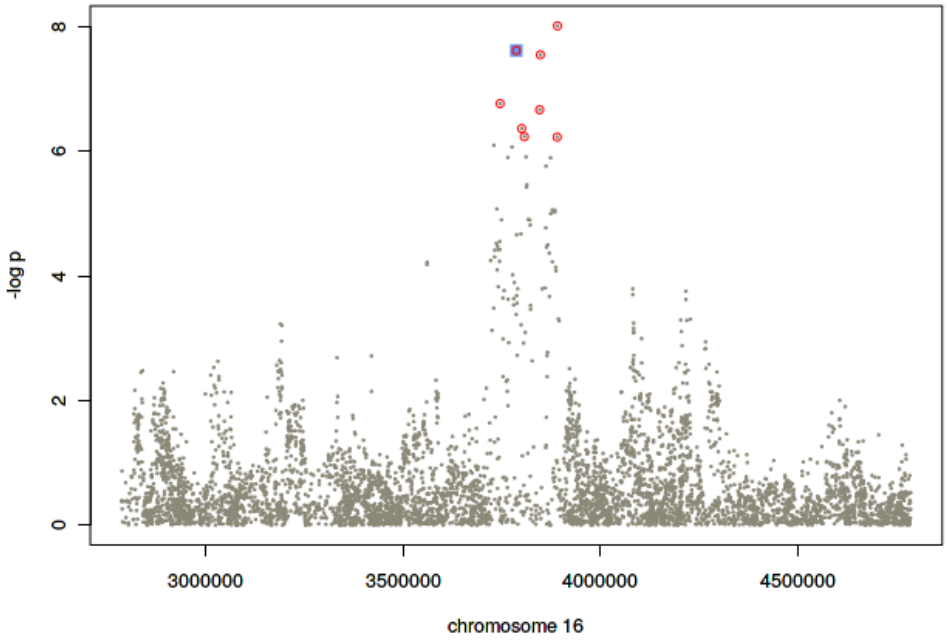
medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Supplementary Figure 3. Regional association plots for the fine mapped loci in chromosomes 2 (upper panel) and 16 (lower panel). Coloured in red, the variants allocated to the credible set at the 95% confidence according to the Bayesian fine mapping. In blue, the sentinel variant.

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



1058

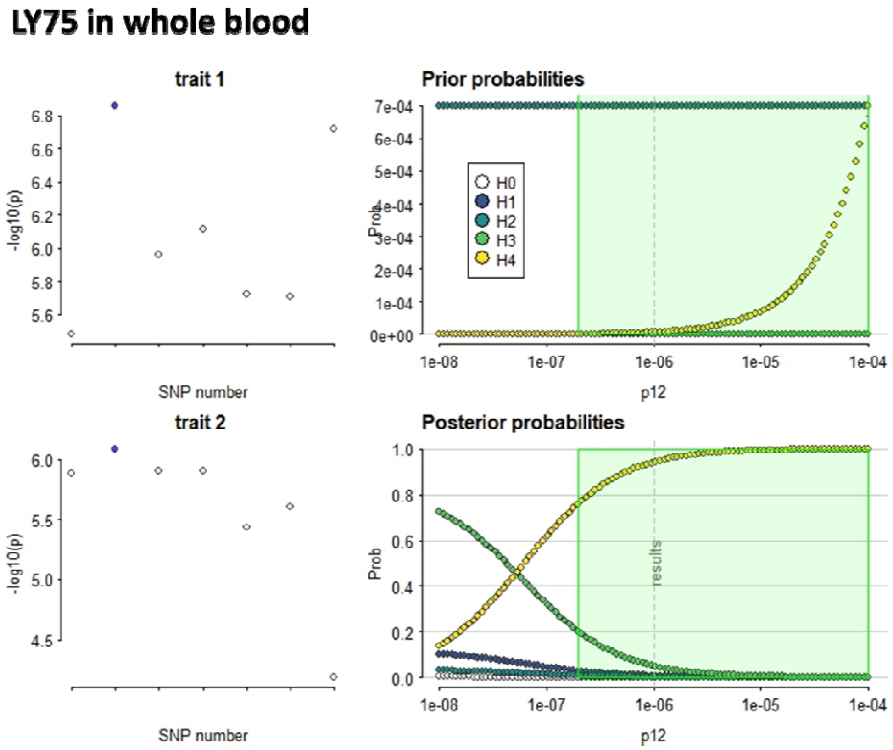


1059

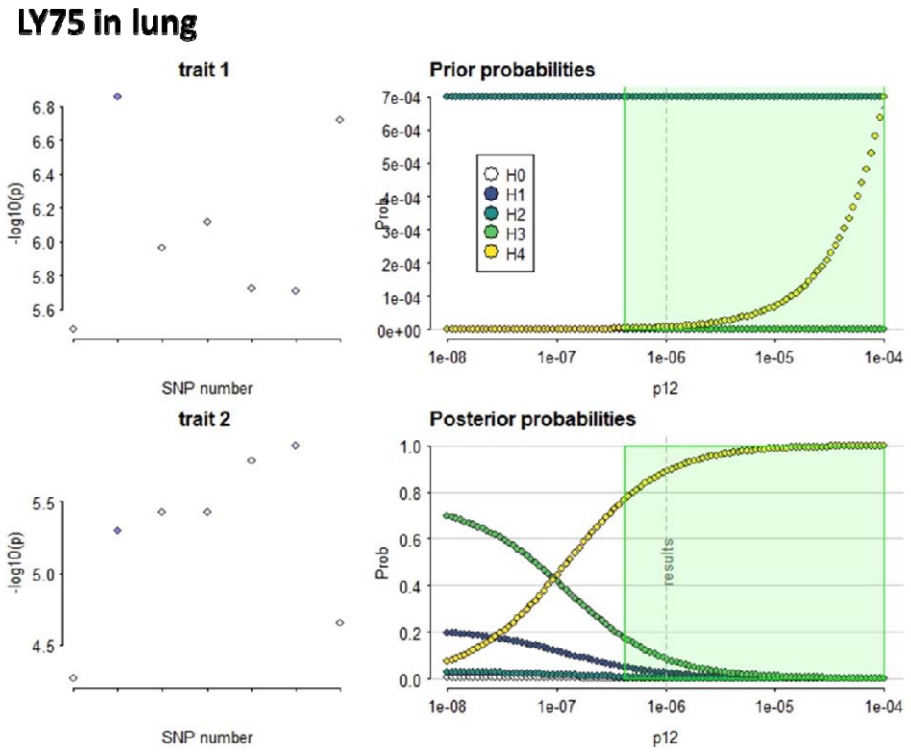
1060

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

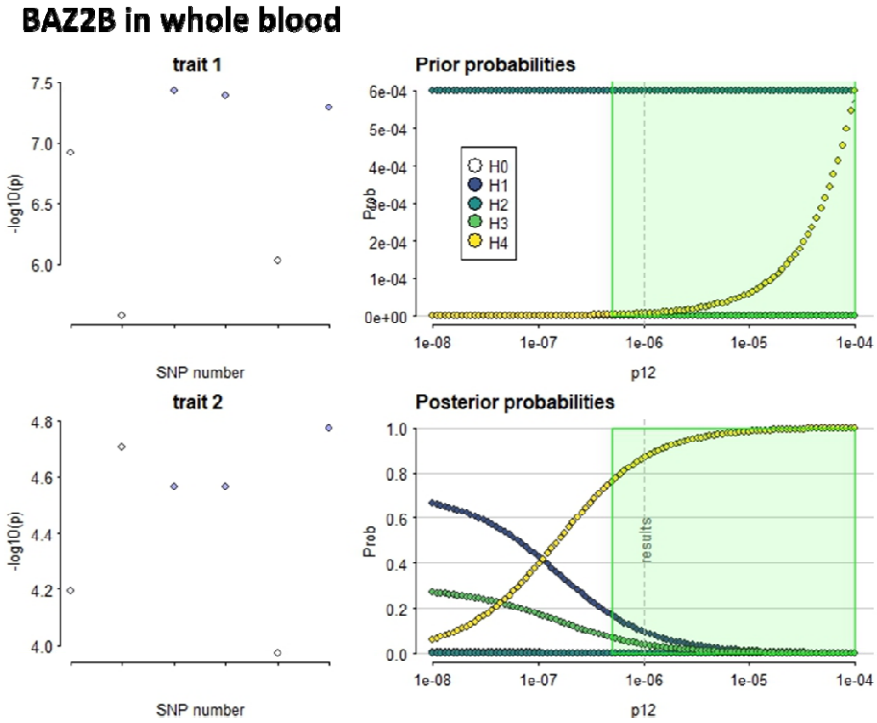
Supplementary Figure 4. Sensitivity plots from COLOC with expression data from GTEx v8. The range of p12 values (probability that a SNP is associated with both traits) for which the rule H_4>0.7 is supported is shown in green in the right plots for each analysis. Plots in the left represent the variants included in the risk region common to both traits along their individual association -log10(p-values) for each trait, whereas the shading shows the posterior probability that the SNP is causal given H_4 is true. Trait 1 corresponds to COVID-19 hospitalization, while trait 2 corresponds to gene expression in each analysis.



medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

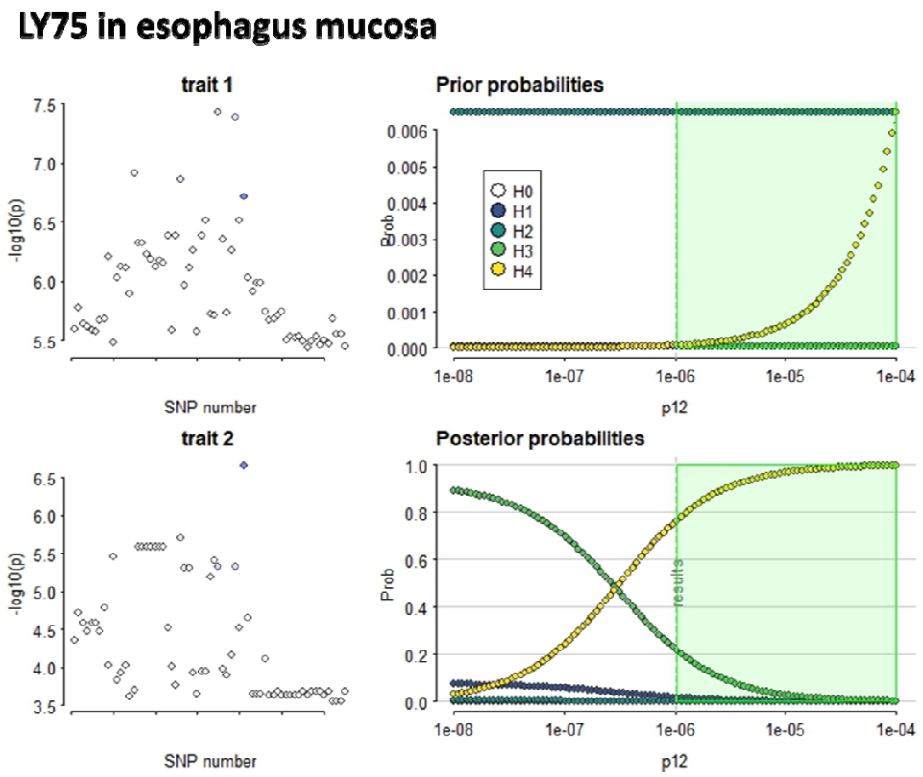


1071



1072

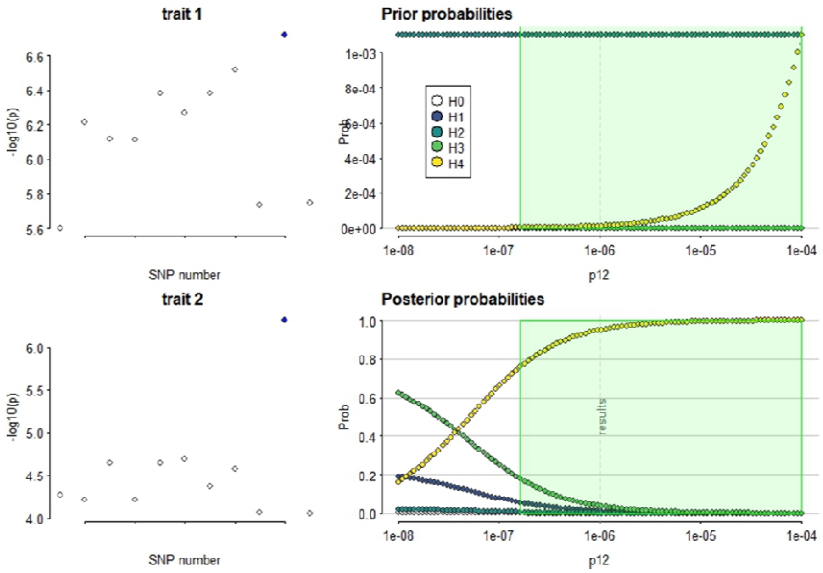
medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



Supplementary Figure 5. Sensitivity plots from COLOC with whole blood expression data from the GALA and SAGE II studies in AMR individuals. *AFRhp5* corresponds to the expression dataset computed in individuals with high African ancestries; *AMRhp5* corresponds to the expression dataset computed individuals with high AMR ancestries; *pooled* corresponds to the dataset computed with the total of individuals from the study. In the right, the plots show in green the range of p12 values (probability that a SNP is associated with both traits) for which the rule H_4>0.7 is supported. Plots in the left represent the variants included in the risk region common to both traits along their individual association -log10(p-values) for each trait, whereas the shading shows the posterior probability that the SNP is causal given H_4 is true. Trait 1 corresponds to COVID-19 hospitalization, while trait 2 corresponds to gene expression.

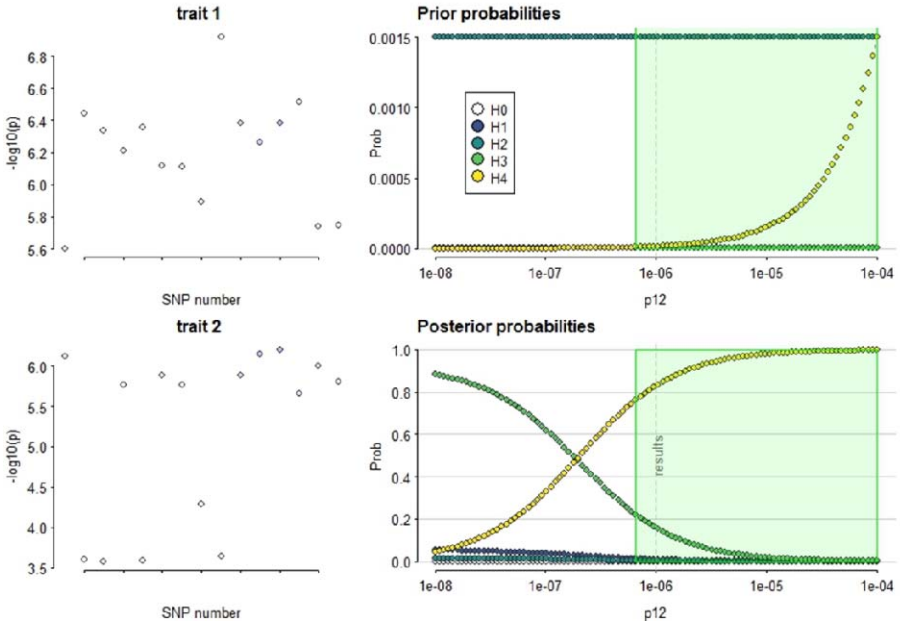
medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

CD302 in whole blood AFRhp5



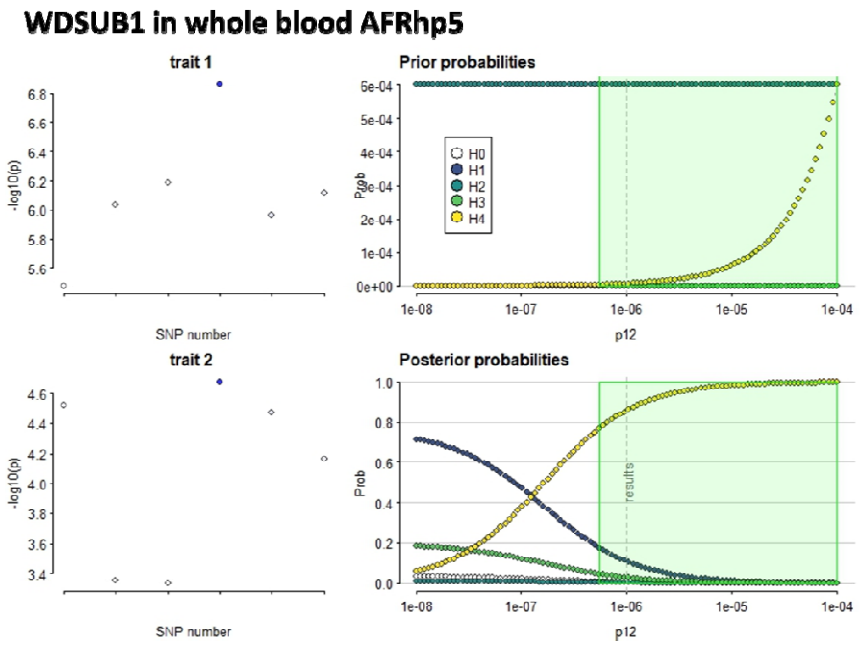
1086

BAZ2B in whole blood AFRhp5

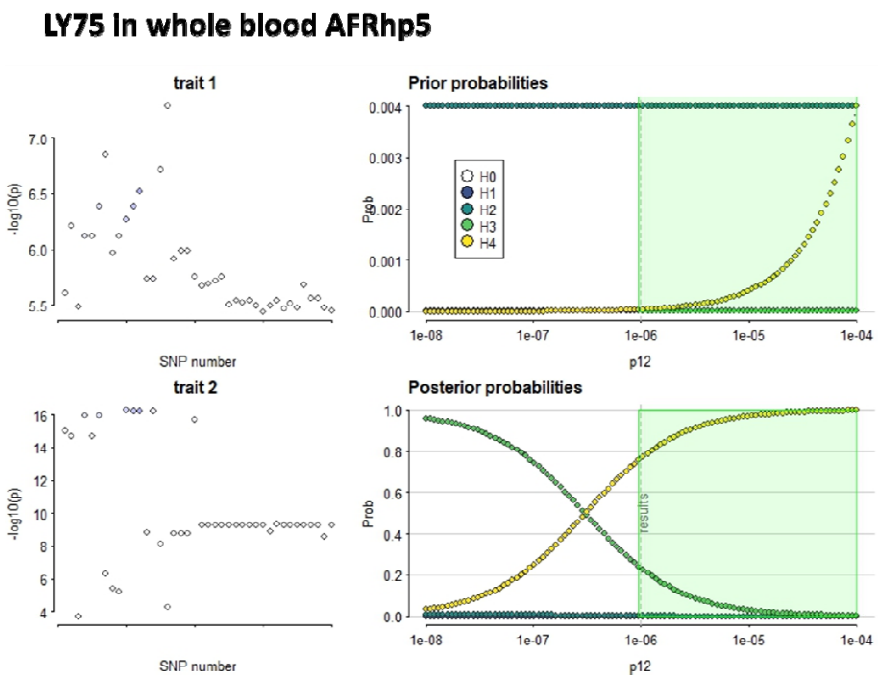


1087

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



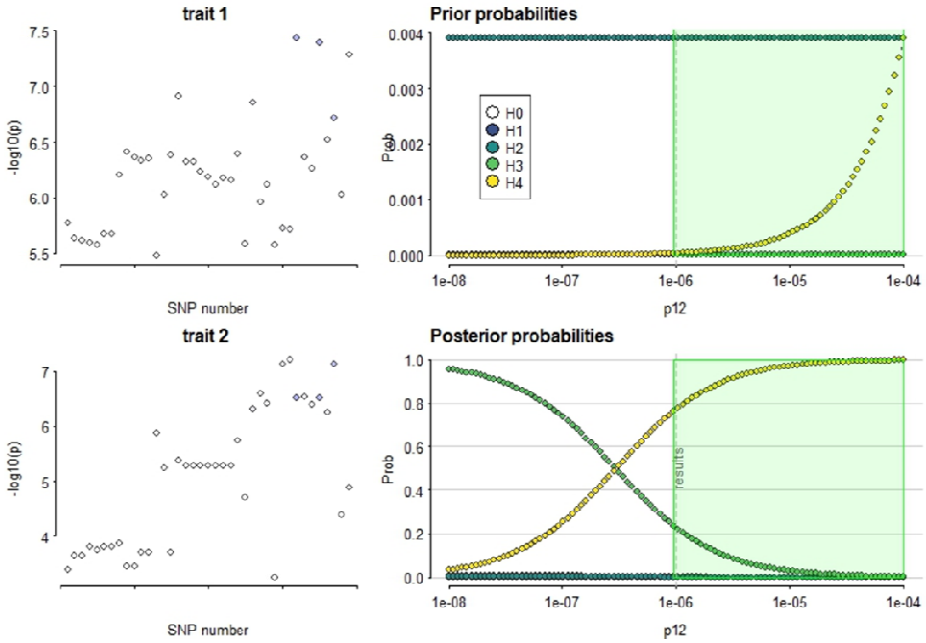
1088



1089

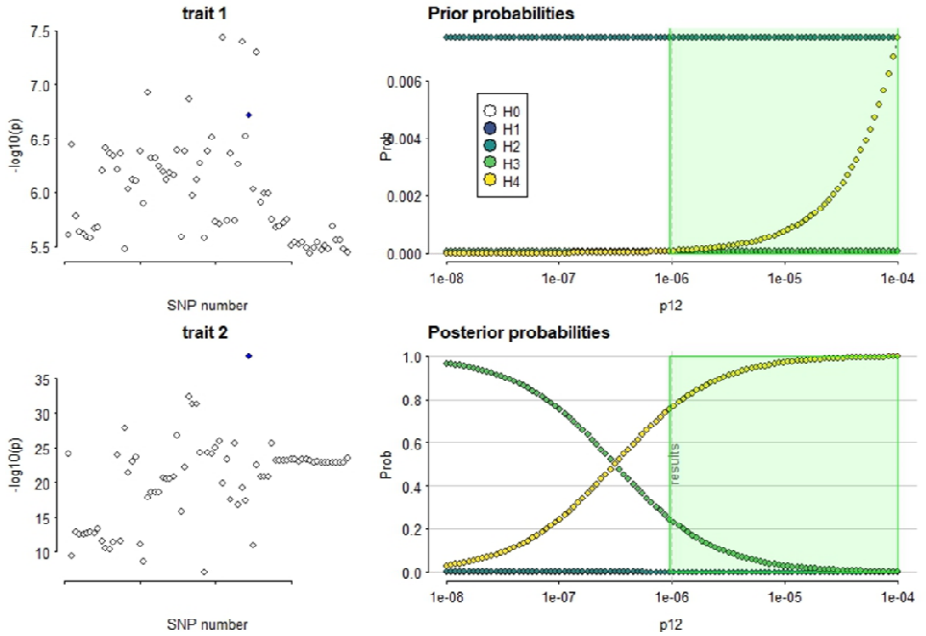
medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

LY75 in whole blood AMRhp5



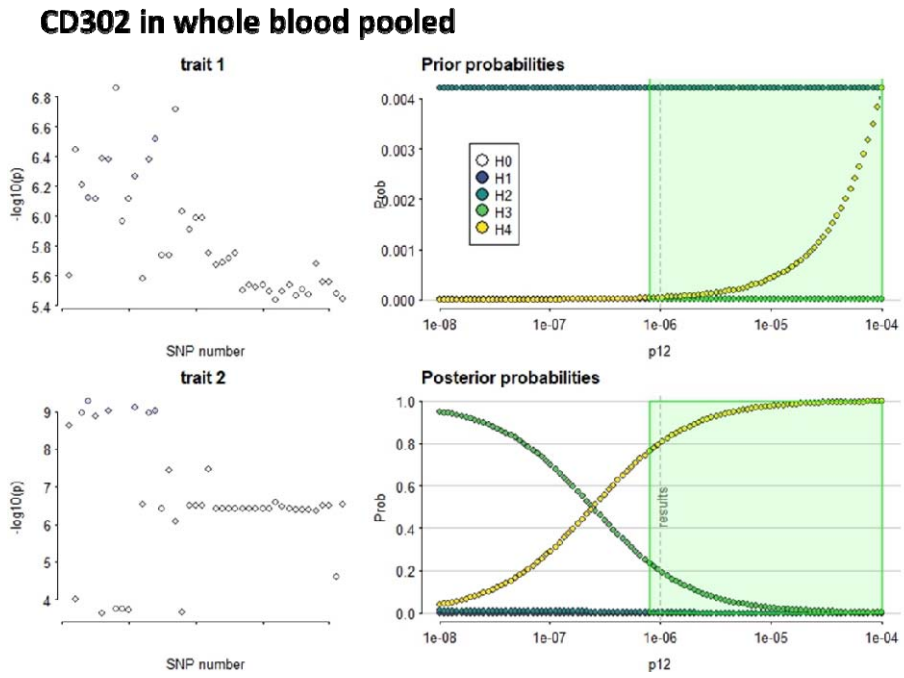
1090

LY75 in whole blood pooled



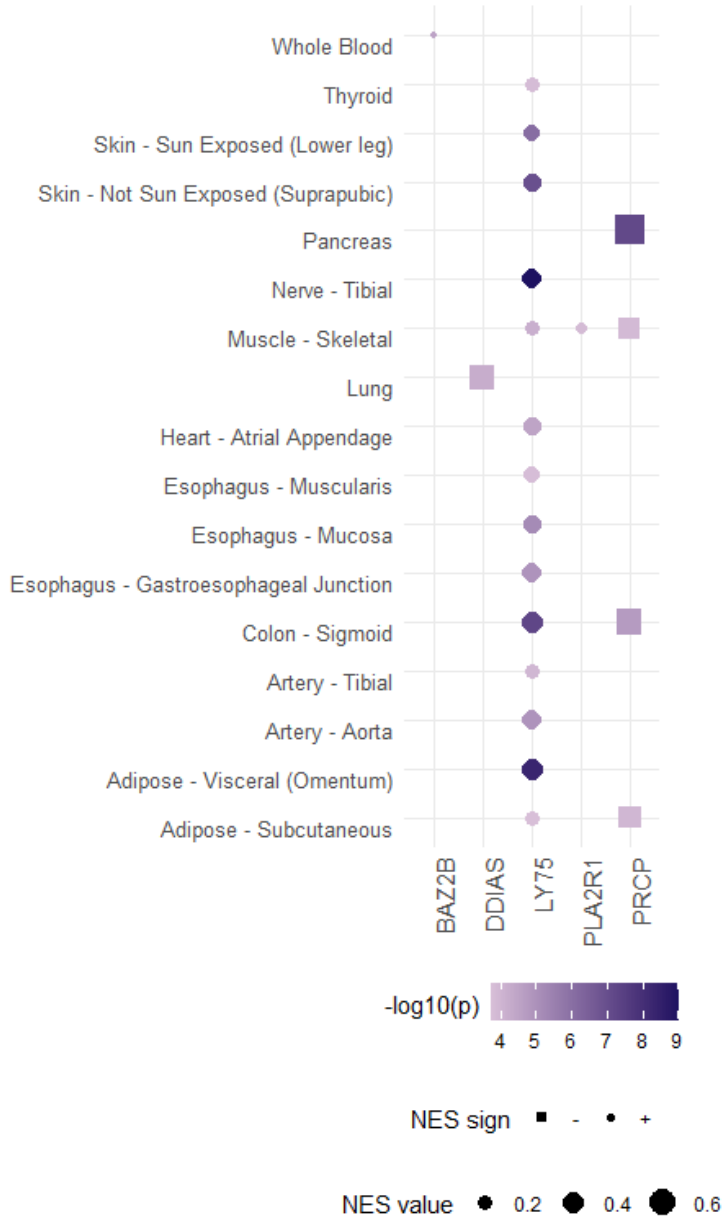
1091

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



Supplementary Figure 6. Gene-tissue pairs for which either rs1003835 or rs60606421 are significant eQTLs at FDR<0.05 (data retrieved from https://gtexportal.org/home/snp/). rs1003835 (chromosome 2) maps to *BAZ2B*, *LY75*, and *PLA2R* genes. As for the lead variant of chromosome 11, rs77599934, since it was not an eQTL, we used an LD proxy variant (rs60606421). *DDIAS* and *PRCP* genes map closely to this variant. NES and p-values correspond to the normalized effect size (and direction) of eQTL-gene associations and the p-value for the tissue, respectively.

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



Scourge Cohort Group

Full list of cohort members and affiliations

Javier Abellan[1,2]; René Acosta-Isaac[3]; Jose María Aguado[4,5,6,7]; Carlos Aguilar[8]; Sergio Aguilera-Albesa[9,10]; Abdolah Ahmadi Sabbagh[11]; Jorge Alba[12]; Sergiu Albu[13,14,15]; Karla A.M. Alcalá-Gallardo[16]; Julia Alcoba-Florez[17]; Sergio Alcolea Batres[18]; Holmes Rafael Algarin-Lara[19,20]; Virginia Almadana[21]; Julia Almeida[22,23,24,25]; Berta Almoguera[26,27]; María R. Alonso[28]; Nuria Alvarez[28]; Rodolfo Alvarez-Sala Walther[18]; Mónica T. Andrade [29,30]; Álvaro Andreu-Bernabeu[31,6]; Maria Rosa Antonijoan[32]; Eunate

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Arana-Arri[33,34]; Carlos Aranda[35,36]; Celso Arango[31,37,6]; Carolina Araque[38,39]; Nathalia K. Araujo[40]; Izabel M.T. Araujo[41]; Ana C. Arcanjo[42,43,44]; Ana Arnaiz[45,46,47]; Francisco Arnalich Fernández[48]; María J. Arranz[49]; José Ramon Arribas Lopez[48]; Maria-Jesus Artiga[50]; Yubelly Avello-Malaver[51]; Carmen Ayuso[26,27]; Ana Margarita Baldión[51]; Belén Ballina Martín[11]; Raúl C. Baptista-Rosas[52,53,54]; Andrea Barranco-Díaz[20]; María Barreda- Sánchez[55,56]; Viviana Barrera-Penagos[51]; Moncef Belhassen-Garcia[57,58]; Enrique Bernal[55]; David Bernal-Bello[59]; Joao F. Bezerra[60]; Marcos A.C. Bezerra[61]; Natalia Blanca-López[62]; Rafael Blancas[63]; Lucía Boix-Palop[64]; Alberto Borobia[65]; Elsa Bravo[66]; María Brion[67,68]; Óscar Brochado-Kith[69,7]; Ramón Brugada[70,71,68,72]; Matilde Bustos[73]; Alfonso Cabello[74]; Juan J. Caceres-Agra[75]; Esther Calbo[76]; Enrique J. Calderón[77,78,79]; Shirley Camacho[80]; Cristina Carbonell[81,58]; Servando Cardona-Huerta[82]; Antonio Augusto F. Carioca[83]; Maria Sanchez Carpintero[35,36]; Carlos Carpio Segura[18]; Thássia M.T. Carratto[84]; José Antonio Carrillo-Avila[85]; Maria C.C. Carvalho[86]; Carlos Casasnovas[87,88,27]; Luis Castano[33,34,27,89,90]; Carlos F. Castaño[35,36]; Jose E. Castelao[91]; Aranzazu Castellano Candalija[92]; María A. Castillo[80]; Yolanda Cañadas[36]; Francisco C. Ceballos[27]; Jessica G. Chaux[39]; Walter G. Chaves-Santiago[93,39]; Sylena Chiquillo-Gómez[19,20]; Marco A. Cid-Lopez[16]; Oscar Cienfuegos-Jimenez[82]; Rosa Conde-Vicente[94]; M. Lourdes Cordero-Lorenzana[95]; Dolores Corella[96,97]; Almudena Corrales[98,99]; Jose L. Cortes-Sanchez[82,100]; Marta Corton[26,27]; Tatiana X. Costa[101]; Raquel Cruz[27,102]; Marina S. Cruz[40]; Luisa Cuesta[103]; Gabriela C.R. Cunha[104]; David Dalmau[105,76]; Raquel C.S. Dantas-Komatsu[40]; M. Teresa Darnaude[106]; Alba De Martino-Rodríguez[107,108]; Juan De la Cruz Troca[109,110,78]; Juan Delgado-Cuesta[111]; Aranzazu Diaz de Bustamante[106]; Covadonga M. Diaz-Caneja[31,37,6]; Beatriz Dietl[76]; Silvia Diz-de Almeida[27,102]; Elena Domínguez-Garrido[112]; Alice M. Duarte[41]; Anderson Díaz-Pérez[20]; Jose Echave-Sustaeta[113]; Rocío Eiros[114]; César O. Enciso-Olivera[38,39]; Gabriela Escudero[115]; Pedro Pablo España[116]; Gladys Mercedes Estigarribia Sanabria[117]; María Carmen Fariñas[45,46,47]; Marianne R. Fernandes[118,119]; Lidia Fernandez-Caballero[26,27]; María J. Fernandez-Nestosa[120]; Ramón Fernández[45,121]; Silvia Fernández Ferrero[11]; Yolanda Fernández Martínez[11]; Ana Fernández-Cruz[122]; Uxía Fernández-Robelo[123]; Amanda Fernández-Rodríguez[69,7]; Marta Fernández-Sampedro[45,47,46]; Ruth Fernández-Sánchez[26,27]; Tania Fernández-Villa[124,78]; Carmen Ferréndez Capitán[92]; Patricia Flores-Pérez[125]; Vicente Friaza[78,79]; Lácides Fuenmayor-Hernández[20]; Marta Fuertes Núñez[11]; Victoria Fumadó[126]; Ignacio Gadea[127]; Lidia Gagliardi[35,36]; Manuela Gago-Domínguez[128,129]; Natalia Gallego[130]; Cristina Galoppo[131]; Carlos Garcia-Cerrada[1,2,132]; Josefina Garcia-García[55]; Inés García[26,27]; Mercedes García[35,36]; Leticia García[35,36]; María Carmen García Torrejón[133,2]; Irene García-García[65]; Carmen García-Ibarbia[45,47,46]; Andrés C. García-Montero[134]; Ana García-Soidán[135]; Elisa García-Vázquez[55]; Aitor García-de-Vicuña[33,136]; Emiliano Garza-Frias[82]; Jesus Gaytán-Martínez[137], Angela Gentile[131]; Belén Gil-Fournier[138]; Fernan Gonzalez Bernaldo de Quirós[139]; Manuel Gonzalez-Sagrado[94]; Hugo Gonzalo Benito[140]; Beatriz González Álvarez[107,108]; Anna González-Neira[28]; Javier González-Peñas[31,6,37]; Oscar Gorgojo-Galindo[141]; Florencia Guaragna[131]; Genilson P. Guegel[142]; Beatriz Guillen-Guio[98]; Encarna Guillen-Navarro[55,143,144,27]; Pablo Guisado-Vasco[113]; Luz D. Gutierrez-Castañeda[145,39]; Juan F. Gutiérrez-Bautista[146]; Luis Gómez Carrera[18]; María

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Gómez García[128]; Ángela Gómez Sacristán[147]; Javier Gómez-Arrue[107,108]; Mario Gómez-Duque[93,39]; Miguel Górgolas[74]; Sarah Heili-Frades[148]; Estefania Hernandez[149]; Luis D. Hernandez-Ortega[150,151]; Cristina Hernández Moro[11]; Guillermo Hernández-Pérez[81]; Rebeca Hernández-Vaquero[152]; Belen Herraez[28]; M. Teresa Herranz[55]; María Herrera[35,36]; María José Herrero[153,154]; Antonio Herrero-Gonzalez[155]; Juan P. Horcajada[156,157,14,158,7]; Natale Imaz-Ayo[33]; Maider Intxausti-Urrutibeaskoa[159]; Rafael H. Jacomo[160]; Rubén Jara[55]; Perez Maria Jazmin[131]; María A. Jimenez-Sousa[69,7]; Ángel Jiménez[35,36]; Pilar Jiménez[146]; Ignacio Jiménez-Alfaro[161]; Iolanda Jordan[162,163,78]; Rocío Laguna-Goya[164,165]; Daniel Laorden[18]; María Lasa-Lazaro[164,165]; María Claudia Lattig[80,166]; Ailen Lauriente[131]; Anabel Liger Borja[167]; Lucía Llanos[169]; Esther Lopez-Garcia[109,110,78,170]; Rosario Lopez-Rodriguez[26,27]; Leonardo Lorente[171]; José E. Lozano[172]; María Lozano-Espinosa[167]; Andre D. Luchessi[173]; Eduardo López Granados[174,175,27]; Amparo López-Bernús[81,58]; Miguel A. López-Ruz[176,177,178]; Aluísio X. Magalhães-Brasil [179];Ignacio Mahillo[180,181,99]; Esther Mancebo[164,165]; Carmen Mar[116]; Cristina Marcelo Calvo[92]; Miguel Marcos[81,58]; Alba Marcos-Delgado[124]; Pablo Mariscal Aguilar[18]; Marta Martin-Fernandez[182]; Laura Martin-Pedraza[62]; Amalia Martinez[183]; Iciar Martinez-Lopez[184,185]; Oscar Martinez-Nieto[51,166]; Pedro Martinez-Paz[140]; Angel Martinez-Perez[186]; Michel F. Martinez-Resendez[82]; María M. Martín[187]; María Dolores Martín[188]; Vicente Martín[124,78]; Caridad Martín-López[167]; José-Ángel Martín-Oterino[81,58]; María Martín-Vicente[69]; Ricardo Martínez[149]; Juan José Martínez[88,27]; Silvia Martínez[45,47]; Violeta Martínez Robles[11]; Eleno Martínez-Aquino[189]; Óscar Martínez-González[190]; Andrea Martínez-Ramas[26,27]; Laura Marzal[26,27]; Alicia Marín Candon[65]; Jose Antonio Mata-Marin,[137] Juliana F. Mazzeu[179,191,192]; Jeane F.P. Medeiros[40]; Francisco J. Medrano[77,78,79]; Xose M. Meijome[193,194]; Natalia Mejuto-Montero[195]; Celso T. Mendes-Junior [84,196,197]; Humberto Mendoza Charris[66,20]; Eleuterio Merayo Macías[198]; Fátima Mercadillo[199]; Arieh R. Mercado-Sesma[150,151]; Pablo Minguez[26,27]; Antonio J J. Molina[124,78]; Elena Molina-Roldán[200]; Juan José Montoya[149]; Patricia Moreira-Escriche[201]; Xenia Morelos-Arnedo[66,20]; Victor Moreno Cuerda[1,2]; Alberto Moreno Fernández[92]; Antonio Moreno-Docón[55]; Junior Moreno-Escalante[20]; Rubén Morilla[79,202]; Patricia Muñoz García[203,99,6]; Ana Méndez-Echevarria[204]; Pablo Neira[131]; Julian Nevado[27,131,205]; Israel Nieto-Gañán[135]; Joana F.R. Nunes[42]; Rocio Nuñez- Torres[28]; Antònia Obrador-Hevia[206,207]; J. Gonzalo Ocejo-Vinyals[45,47]; Virginia Olivar[131]; Silviene F. Oliveira[179,208,208,210,211]; Lorena Ondo[26,27]; Alberto Orfao[22,23,24,25]; Luis Ortega[212]; Eva Ortega-Paino[50]; Fernando Ortiz-Flores[45,47]; Rocio Ortiz-Lopez[213,82]; José A. Oteo[12,214]; Harry Pachajoa[215,216]; Manuel Pacheco[149]; Fredy Javier Pacheco-Miranda[20]; Irene Padilla Conejo[11]; Sonia Panadero-Fajardo[85]; Mara Parellada[31,37,6]; Roberto Pariente-Rodríguez[131]; Estela Paz-Artal[164,165,217]; Germán Peces-Barba[218,99]; Miguel S. Pedromingo Kus[219]; Celia Perales[127]; Patricia Perez[220]; Gustavo Perez-de-Nanclares[33,221]; Teresa Perucho[222]; Aline Pic-Taylor [42,208,209,211], Lisbeth A. Pichardo[11]; Mel·lina Pinsach-Abuin[70,68]; Luz Adriana Pinzón[93,39]; Guillermo Pita[30]; Francesc Pla-Junca[223,27]; Laura Planas-Serra[88,27]; Ericka N. Pompa-Mera[224,137] ; Gloria L. Porras-Hurtado[149]; Aurora Pujol[88,27,225]; César Pérez[226]; Felipe Pérez-García[227,228]; Patricia Pérez-Matute[214]; Alexandra Pérez-Serra[70,68]; M. Elena Pérez-Tomás[55]; María Eugenia Quevedo Chávez[19,20]; Maria Angeles Quijada[30,229]; Inés Quintela[128]; Diana Ramirez-

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Montaño[230]; Soraya Ramiro León[131]; Pedro Rascado Sedes[231]; Delia Recalde[107,108]; Emma Recio-Fernández[214]; Salvador Resino[69,7]; Adriana P. Ribeiro[29,30,232]; Carlos S. Rivadeneira-Chamorro[39]; Diana Roa-Agudelo[51]; Montserrat Robelo Pardo[231]; Marilyn Johanna Rodriguez[39]; German Ezequiel Rodriguez Novoa[131]; Fernando Rodriguez-Artalejo[109,110,78,170]; Carlos Rodriguez-Gallego[233,234]; José A. Rodriguez-Garcia[11]; María A. Rodriguez-Hernandez[73]; Antonio Rodriguez-Nicolas[146]; Agustí Rodriguez-Palmero[235,88]; Paula A. Rodriguez-Urrego[51]; Belén Rodríguez Maya[1]; Marena Rodríguez-Ferrer[20]; Emilio Rodríguez-Ruiz[231,129]; Federico Rojo[236,25]; Andrea Romero-Coronado[20]; Filomeno Rondón García[11]; ~~Lidia S. Rosa[236]~~; Antonio Rosales-Castillo[237]; Cladelis Rubio[238,239]; María Rubio Olivera[35,36]; Montserrat Ruiz[88,27]; Francisco Ruiz-Cabello[146,177,240]; Eva Ruiz-Casares[222]; Juan J. Ruiz-Cubillan[45,47]; Javier Ruiz-Hornillos[241,36,242]; Pablo Ryan[243,244,245,7]; Hector D. Salamanca[38,39]; Lorena Salazar-García[80]; Giorgina Gabriela Salgueiro Origlia [92]; Cristina Sancho-Sainz[159]; Jorge Luis Sandoval-Ramírez,[137] Anna Sangil[64]; Arnoldo Santos[226]; Ney P.C. Santos[118]; Amanda C.M. Saúde [30,246] Agatha Schlüter[88,27]; Sonia Segovia[223,247,248]; Alex Serra-Llovich[249]; Fernando Sevil Puras[8]; Marta Sevilla Porras[27,130]; Miguel A. Sicolo[250,251]; Vivian N. Silbiger[173]; Nayara S. Silva[252]; ~~Fabiola T.C. Silva[40]~~; Cristina Silván Fuentes[27]; Jordi Solé-Violán[253,99,254]; José Manuel Soria[186]; Jose V. Sorlí[96,97]; Renata R. Sousa[179]; Juan Carlos Souto[3]; Karla S.C. Souza[86]; Vanessa S. Souza[104]; John J. Sprockel[93,39]; David A. Suarez-Zamora[51]; José Javier Suárez-Rama[128]; Pedro-Luis Sánchez[114,58]; Antonio J. Sánchez López[255]; María Concepción Sánchez Prados[18]; Javier Sánchez Real[11]; Jorge Sánchez Redondo[1,256]; Clara Sánchez-Pablo[114]; Olga Sánchez-Pernaute[257]; Xiana Taboada-Fraga[195]; Eduardo Tamayo[258,140,7]; Alvaro Tamayo-Velasco[259]; Juan Carlos Taracido-Fernandez[155]; Nathali A.C. Tavares[260]; Carlos Tellería[107,108]; Jair Antonio Tenorio Castaño[27,130,205]; Alejandro Teper[131]; Ronald P. Torres Gutiérrez[221]; Juan Torres-Macho[261]; Lilian Torres-Tobar[39]; Jesús Troya[243]; Miguel Urioste[199]; Juan Valencia-Ramos[262]; Agustín Valido[21,263]; Juan Pablo Vargas Gallo[264,265]; Belén Varón[266]; Romero H.T. Vasconcelos[260]; Tomas Vega[267]; Santiago Velasco-Quirce[268]; Julia Vidán Estévez[11]; Miriam Vieitez-Santiago[45,47]; Carlos Vilches[269]; Lavinia Villalobos[11]; Felipe Villar[218]; Judit Villar-Garcia[270,271,272]; Cristina Villaverde[26,27]; Pablo Villoslada-Blanco[214]; Ana Virseda-Berdices[69]; Valentina Vélez-Santamaría[87,88]; Virginia Víctor[35,36]; Zuleima Yáñez[20]; Antonio Zapatero-Gaviria[273]; Ruth Zarate[274]; Sandra Zazo[236]; Gabriela V. da Silva[41]; Raimundo de Andrés[275]; Jéssica N.G. de Araújo[252]; Carmen de Juan[201]; Julianna Lys de Sousa Alves Neri[276]; Carmen de la Horra[79]; Ana B. de la Hoz[33]; Victor del Campo-Pérez[277]; Manoella do Monte Alves[278,279]; Katiusse A. dos Santos[86]; Yady Álvarez-Benítez[19,20]; Felipe Álvarez-Navia[81,58]; María Íñiguez[214]; Miguel López de Heredia[27]; Ingrid Mendes[27]; Rocío Moreno[27]; Esther Sande[27,129,102]; Carlos Flores[280,98,99,234]; José A. Riancho[45,46,47,27]; Augusto Rojas-Martinez[82]; Pablo Lapunzina[27,130,205]; Angel Carracedo[27,129,102,128]

Scourge Cohort Group's filiations

[1], Hospital Universitario Mostoles, Medicina Interna, Madrid, Spain

[2], Universidad Francisco de Vitoria, Madrid, Spain

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[3], Haemostasis and Thrombosis Unit, Hospital de la Santa Creu i Sant Pau, IIB Sant Pau, Barcelona, Spain

[4], Unit of Infectious Diseases, Hospital Universitario 12 de Octubre, Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas12), Madrid, Spain

[5], Spanish Network for Research in Infectious Diseases (REIPI RD16/0016/0002), Instituto de Salud Carlos III, Madrid, Spain

[6], School of Medicine, Universidad Complutense, Madrid, Spain

[7], Centro de Investigación Biomédica en Red de Enfermedades Infecciosas (CIBERINFEC), Instituto de Salud Carlos III, Madrid, Spain

[8], Hospital General Santa Bárbara de Soria, Soria, Spain

[9], Pediatric Neurology Unit, Department of Pediatrics, Navarra Health Service Hospital, Pamplona, Spain

[10], Navarra Health Service, NavarraBioMed Research Group, Pamplona, Spain

[11], Complejo Asistencial Universitario de León, León, Spain

[12], Hospital Universitario San Pedro, Infectious Diseases Department, Logroño, Spain

[13], Fundación Institut Guttmann, Institut Universitari de Neurorehabilitació adscrit a la UAB, Hospital de Neurorehabilitació, Barcelona, Spain

[14], Universitat Autònoma de Barcelona (UAB), Barcelona, Spain

[15], Fundació Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol, Barcelona, Spain

[16], Hospital General de Occidente, Guadalajara, Mexico

[17], Microbiology Unit, Hospital Universitario N. S. de Candelaria, Santa Cruz de Tenerife, Spain

[18], Hospital Universitario La Paz-IDIPAZ, Servicio de Neumología, Madrid, Spain

[19], Camino Universitario Adelita de Char, Mired IPS, Barranquilla, Colombia

[20], Universidad Simón Bolívar, Facultad de Ciencias de la Salud, Barranquilla, Colombia

[21], Hospital Universitario Virgen Macarena, Neumología, Seville, Spain

[22], Departamento de Medicina, Universidad de Salamanca, Salamanca, Spain

[23], Centro de Investigación del Cáncer (IBMCC) Universidad de Salamanca - CSIC, Salamanca, Spain

[24], Biomedical Research Institute of Salamanca (IBSAL) Salamanca, Spain

[25], Centre for Biomedical Network Research on Cancer (CIBERONC), Instituto de Salud Carlos III, Madrid, Spain

[26], Department of Genetics & Genomics, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[27], Centre for Biomedical Network Research on Rare Diseases (CIBERER), Instituto de Salud Carlos III, Madrid, Spain

[28], Spanish National Cancer Research Centre, Human Genotyping-CEGEN Unit, Madrid, Spain

[29], Hospital das Forças Armadas, Brazil

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[30], Exército Brasileiro, Brazil

[31], Department of Child and Adolescent Psychiatry, Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón (IiSGM), Madrid, Spain

[32], Clinical Pharmacology Service, Hospital de la Santa Creu i Sant Pau, IIB Sant Pau, Barcelona, Spain

[33], Biocruces Bizkai HRI, Barakaldo, Bizkaia, Spain

[34], Cruces University Hospital, Osakidetza, Barakaldo, Bizkaia, Spain

[35], Hospital Infanta Elena, Valdemoro, Madrid, Spain

[36], Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[37], Centre for Biomedical Network Research on Mental Health (CIBERSAM), Instituto de Salud Carlos III, Madrid, Spain

[38], Fundación Hospital Infantil Universitario de San José, Bogotá, Colombia

[39], Fundación Universitaria de Ciencias de la Salud, Bogotá, Colombia

[40], Universidade Federal do Rio Grande do Norte, Programa de Pós-graduação em Ciências da Saúde, Natal, Brazil

[41], Universidade Federal do Rio Grande do Norte, Departamento de Medicina Clínica, Natal, Brazil

[42], Departamento de Genética e Morfologia, Instituto de Ciências Biológicas, Universidade de Brasília, Brasilia, Brazil

[43], Colégio Marista de Brasilia, Brazil

[44], Associação Brasileira de Educação e Cultura, Brazil

[45], IDIVAL, Santander, Spain

[46], Universidad de Cantabria, Santander, Spain

[47], Hospital U M Valdecilla, Santander, Spain

[48], Hospital Universitario La Paz-IDIPAZ, Servicio de Medicina Interna, Madrid, Spain

[49], Fundació Docència I Recerca Mutua Terrassa, Barcelona, Spain

[50], Spanish National Cancer Research Center, CNIO Biobank, Madrid, Spain

[51], Fundación Santa Fe de Bogota, Departamento Patologia y Laboratorios, Bogotá, Colombia

[52], Hospital General de Occidente, Zapopan, Jalisco, Mexico

[53], Centro Universitario de Tonalá, Universidad de Guadalajara, Tonalá, Jalisco, Mexico

[54], Centro de Investigación Multidisciplinario en Salud, Universidad de Guadalajara, Tonalá, Jalisco, Mexico

[55], Instituto Murciano de Investigación Biosanitaria (IMIB-Arrixaca), Murcia, Spain

[56], Universidad Católica San Antonio de Murcia (UCAM), Murcia, Spain

[57], Hospital Universitario de Salamanca-IBSAL, Servicio de Medicina Interna-Unidad de Enfermedades Infecciosas, Salamanca, Spain

[58], Universidad de Salamanca, Salamanca, Spain

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[59], Hospital Universitario de Fuenlabrada, Department of Internal Medicine, Madrid, Spain

[60], Escola Tecnica de Saúde, Laboratorio de Vigilancia Molecular Aplicada, Pará, Brazil

[61], Federal University of Pernambuco, Genetics Postgraduate Program, Recife, PE, Brazil

[62], Hospital Universitario Infanta Leonor, Servicio de Alergia, Madrid, Spain

[63], Hospital Universitario del Tajo, Servicio de Medicina Intensiva, Aranjuez, Spain

[64], Hospital Universitario Mutua Terrassa, Barcelona, Spain

[65], Hospital Universitario La Paz-IDIPAZ, Servicio de Farmacología, Madrid, Spain

[66], Alcaldía de Barranquilla, Secretaría de Salud, Barranquilla, Colombia

[67], Instituto de Investigación Sanitaria de Santiago (IDIS), Xenética Cardiovascular, Santiago de Compostela, Spain

[68], Centre for Biomedical Network Research on Cardiovascular Diseases (CIBERCV), Instituto de Salud Carlos III, Madrid, Spain

[69], Unidad de Infección Viral e Inmunidad, Centro Nacional de Microbiología (CNM), Instituto de Salud Carlos III (ISCIII), Madrid, Spain

[70], Cardiovascular Genetics Center, Institut d'Investigació Biomèdica Girona (IDIBGI), Girona, Spain

[71], Medical Science Department, School of Medicine, University of Girona, Girona, Spain

[72], Hospital Josep Trueta, Cardiology Service, Girona, Spain

[73], Institute of Biomedicine of Seville (IBiS), Consejo Superior de Investigaciones Científicas (CSIC)- University of Seville- Virgen del Rocio University Hospital, Seville, Spain

[74], Division of Infectious Diseases, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[75], Intensive Care Unit, Hospital Universitario Insular de Gran Canaria, Las Palmas de Gran Canaria, Spain

[76], Hospital Universitario Mutua Terrassa, Terrassa, Spain

[77], Departamento de Medicina, Hospital Universitario Virgen del Rocío,Universidad de Sevilla, Seville, Spain

[78], Centre for Biomedical Network Research on Epidemiology and Public Health (CIBERESP), Instituto de Salud Carlos III, Madrid, Spain

[79], Institute of Biomedicine of Seville (IBiS), Consejo Superior de Investigaciones Científicas (CSIC)- University of Seville- Virgen del Rocio University Hospital, Seville, Spain

[80], Universidad de los Andes, Facultad de Ciencias, Bogotá, Colombia

[81], Hospital Universitario de Salamanca-IBSAL, Servicio de Medicina Interna, Salamanca, Spain

[82], Tecnologico de Monterrey, Escuela de Medicina y Ciencias de la Salud and Hospital San Jose TecSalud, Monterrey, Mexico

[83], University of Fortaleza (UNIFOR), Department of Nutrition. Fortaleza, Brazil

[84], Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Brazil

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[85], Andalusian Public Health System Biobank, Granada, Spain

[86], Universidade Federal do Rio Grande do Norte, Programa de Pós-Graduação em Ciências Farmacêuticas, Natal, Brazil

[87], Neuromuscular Unit, Neurology Department, Hospital Universitari de Bellvitge, L'Hospitalet de Llobregat (Barcelona), Spain

[88], Bellvitge Biomedical Research Institute (IDIBELL), Neurometabolic Diseases Laboratory, L'Hospitalet de Llobregat, Spain

[89], Centre for Biomedical Network Research on Diabetes and Metabolic Associated Diseases (CIBERDEM), Instituto de Salud Carlos III, Madrid, Spain

[90], University of Pais Vasco, UPV/EHU, Bizkaia, Spain

[91], Oncology and Genetics Unit, Instituto de Investigacion Sanitaria Galicia Sur, Xerencia de Xestion Integrada de Vigo-Servizo Galego de Saúde, Vigo, Spain

[92], Hospital Universitario La Paz, Hospital Carlos III, Madrid, Spain

[93], Hospital de San José, Sociedad de Cirugía de Bogota, Bogotá, Colombia

[94], Hospital Universitario Río Hortega, Valladolid, Spain

[95], Servicio de Medicina intensiva, Complejo Hospitalario Universitario de A Coruña (CHUAC), Sistema Galego de Saúde (SERGAS), A Coruña, Spain

[96], Valencia University, Preventive Medicine Department, Valencia, Spain

[97], Centre for Biomedical Network Research on Physiopatology of Obesity and Nutrition (CIBEROBN), Instituto de Salud Carlos III, Madrid, Spain

[98], Research Unit, Hospital Universitario N.S. de Candelaria, Santa Cruz de Tenerife, Spain

[99], Centre for Biomedical Network Research on Respiratory Diseases (CIBERES), Instituto de Salud Carlos III, Madrid, Spain

[100], Otto von Guericke University, Departament of Microgravity and Translational Regenerative Medicine, Magdeburg, Germany

[101], Maternidade Escola Janário Cicco, Natal, Brazil

[102], Centro Singular de Investigación en Medicina Molecular y Enfermedades Crónicas (CIMUS), Universidade de Santiago de Compostela, Santiago de Compostela, Spain

[103], Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón (IiSGM), Madrid, Spain

[104], Programa de Pós Graduação em Ciências da Saúde, Faculdade de Medicina, Universidade de Brasília, Brasilia, Brazil

[105], Fundació Docència I Recerca Mutua Terrassa, Terrassa, Spain

[106], Hospital Universitario Mostoles, Unidad de Genética, Madrid, Spain

[107], Instituto Aragonés de Ciencias de la Salud (IACS), Zaragoza, Spain

[108], Instituto Investigación Sanitaria Aragón (IIS-Aragon), Zaragoza, Spain

[109], Department of Preventive Medicine and Public Health, School of Medicine, Universidad Autónoma de Madrid, Madrid, Spain

[110], IdiPaz (Instituto de Investigación Sanitaria Hospital Universitario La Paz), Madrid, Spain

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[111], Hospital Universitario Virgen del Rocío, Servicio de Medicina Interna, Seville, Spain

[112], Unidad Diagnóstico Molecular. Fundación Rioja Salud, La Rioja, Spain

[113], Hospital Universitario Quironsalud Madrid, Madrid, Spain

[114], Hospital Universitario de Salamanca-IBSAL, Servicio de Cardiología, Salamanca, Spain

[115], Hospital Universitario Puerta de Hierro, Servicio de Medicina Interna, Majadahonda, Spain

[116], Biocruces Bizkaia Health Research Institute, Galdakao University Hospital, Osakidetza, Bizkaia, Spain

[117], Instituto Regional de Investigación en Salud-Universidad Nacional de Caaguazú, Caaguazú, Paraguay

[118], Universidade Federal do Pará, Núcleo de Pesquisas em Oncologia, Belém, Pará, Brazil

[119], Hospital Ophir Loyola, Departamento de Ensino e Pesquisa, Belém, Pará, Brazil

[120], Universidad Nacional de Asunción, Facultad de Politécnica, Paraguay

[121], Fundación Asilo San Jose, Santander, Spain

[122], Unidad de Enfermedades Infecciosas, Servicio de Medicina Interna, Hospital Universitario Puerta de Hierro, Instituto de Investigación Sanitaria Puerta de Hierro - Segovia de Arana, Madrid, Spain

[123], Urgencias Hospitalarias, Complejo Hospitalario Universitario de A Coruña (CHUAC), Sistema Galego de Saúde (SERGAS), A Coruña, Spain

[124], Grupo de Investigación en Interacciones Gen-Ambiente y Salud (GIIGAS) - Instituto de Biomedicina (IBIOMED), Universidad de León, León, Spain

[125], Hospital Universitario Niño Jesús, Pediatrics Department, Madrid, Spain

[126], Unitat de Malalties Infeccioses i Importades, Servei de Pediatría, Infectious and Imported Diseases, Pediatric Unit, Hospital Universitari Sant Joan de Deú, Barcelona, Spain

[127], Microbiology Department, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[128], Fundación Pública Galega de Medicina Xenómica, Sistema Galego de Saúde (SERGAS) Santiago de Compostela, Spain

[129], Instituto de Investigación Sanitaria de Santiago (IDIS), Santiago de Compostela, Spain

[130], Instituto de Genética Médica y Molecular (INGEMM), Hospital Universitario La Paz-IDIPAZ, Madrid, Spain

[131], Hospital de Niños Ricardo Gutierrez, Buenos Aires, Argentina

[132], Centre for Biomedical Network Research on Rare Diseases (CIBERER), Instituto de Salud Carlos III, Madrid, Spain Universidad Francisco de Vitoria, Madrid,Spain

[133], Hospital Infanta Elena, Servicio de Medicina Intensiva, Valdemoro, Madrid, Spain

[134], University of Salamanca, Biomedical Research Institute of Salamanca (IBSAL), Salamanca, Spain

[135], Department of Immunology, IRYCIS, Hospital Universitario Ramón y Cajal, Madrid, Spain

[136], Osakidetza, Cruces University Hospital, Bizkaia, Spain

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[137], Instituto Mexicano del Seguro Social, IMSS. Centro Médico Nacional La Raza. Hospital de Infectología. Mexico City, Mexico.

[138], Hospital Universitario de Getafe, Servicio de Genética, Madrid, Spain

[139], Ministerio de Salud Ciudad de Buenos Aires, Buenos Aires, Argentina

[140], Hospital Clinico Universitario de Valladolid, Unidad de Apoyo a la Investigación, Valladolid, Spain

[141], Universidad de Valladolid, Departamento de Cirugía, Valladolid, Spain

[142], Secretaria Municipal de Saude de Apodi, Natal, Brazil

[143], Sección Genética Médica - Servicio de Pediatría, Hospital Clínico Universitario Virgen de la Arrixaca, Servicio Murciano de Salud, Murcia, Spain

[144], Departamento Cirugía, Pediatría, Obstetricia y Ginecología, Facultad de Medicina, Universidad de Murcia (UMU), Murcia, Spain

[145], Hospital Universitario Centro Dermatológico Federico Lleras Acosta, Bogotá, Colombia

[146], Hospital Universitario Virgen de las Nieves, Servicio de Análisis Clínicos e Inmunología, Granada, Spain

[147], Pneumology Department, Hospital General Universitario Gregorio Marañón (iiSGM), Madrid, Spain

[148], Intermediate Respiratory Care Unit, Department of Pneumology, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[148], Clinica Comfamiliar Risaralda, Pereira, Colombia

[150], Centro Universitario de Tonalá, Universidad de Guadalajara, Guadalajara, Mexico

[151], Centro de Investigación Multidisciplinario en Salud, Universidad de Guadalajara, Guadalajara, Mexico

[152], Unidad de Cuidados, Intensivos Hospital Clínico Universitario de Santiago (CHUS), Sistema Galego de Saúde (SERGAS), Santiago de Compostela, Spain

[153], IIS La Fe, Plataforma de Farmacogenética, Valencia, Spain

[154], Universidad de Valencia, Departamento de Farmacología, Valencia, Spain

[155], Data Analysis Department, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[156], Hospital del Mar, Infectious Diseases Service, Barcelona, Spain

[157], Institut Hospital del Mar d'Investigacions Mèdiques (IMIM), Barcelona, Spain

[158], CEXS-Universitat Pompeu Fabra, Spanish Network for Research in Infectious Diseases (REIPI), Barcelona, Spain

[159], Biocruces Bizkaia Health Research Institute, Basurto University Hospital, Osakidetza, Bizkaia, Spain

[160], Sabin Medicina Diagnóstica, Brazil

[161], Opthalmology Department, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[162], Hospital Sant Joan de Deu,Pediatric Critical Care Unit, Barcelona, Spain

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[163], Paediatric Intensive Care Unit, Agrupación Hospitalaria Clínic-Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain

[164], Hospital Universitario 12 de Octubre, Department of Immunology, Madrid, Spain

[165], Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas12), Transplant Immunology and Immunodeficiencies Group, Madrid, Spain

[166], SIGEN Alianza Universidad de los Andes - Fundación Santa Fe de Bogotá, Bogotá, Colombia

[167], Hospital General de Segovia, Medicina Intensiva, Segovia, Spain

[168], Programa de Pós-Graduação em Biologia Animal, Universidade de Brasília, Brasília, Brazil

[169], Clinical Trials Unit, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[170], IMDEA-Food Institute, CEI UAM+CSIC, Madrid, Spain

[171], Intensive Care Unit, Hospital Universitario de Canarias, La Laguna, Spain

[172], Dirección General de Salud Pública, Consejería de Sanidad, Junta de Castilla y León, Valladolid, Spain

[173], Universidade Federal do Rio Grande do Norte, Departamento de Analises Clinicas e Toxicologicas, Natal, Brazil

[174], Hospital Universitario La Paz-IDIPAZ, Servicio de Inmunología, Madrid, Spain

[175], La Paz Institute for Health Research (IdiPAZ), Lymphocyte Pathophysiology in Immunodeficiencies Group, Madrid, Spain

[176], Hospital Universitario Virgen de las Nieves, Servicio de Enfermedades Infecciosas, Granada, Spain

[177], Instituto de Investigación Biosanitaria de Granada (ibs GRANADA), Granada, Spain

[178], Universidad de Granada, Departamento de Medicina, Granada, Spain

[179], Faculdade de Medicina, Universidade de Brasília, Brasilia, Brazil

[180], Fundación Jiménez Díaz, Epidemiology, Madrid, Spain

[181], Universidad Autónoma de Madrid, Department of Medicine, Madrid, Spain

[182], Universidad de Valladolid, Departamento de Medicina, Valladolid, Spain

[183], Hospital Universitario Infanta Leonor, Servicio de Medicina Intensiva, Madrid, Spain

[184], Unidad de Genética y Genómica Islas Baleares, Islas Baleares, Spain

[185], Hospital Universitario Son Espases, Unidad de Diagnóstico Molecular y Genética Clínica, Islas Baleares, Spain

[186], Genomics of Complex Diseases Unit, Research Institute of Hospital de la Santa Creu i Sant Pau, IIB Sant Pau, Barcelona, Spain

[187], Intensive Care Unit, Hospital Universitario N. S. de Candelaria, Santa Cruz de Tenerife, Spain

[188], Preventive Medicine Department, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[189], Servicio de Medicina Interna, Sanatorio Franchin, Buenos Aires, Argentina

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[190], Hospital Universitario del Tajo, Servicio de Medicina Intensiva, Toledo, Spain

[191], Programa de Pós-Graduação em Ciências Médicas, Universidade de Brasília, Brasilia, Brazil

[192], Programa de Pós-Graduação em Ciências da Saúde, Universidade de Brasília, Brasilia, Brazil

[193], Hospital El Bierzo, Gerencia de Asistencia Sanitaria del Bierzo (GASBI), Gerencia Regional de Salud (SACYL), Ponferrada, Spain

[194], Grupo INVESTEN, Instituto de Salud Carlos III, Madrid, Spain

[195], Unidad de Cuidados Intensivos, Complejo Universitario de A Coruña (CHUAC), Sistema Galego de Saúde (SERGAS), A Coruña, Spain

[196] Programa de Pós-Graduação em Genética da Faculdade de Medicina de Ribeirão Preto

[197] Programa de Pós-Graduação em Química da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto

[198], Hospital El Bierzo, Unidad Cuidados Intensivos, León, Spain

[199], Spanish National Cancer Research Centre, Familial Cancer Clinical Unit, Madrid, Spain

[200], Instituto de Investigación Sanitaria San Carlos (IdISSC), Hospital Clínico San Carlos (HCSC), Madrid, Spain

[201], Hospital Universitario Severo Ochoa, Servicio de Medicina Interna, Madrid, Spain

[202], Universidad de Sevilla, Departamento de Enfermería, Seville, Spain

[203], Hospital General Universitario Gregorio Marañón (IiSGM), Madrid, Spain

[204], Hospital Universitario La Paz-IDIPAZ, Servicio de Pediatría, Madrid, Spain

[205], ERN-ITHACA-European Reference Network

[206], Unidad de Genética y Genómica Islas Baleares, Unidad de Diagnóstico Molecular y Genética Clínica, Hospital Universitario Son Espases, Islas Baleares, Spain

[207], Instituto de Investigación Sanitaria Islas Baleares (IdISBa), Islas Baleares, Spain

[208], Programa de Pós-Graduação em Biologia Animal, Universidade de Brasília, Brasília, Brazil

[209], Programa de Pós-Graduação em Ciências da Saúde, Universidade de Brasília, Brasília, Brazil

[210], Programa de Pós-Graduação Profissional em Ensino de Biologia, Universidade de Brasília, Brasília, Brazil

[211], Programa de Pós-Graduação em Ciências Médicas, Universidade de Brasília, Brasília, Brazil

[212], Anatomía Patológica, Instituto de Investigación Sanitaria San Carlos (IdISSC), Hospital Clínico San Carlos (HCSC), Madrid, Spain

[213], Tecnológico de Monterrey, Monterrey, Mexico

[214], Infectious Diseases, Microbiota and Metabolism Unit, CSIC Associated Unit, Center for Biomedical Research of La Rioja (CIBIR), Logroño, Spain

[215], Centro de Investigación en Anomalías Congénitas y Enfermedades Raras (CIACER), Universidad Icesi

[216], Departamento de Genetica, Fundación Valle del Lili

[217], Universidad Complutense de Madrid, Department of Immunology, Ophthalmology and ENT, Madrid, Spain

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[218], Department of Neumology, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[219], Hospital Nuestra Señora de Sonsoles, Ávila, Spain

[220], Inditex, A Coruña, Spain

[221], Osakidetza, Cruces University Hospital, Barakaldo, Bizkaia, Spain

[222], GENYCA, Madrid, Spain

[223], Neuromuscular Diseases Unit, Department of Neurology, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona, Barcelona, Spain

[224], Instituto Mexicano del Seguro Social (IMSS), Centro Médico Nacional Siglo XXI, Unidad de Investigación Médica en Enfermedades Infecciosas y Parasitarias, Mexico City, Mexico

[225], Catalan Institution of Research and Advanced Studies (ICREA), Barcelona, Spain

[226], Intensive Care Department, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[227], Hospital Universitario Príncipe de Asturias, Servicio de Microbiología Clínica, Madrid, Spain

[228], Universidad de Alcalá de Henares, Departamento de Biomedicina y Biotecnología, Facultad de Medicina y Ciencias de la Salud, Madrid, Spain

[229], Drug Research Centre, Institut d'Investigació Biomèdica Sant Pau, IIB-Sant Pau, Barcelona, Spain

[230], Departamento de Genetica, Clinica imbanaco

[231], Unidad de Cuidados Intensivos, Hospital Clínico Universitario de Santiago (CHUS), Sistema Galego de Saúde (SERGAS), Santiago de Compostela, Spain

[232], Universidade de Brasília, Brasilia, Brazil

[233], Department of Immunology, Hospital Universitario de Gran Canaria Dr. Negrín, Las Palmas de Gran Canaria, Spain

[234], Department of Clinical Sciences, University Fernando Pessoa Canarias, Las Palmas de Gran Canaria, Spain

[235], University Hospital Germans Trias i Pujol, Pediatrics Department, Badalona, Spain

[236], Department of Pathology, Biobank, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[237], Hospital Universitario Virgen de las Nieves, Servicio de Medicina Interna, Granada, Spain

[238], Fundación Universitaria de Ciencias de la Salud, Grupo de Ciencias Básicas en Salud (CBS), Bogotá, Colombia

[239], Sociedad de Cirugía de Bogotá, Hospital de San José, Bogotá, Colombia

[240], Universidad de Granada, Departamento Bioquímica, Biología Molecular e Inmunología III, Granada, Spain

[241], Hospital Infanta Elena, Allergy Unit, Valdemoro, Madrid, Spain

[242], Faculty of Medicine, Universidad Francisco de Vitoria, Madrid, Spain

[243], Hospital Universitario Infanta Leonor, Madrid, Spain

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[244], Complutense University of Madrid, Madrid, Spain

[245], Gregorio Marañón Health Research Institute (IiSGM), Madrid, Spain

[246] Colégio Militar de Brasília

[247], The John Walton Muscular Dystrophy Research Centre, Newcastle University and Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK.

[248], Neuromuscular Unit, Neuropediatrics Department, Institut de Recerca Sant Joan de Déu, Hospital Sant Joan de Déu, Spain

[249], Fundació Docència i Recerca Mutua Terrassa, Terrassa, Spain

[250], Casa de Saúde São Lucas, Natal, Brazil

[251], Hospital Rio Grande, Rio Grande do Norte, Natal, Brazil

[252], Universidade Federal do Rio Grande do Norte, Pós-graduação em Biotecnologia - Rede de Biotecnologia do Nordeste (Renorbio), Natal, Brazil

[253], Intensive Care Unit, Hospital Universitario de Gran Canaria Dr. Negrín, Las Palmas de Gran Canaria, Spain

[254], Universidad Fernando Pessoa Canarias, Las Palmas de Gran Canaria, Spain

[255], Biobank, Puerta de Hierro-Segovia de Arana Health Research Institute, Madrid, Spain

[256], Universidad Rey Juan Carlos, Madrid, Spain

[257], Reumathology Service, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[258], Hospital Clinico Universitario de Valladolid, Servicio de Anestesiologia y Reanimación, Valladolid, Spain

[259], Hospital Clinico Universitario de Valladolid, Servicio de Hematologia y Hemoterapia, Valladolid, Spain

[260], Hospital Universitario Lauro Wanderley, Brazil

[261], Hospital Universitario Infanta Leonor, Servicio de Medicina Interna, Madrid, Spain

[262], University Hospital of Burgos, Burgos, Spain

[263], Universidad de Sevilla, Seville, Spain

[264], Fundación Santa Fe de Bogota, Instituto de servicios medicos de Emergencia y trauma, Bogotá, Colombia

[265], Universidad de los Andes, Bogotá, Colombia

[266], Quironprevención, A Coruña, Spain

[267], Junta de Castilla y León, Consejería de Sanidad, Valladolid, Spain

[268], Gerencia Atención Primaria de Burgos, Burgos, Spain

[269], Immunogenetics-Histocompatibility group, Servicio de Inmunología, Instituto de Investigación Sanitaria Puerta de Hierro - Segovia de Arana, Madrid, Spain

[270], Hospital del Mar, Department of Infectious Diseases, Barcelona, Spain

[271], IMIM (Hospital del Mar Medical Research Institute, Institut Hospital del Mar d'Investigacions Mediques), Barcelona, Spain

[272], Universitat Autònoma de Barcelona, Department of Medicine, Spain

medRxiv preprint doi: https://doi.org/10.1101/2023.08.11.23293871; this version posted February 8, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[273], Consejería de Sanidad, Comunidad de Madrid, Madrid, Spain

[274], Centro para el Desarrollo de la Investigación Científica, Asunción, Paraguay

[275], Internal Medicine Department, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital - Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

[276], Universidade Federal do Rio Grande do Norte, Programa de Pós Graduação em Nutrição, Natal, Brazil

[277], Preventive Medicine Department, Instituto de Investigacion Sanitaria Galicia Sur, Xerencia de Xestion Integrada de Vigo-Servizo Galego de Saúde, Vigo, Spain

[278], Universidade Federal do Rio Grande do Norte, Departamento de Infectologia, Natal, Brazil

[279], Hospital de Doenças Infecciosas Giselda Trigueiro, Rio Grande do Norte, Natal, Brazil

[280], Genomics Division, Instituto Tecnológico y de Energías Renovables, Santa Cruz de Tenerife, Spain