Neuromuscular disease genetics in underrepresented populations: increasing data diversity

Lindsay A. Wilson,[1,†] William L. Macken,[1,†] Luke D. Perry,[2,3,†] Christopher J. Record,[1] Katherine R. Schon,[4] Rodrigo S. S. Frezatti,[5] Sharika Raga,[6,7] Kireshnee Naidu,[8,9] Özlem Yayıcı Köken,[10] Ipek Polat,[11,12] Musambo M. Kapapa,[13] Natalia Dominik,[1] Stephanie Efthymiou,[1] Heba Morsy,[1] Melissa Nel,[6,8] Mahmoud R. Fassad,[14] Fei Gao,[4] Krutik Patel,[14] Maryke Schoonen,[15] Michelle Bisschoff,[15] Armand Vorster,[15] Hallgeir Jonvik,[1] Ronel Human,[16] Elsa Lubbe,[16] Malebo Nonyane,[16] Seena Vengalil,[17] Saraswati Nashi,[17] Kosha Srivastava,[17] Richard J. L. F. Lemmers,[18] Alisha Reyaz,[19] Rinkle Mishra,[19] Ana Töpf,[20] Christina I. Trainor,[20] Elizabeth C. Steyn,[8] Amokelani C. Mahungu,[6,8] Patrick J. van der Vliet,[18] Ahmet Cevdet Ceylan,[21.22] A Semra Hiz,[11,23] Büşranur Çavdarlı,[21] C. Nur Semerci Gündüz,[21,22] Gülay Güleç Ceylan,[21,22] Madhu Nagappa,[17] Karthik B. Tallapaka,[24] Periyasamy Govindaraj,[25] Silvère M. van der Maarel,[18] Gayathri Narayanappa,[26] Bevinahalli N. Nandeesh,[26] Somwe wa Somwe,[27] David R. Bearden,[28,29] Michelle P. Kvalsund,[29,30] Gita M. Ramdharry,[1] Yavuz Oktay,[12,23] Uluç Yiş,[11] Haluk Topaloğlu,[31] Anna Sarkozy,[3] Enrico Bugiardini,[1] Franclo Henning,[9] Jo M. Wilmshurst,[6,7] Jeannine M. Heckmann,[6,8] Robert McFarland,[14,32] Robert W. Taylor,[14,32] Izelle Smuts,[16] Francois H. van der Westhuizen,[15] Claudia Ferreira da Rosa Sobreira,[5] Pedro J. Tomaselli,[5] Wilson Marques Jr,[5] Rohit Bhatia,[19] Ashwin Dalal,[25] M. V. Padma Srivastava,[19] Sireesha Yareeda,[33] Atchayaram Nalini,[17] Venugopalan Y. Vishnu,[19] Kumarasamy Thangaraj,[24] Volker Straub,[20] Rita Horvath,[4] Patrick F. Chinnery,[4] Robert D. S. Pitceathly,[1] Francesco Muntoni,[2,3] Henry Houlden,[1] Jana Vandrovcova,[1] Mary M. Reilly[1] and Michael G. Hanna[1]

[†]These authors contributed equally to this work.

Abstract

Neuromuscular diseases (NMDs) affect ~15 million people globally. In high income settings DNA-based diagnosis has transformed care pathways and led to gene-specific therapies. However, most affected families are in low-middle income countries (LMICs) with limited access to DNA-based diagnosis. Most (86%) published genetic data is derived from European ancestry. This marked genetic data inequality hampers understanding of genetic diversity and hinders accurate genetic diagnosis in all income settings. We developed a cloud-based transcontinental partnership to build diverse, deeply-phenotyped and genetically characterised

© The Author(s) 2023. Published by Oxford University Press on behalf of the Guarantors of Brain. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

cohorts to improve genetic architecture knowledge, and potentially advance diagnosis and clinical management.

We connected 18 centres in Brazil, India, South Africa, Turkey, Zambia, Netherlands and the UK. We co-developed a cloud-based data solution and trained 17 international neurology fellows in clinical genomic data interpretation. Single gene and whole exome data was analysed via a bespoke bioinformatics pipeline and reviewed alongside clinical and phenotypic data in global webinars to inform genetic outcome decisions.

We recruited 6001 participants in the first 43 months. Initial genetic analyses "solved" or "possibly solved" ~56% probands overall. In-depth genetic data review of the four commonest clinical categories (limb girdle muscular dystrophy, inherited peripheral neuropathies, congenital myopathy/muscular dystrophies and Duchenne/Becker muscular dystrophy) delivered a ~59% "solved and ~13% "possibly solved" outcome. Almost 29% of disease causing variants were novel, increasing diverse pathogenic variant knowledge. Unsolved participants represent a new discovery cohort. The dataset provides a large resource from underrepresented populations for genetic and translational research.

In conclusion, we established a remote transcontinental partnership to assess genetic architecture of NMDs across diverse populations. It supported DNA-based diagnosis potentially enabling genetic counselling, care pathways and eligibility for gene-specific trials. Similar virtual partnerships could be adopted by other areas of global genomic neurological practice, to reduce genetic data inequality and benefit patients globally.

Author affiliations:

1 Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology and The National Hospital for Neurology and Neurosurgery, London WC1N 3BG, UK

2 The Dubowitz Neuromuscular Centre, University College London, UCL Great Ormond Street, Institute of Child Health and Centre for Neuromuscular Diseases, Neurosciences Unit, Great Ormond Street Hospital, London WC1N 3JH, UK

3 NIHR Great Ormond Street Hospital Biomedical Research Centre, UCL Great Ormond Street Institute of Child Health, University College London WC1N 1EH, UK

4 Department of Clinical Neurosciences, University of Cambridge, Cambridge CB2 0QQ, UK

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

5 Department of Neurosciences, School of Medicine of Ribeirão Preto, University of São Paulo, São Paulo, Brazil

6 Neuroscience Institute, University of Cape Town, Cape Town, South Africa

7 Division of Paediatric Neurology, Department of Paediatrics and Child Health, Red Cross War Memorial Children's Hospital, Cape Town, South Africa

8 Neurology Research Group, Division of Neurology, Department of Medicine, University of Cape Town, Cape Town, South Africa

9 Division of Neurology, Department of Medicine, Stellenbosch University, Cape Town, South Africa

10 Akdeniz University, Faculty of Medicine, Department of Pediatric Neurology, Antalya, Turkey

11 Dokuz Eylül University, Faculty of Medicine, Pediatric Neurology Department, Izmir, Turkey

12 Izmir International Biomedicine and Genome Institute, Dokuz Eylül University, Izmir, Turkey

13 University of Zambia School of Health Sciences & University Teaching Hospital Neurology Research Office, Lusaka, Zambia

14 Wellcome Centre for Mitochondrial Research, Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne NE2 4HH, UK

15 Focus Area for Human Metabolomics, North-West University, Potchefstroom, South Africa

16 Department of Paediatrics, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa

17 Department of Neurology, National Institute of Mental Health and Neurosciences (NIMHANS), Bengaluru, India

18 Department of Human Genetics, Leiden University Medical Center (LUMC), Leiden, The Netherlands

19 Department of Neurology, All India Institute of Medical Sciences (AIIMS), Delhi, India

20 John Walton Muscular Dystrophy Research Centre, Newcastle University Translational and Clinical Research Institute and Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

21 Ankara Bilkent City Hospital, Department of Medical Genetics, Ankara, Turkey

22 Ankara Yıldırım Beyazıt University, Faculty of Medicine, Department of Medical Genetics, Ankara, Turkey

23 Izmir Biomedicine and Genome Center (IBG), Izmir, Turkey

24 CSIR - Centre for Cellular and Molecular Biology (CCMB), Hyderabad, Telangana, India

25 Centre for DNA Fingerprinting and Diagnostics, Hyderabad, Telangana, India

26 Department of Neuropathology, National Institute of Mental Health and Neurosciences (NIMHANS), Bengaluru, India

27 Department of Clinical Sciences, School of Medicine and Health Sciences, University of Lusaka, Lusaka, Zambia

28 University of Zambia Department of Educational Psychology, Lusaka, Zambia

29 University of Rochester Medical Center School of Medicine and Dentistry, Rochester, New York, USA

30 29 University of Zambia School of Medicine, Lusaka, Zambia

31 Yeditepe University Hospitals, Istanbul, Turkey

32 NHS Highly Specialised Service for Rare Mitochondrial Disorders, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne NE1 4LP, UK

33 Nizam's Institute of Medical Sciences (NIMS), Hyderabad, Telangana, India

Correspondence to: Professor Michael G. Hanna

Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology and The National Hospital for Neurology and Neurosurgery, London WC1N 3BG, UK

E-mail: m.hanna@ucl.ac.uk

Running title: The ICGNMD Consortium

Keywords: genomic medicine; inherited neuromuscular disease; capacity building; low-to-middle income country; equality and diversity

Abbreviations: ACMG = American College of Medical Genetics; CM/CMD = congenital myopathies/congenital muscular dystrophies; CMT1 = Charcot-Marie-Tooth disease type 1;

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

DSA = Data Sharing Agreement; DMD/BMD = Duchenne/Becker muscular dystrophy; DUO code = Data Use Ontology Code; EMBL EBI EGA = European Molecular Biology Laboratory European Bioinformatics Institute European Genome-Phenome Archive; FSHD = facioscapulohumeral muscular dystrophy; GA4GH = Global Alliance for Genomics and Health; GDPR = UK General Data Protection Regulations; ICGNMD = International Centre for Genomic Medicine in Neuromuscular Diseases; LGMD = Limb Girdle Muscular Dystrophy; LMIC = low-to-middle income country; LP = Likely Pathogenic; MLPA = multiplex ligation-dependent probe amplification; NMD = neuromuscular diseases; PCR = polymerase chain reaction; PN = Genetic Peripheral Neuropathies; SGT = single gene test; SMA = spinal muscular atrophy; VUS = variant of uncertain significance; WES = whole exome sequencing; WGS = whole genome sequencing

Introduction

Neuromuscular diseases (NMD) affect an estimated 15 million children and adults globally.[1] They cause shortened life expectancy or chronic lifelong disability with personal and economic impact. Although individually rare, collectively they account for approximately ~20% of all non-infectious neurological diseases. In low-to-middle income countries (LMICs), NMD prevalence and incidence are under-reported, as diagnosis is limited to a few specialist centres which may be geographically distant from much of the eligible population.

In high-income settings, improved diagnostics, especially genetic analysis, have delivered important advances in patient care pathways. Many interventions enabled by an accurate DNA-based diagnosis are relatively inexpensive, including genetic counselling, tailored application of widely-available medicines, screening for known complications (e.g. cardiac, respiratory, gastroenterological, metabolic) and physiotherapy. Genetic advances have also led to new advanced therapies, for example RNA targeting approaches (for Spinal Muscular Atrophy, SMA) and AAV-mediated gene therapies or trials for SMA and Duchenne's Muscular Dystrophy (DMD).[2,3]

A key challenge to realising DNA-based diagnostic benefits for patients in LMICs is that ~86% of published genomic studies are derived from populations of primarily European ancestry and non-European populations are under-represented in control databases.[4] Knowledge of genetic diversity outside European populations is limited.[5-7] Improved understanding of NMD distribution and associated phenotypic and genetic variability requires a large, diverse, accurately-phenotyped cohort of patients and families with linked genetic data. Additional

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

cohort benefits include generating allele frequency data for under-represented populations to aid variant classification, and potential to connect participants to clinical trials and novel therapies.

Here we describe our approach and genetic results in setting up a NMD genomic medicine partnership across 18 centres spanning seven countries.

Materials and methods

ICGNMD structure

The International Centre for Genomic Medicine in Neuromuscular Diseases (ICGNMD) was launched in June 2019 and is ongoing. Partner sites (**Supplementary Table 1**) established aligned, locally-approved studies to recruit NMD patients and relatives to an international cohort and share materials and data in full compliance with all ethics and legislation (for further details about ethics and data storage and sharing see **Supplementary material**). The inclusion criterion for participants was a suspected inherited neuromuscular disease clinically diagnosed by a trained clinical neurologist, or being a close relative. Participants with a local genetic diagnosis could be included, or unsolved local WES data re-analysed. UK partner sites may also recruit participants (typically with existing genetic test data) living in the UK; however this paper considers only participant data from low-to-middle income partner sites.

The international regulatory and ethics landscape is complex and securing all regulatory and ethical approvals to balance data and material accessibility with patient rights was highly challenging and required the nuanced input of experienced local teams.

Building future genomics medicine capacity was crucial to all partners realising precision medicine benefits. Therefore, 17 fellows were appointed to support recruitment, data entry and results interpretation. Twelve were based in LMICs and all are pursuing their careers locally. Fellows were assigned one LMIC Principal Investigator (PI) and one UK PI for mentorship and capacity building; supervision ran alongside regular remote training and data interpretation and was considered highly effective by both fellows and PIs.

In-person face-to-face study induction training of all neurology fellows focused on standardised data entry to the ICGNMD REDCap database,[8] including Human Phenotype Ontology (HPO) terms,[9] standard clinical assessment scales and summary genetic data (for database instrument, see **Supplementary material**), followed by regular online refresher

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

training. As not all sites could access all investigations (e.g. MRI often unavailable), the only mandatory data entries were: proband or relationship to proband, diagnostic category (provisional clinical diagnosis), sex, age at recruitment and positive and negative HPO terms. Repeat measurements (e.g. blood creatine kinase) and progression indicators could be recorded but the study was primarily cross-sectional, reflecting challenges of re-evaluating participants who may struggle to travel to clinics.

ICGNMD genetic analysis and data report generation

After consent and data collection, international, remote expert group "genetic analysis decision" meetings discussed the most appropriate initial genetic analysis. These meetings, attended by all PIs, also served as a Fellows' training forum. Clinical phenotype and any investigational data underpinned the decision to apply specific single gene tests (e.g. MLPA for SMA) or whole exome sequencing (WES) (see **Supplementary material**), with optional genotyping arrays to detect large structural or copy number variation and/or for linkage analysis. Testing was typically proband first, and extended to relatives if needed and/or available (**Fig. 1**). Partners in Brazil, South Africa and Zambia sent DNA to the UK, however partners in India and Turkey generated pseudonymised raw data to agreed standards and shared this for centralised analysis.

Single gene test results (SGTs) were reviewed by trained staff at ICGNMD partner sites. Raw whole exome sequencing (WES) data were analysed via a common bioinformatics pipeline (see **Supplementary material**), with ICGNMD Fellows supported to interpret and present genetic results. WES candidate variant prioritisation followed a modified protocol developed by the 100,000 Genomes Project.[10] Initial analysis focused on variants with a minor allele frequency of of <0.01 (autosomal recessive inheritance) and <0.001 (autosomal dominant inheritance) within a subset of genes present in expertly reviewed gene panels (https://panelapp.genomicsengland.co.uk/) aligning to the participant HPO terms and phenotype. If no significant genetic variants were reported, extended pipeline analyses (structural and copy number, mitochondrial, repeat expansion, *de novo*) were applied. To maximise reproducibility, open-source and well maintained software tools and databases were implemented.[11,12] Given their relatively limited non-European ancestry data and lack of subpopulation resolution, e.g. for specific Indian and African ethnicities, we supplemented large-scale population data resources including gnomAD[13] with additional allele frequency data generated in local populations [14-18] and the growing ICGNMD in-house dataset.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

Prioritised variants were reviewed and potentially causative variants were classified using ACMG criteria.[19] The outcome was classed as 'solved' where pathogenic/likely pathogenic variant(s) were identified and fitted with the phenotype (two variants/homozygous in recessive disorders). The outcome was classed as 'possibly solved' if there was a strong candidate variant (two variants/homozygous recessive disorders) based on population frequency (<0.01% frequency), bioinformatic predictions and clinical phenotype, but at least one variant was classified as a Variant of Uncertain Significance (VUS) according to ACMG criteria. Where further manual curation was performed for subgroups of disease categories, rare variants considered relevant for each proband were reviewed against gnomAD, Clinvar, Decipher, Varsome, Pubmed and Google to ascertain if previously reported. Variants were classified as "novel" if absent from all these sources.

Results

Despite the SARS-CoV-2 pandemic we established a phenotyping, genetic analysis and data sharing platform connecting centres across Brazil, India, Netherlands, South Africa, Turkey, UK and Zambia. We developed remote training and global webinars to discuss participants and supported decision-making for genetic analysis and result interpretation. As of January 2023, 6001 participants (including 3631 probands) had consented and provided DNA (**Supplementary Table 2**). The majority were in India (3578 participants, 60% of the total), followed by Brazil (979 participants, 16%), South Africa (737 participants, 12%), Turkey (578 participants, 10%) and Zambia (129 participants, 2%). The cohort included 337 (9% probands) participants "locally genetically solved" or with existing genetic data to review at study start. The majority (3294, 91%) of participants had no previous genetic test.

Sixty five percent of participants were male, 35% female. The median age at proband recruitment was 26 years of age with 35% of the cohort aged 18 or under (**Supplementary Fig. 1**). Using 1000 Genomes populations as a background for ancestry estimation, 82% percent of individuals tested by WES were of non-European ancestry. (**Supplementary Fig. 2**). Based on current recruitment, we estimate cohort size will exceed 10,000 at year 5 end; June 2024.

Phenotypic spectrum

We recruited a broad range of NMDs (**Fig. 2** and **Supplementary Table 3**). Total recruitment to mid-January 2023 by initial clinical diagnosis included 18.1% of limb girdle muscular dystrophy (LGMD), 15.5% genetic peripheral neuropathies (PN), 9.4% congenital muscular

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

dystrophy or myopathy (CM/CMD) and 8.6% Duchenne muscular dystrophy (DMD) or Becker muscular dystrophy (BMD). Other categories each contributed less than 7%. The four most common NMD categories were in line with those reported by centres worldwide.[1]

Genetic data results

We report 978 new genetic analyses (including 547 proband WES and 274 proband SGTs) by January 2023, following the process in **Fig. 1**. 50% of probands receiving a SGT were solved (**Supplementary Table 4**). The first proband WES data review identified 223 (41%) variants to "solve" and 83 (15%) variants to "possibly solve" participants' NMDs (combined WES solved/possibly solved rate 56%) . Single gene tests and WES combined yielded 43.8% "solved" outcomes. Below are genetic summary data following in-depth review of proband single gene test and WES results for the four clinical diagnostic categories with highest recruitment levels (PN, LGMD, CMD/CM and DMD/BMD) to demonstrate project value at Year 4 stage. These categories combined represent 1875 of 3631 study probands (51.6% of cohort) and 340 of 547 (62%) exomes available in January 2023, plus 182 single gene tests.

Genetic peripheral neuropathies

One hundred and eighty one participants with a presumed genetic peripheral neuropathy were analysed (114 Brazilian, 16 Indian, 51 South African). Single gene tests in 80 probands had a diagnostic yield of 54% (43/80) (*GJB1* (26 participants; 32.5%), *PMP22* duplication (14 participants; 17.5%), *FXN* (2 participants; 2.5%), *PMP22* deletion for HNPP genetic diagnosis (1 participant; 1.25%)). All but seven probands with a negative SGT had WES afterwards. WES in 131 probands yielded 40 (31%) solved, 28 (21%) possibly solved and 63 (48%) unsolved outcomes. Combined diagnostic yield (Solved) of SGT and WES was 46% (83/181) with 15% (28/181) classified as possibly solved (**Fig. 3A**). Combined solved outcomes involved twenty seven genes (**Fig. 3B**) with *GJB1*, *PMP22* duplication and *MFN2* most common across 29 (16%), 14 (8%) and 4 (2%) participants respectively. Most common WES-identified variants were in genes *MFN2* (4 participants, 3%), *GJB1* (3 participants, 2.3%), *MPZ* (3 participants, 2.3%), *PRX* (3 participants, 2.3%), and *SH3TC2* (3 participants, 2.3%). Twenty-two novel genetic variants (13 pathogenic/likely pathogenic, 9 VUS) were identified in 19 genes (**Table 1**).

Congenital myopathies and muscular dystrophies

One hundred and seven probands underwent WES (72 South African, 22 Turkish, 10 Brazilian, 3 Indian) with 57 (53%) probands solved (16 genes), 20 (19%) possibly solved (18 genes), and

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

30 (28%) unsolved (**Fig. 3A**). Diagnostic yield ("solved/possibly solved" outcomes) varied between 0-60% across populations. *STAC3* (28), *RYR1* (8) and *COL6A2/3* (5) were the most common genes in solved probands (**Fig. 3B**). Twenty-one novel genetic variants (12 likely pathogenic, 9 VUS) were identified in 14 genes (**Table 2**).

Limb girdle muscular dystrophies

Eighty-five probands underwent WES (47 Indian, 13 Turkish, 11 Brazilian, 12 South African, 2 Zambian) with 37 (44%) patients solved (16 genes), 14 (16%) possibly solved (12 genes) and 34 (40%) unsolved (**Fig. 3A**). *DYSF* (10), *CAPN3* (9), and *GNE* (3) were the most common genes in solved probands (**Fig. 3B**). Fifteen novel genetic variants (7 pathogenic/likely pathogenic, 8 VUS) were identified in ten genes (**Table 3**).

Duchenne and Becker muscular dystrophy

Multiplex ligation probe amplification (MLPA) test for dystrophin (*DMD*) gene deletions and duplications is relatively accessible in partner sites with the exception of Zambia; enabling strong "already solved" ICGNMD participant recruitment for DMD and BMD to combine with new data. DMD MLPA data for 102 probands (94%) was reviewed, plus 17 ICGNMD exomes across a total of 108 probands (64 Indian, 41 South African, 3 Brazilian). One hundred and six (98%) patients were solved (81 DMD, 23 BMD, 2 symptomatic female carriers) and 2 (2%) unsolved after both MLPA and WES. Diagnosis came from *DMD* MLPA in 91 (86%) and WES in 15 (14%) solved outcomes (**Fig. 3A**). In the solved cohort, 85 (80%) patients had a deletion, 12 (11%) had a nonsense variant (point mutation), 7 (7%) had a duplication and 2 (2%) had a splice variant (**Fig. 4**). All BMD patients (23/23) had in frame deletions/duplications. Of the 69 DMD results involving deletions/duplications, 59 (86%) were frameshifts, eight (11%) were in frame and two (3%) involved the initial or terminal *DMD* exon. Of the two symptomatic female carriers one harboured a nonsense variant and the other an out of frame deletion. The prevalence of different types of *DMD* variant varied significantly (Fisher exact test ($P<0.001$)) between Indian and South African cohorts. Of 64 solved Indian participants, 60 (94%) carried a deletion, three (5%) a nonsense variant and one (1%) a splice variant. No Indian patients had a duplication. Of the 40 solved South African participants, 24 (60%) carried a deletion, eight (20%) a nonsense variant, seven (17%) a duplication and one a splice variant in *DMD* (**Fig. 4**). Two solved Brazilian outcomes were due to a *DMD* deletion and a nonsense variant. Analysis of *DMD* breakpoints for the entire cohort revealed two hotspots for deletions and duplications around exons 45-50 and 10-20 (**Fig. 5A**). There were recurrent deletion/duplication breakpoints in intron 45 (29/92; 32%) 46 (10/92; 11%)) and 48

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

(15/92; 16%); with 59% of deletions/duplications having one breakpoint within introns 45-48. Duplications were spread across the gene with 2/7 occurring at exon 2 (one spanning exons 2-7), 2/7 at exon 18 and 1/7 each at exons 17, 45 and 55. Nonsense variants were spread across the gene between exons 4 and 72 (**Fig. 5A**). There was a notable difference in proportion of breakpoints affecting different areas of *DMD* between Indian and South African cohorts. Of 54 deletions/duplications involving exon 45-48 breakpoints, 40/54 (74%) were from the Indian cohort, compared to 13/54 (24%) from South Africa. There was also a comparatively higher proportion of deletions/duplications with breakpoints in the proximal 5' end of the gene (before exon 10) in the South African cohort 9/31 (29%) (6 involving exons 1-3), compared to the Indian cohort 3/60 (5%) (**Fig. 5B**). WES identified two novel pathogenic *DMD* variants (one splice variant and one frameshift deletion). Of the 85 patients with deletions, 44 (52%; 33 Indian and 11 South African) are potentially amenable to antisense oligonucleotide (ASO) exon skipping therapies. Twenty nine patients (22 Indian, 7 South African) carry deletions amenable to licenced exon skipping ASO therapies targeting exons 45 (9), 51 (14) and 53 (5); a single patient carried an exon 52 deletion amenable to either exon 51 or 53 skipping. Of the remaining 15 patients with deletions amenable to exon skipping, six (40%) would be amenable to skipping of exon 44, an ASO currently in clinical trials.

Discussion

Recent years have seen dramatic advances in gene discovery and genetic diagnosis in NMD. Resulting patient benefits include accurate diagnosis, genetic counselling, improved care pathways including complication screening and prevention, and potential access to clinical trials and disease modifying genetic therapies. However, these benefits and impacts have so far been limited or non-existent in lower-income contexts, despite most NMD patients living in LMICs. Here we explored a principally remote method of connecting academic partners and building international capacity for cohort development and genetic analysis of NMD patients in LMICs. We harnessed key features of genetic analysis; i.e. that samples can be collected remotely and shipped for DNA extraction at relatively low-cost, enabling inclusion of geographically-dispersed patient populations and economies of scale. We took advantage of recent computational tools and high-performance clusters to enable efficient, remote processing of sequence data. We used low-cost cloud-based databases to support rapid and secure sharing of phenotypic data between distant sites. Specifically, we 1) tested feasibility of a distributed transcontinental genomic medicine partnership to build diverse, deeply-

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

phenotyped and genetically characterised cohorts and 2) evaluated deploying this partnership and cohort to understand genetic architecture and advance diagnosis. We report 3,600 probands and 6,001 participants who together represent the first recruits to a new global cohort including previously underrepresented populations.

The network of trained fellows working with local PIs and PIs in the UK assembled a deeply phenotyped cohort of children and adults with NMDs. Over 3,600 probands' detailed clinical phenotype and medical histories were recorded in the REDCap database after 43 months (June 2019 to January 2023), and over 2,300 affected and unaffected (mainly first degree) family members. The male:female proband ratio (mean 1.86) is higher than reported by other NMD registries.[20] This may be influenced by referral patterns to some recruitment sites and/or socio-economic circumstances differentially impacting ability to attend appointments.[21]

The data indicates patients with a wide spectrum of neuromuscular diseases joined the ICGNMD study, with a frequency broadly similar to reports from European centres.[1] The most common NMD clinical diagnoses were genetic peripheral neuropathies, limb girdle muscular dystrophy, congenital muscular dystrophy or myopathy and Duchenne or Becker muscular dystrophy, together comprising over half the cohort.

Review of over 820 first proband results (single gene tests and exomes), demonstrated a "solved" rate of approximately 44% and "possibly solved" rate of approximately 15%, (i.e. 59% solved/possibly solved). Solved rate increased to over 58% for the four most common disease categories reviewed in depth (LGMD 44%, PN 46%, CM/CMD 53%, DMD/BMD 98%, increasing to LGMD 60%, PN 61% and CM/CMD 72% (no change to DMD/BMD) when "possibly solved" numbers are included.

The 44% of limb girdle muscular dystrophies solved by WES included *CAPN3*, *DYSF* and *GNE* variants found at frequencies similar to previously reported,[29] alongside 16 novel variants. Overall, 1 in 3.5 LGMD WES variants considered disease-causing disease are novel.

Genetic peripheral neuropathies' solved rate of 46% included single gene tests for common genetic causes. The most common causative genes identified in the genetic peripheral neuropathy cohort (*PMP22* duplication and *GJB1* variants) are similar to those described in previous European and US studies.[22-24] There were more probands with *GJB1* variants (29 participants, 16%) than the *PMP22* duplication (14 participants, 8%) in our cohort whereas other cohorts describe higher rate of *PMP22* duplication participants, usually greater than 50% (e.g. US/UK/European study 61% *PMP22* duplication; 10% *GJB1* variants).[24] A possible cause

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

of this is that 114/181 probands in the current study were from Brazil where the *PMP22* duplication was excluded in many enrolled participants.

Congenital myopathies and muscular dystrophies yielded a comparable "solved" rate of 53%, spanning 16 genes, including known variants in *STAC3*, *RYR1* and *COL6A2/3* in addition to 21 novel variants across 14 genes. There is a notable lack of solved patients with *COL6A1* and *TTN* related CM/CMD (two of the most common forms)[25-28] and further WES is underway.

The dystrophin gene diagnostic rate was high (98%), with MLPA and WES contributing 86% and 14% of diagnoses respectively. Deletions and duplications were most common (87% solved participants), in keeping with previous studies.[30,31] Comparisons between Indian and South African cohorts identified differences in both type of genetic variants (deletions, duplications, nonsense, splice variants) and their distribution (including intronic breakpoints) within the DMD gene. The South African cohort demonstrated a higher number of duplications and nonsense mutations and a higher proportion of intronic breakpoints in the proximal 5' end of the gene. This could be due to differences in cohort size and variation in patient recruitment (e.g. locally solved via MLPA vs unsolved patients) however a comparatively low proportion of large deletions in South African populations is reported.[32] It will be important to interrogate reasons for this observation as there may be implications for applicability of exon-skipping therapies.

Overall, our data indicates that, depending on NMD diagnostic category, 44-98% of patients in LMIC settings may receive an accurate genetic diagnosis with single gene tests and/or whole exome sequencing, creating potential for benefit. We increased the reported genetic diversity associated with NMD since 1 in 3.5 mutations were novel variants. These data also indicate an additional 15% of probands with a strong VUS candidate and "possibly solved" classification require further evaluation/functional studies to confirm pathogenicity with corollary benefits for discovery research and pharmaceutical insights. The 28% of probands for whom no convincing variants were identified represent an important new diverse discovery cohort for further analysis including whole genome and long read approaches.

The ICGNMD team and results depended on international collaborations established at a smaller scale over the preceding decade, building trust and mutual understanding of local populations, facilities and perspectives. Such collaborations benefited from local computing and data- and material storage infrastructure, small-scale pump-priming, initiatives promoting clinical and genetic data interoperability, and gene-matching platforms. The partnership has potential for additional bidirectional benefit, including enabling deeper understanding of VUS,

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

which is relevant to NMD patients in all countries, and the partnership is a foundation on which to build expanded testing capacity, tailored care guidelines and clinical trial readiness.

In conclusion, we demonstrated that it is feasible to set up a virtual transcontinental partnership using a cloud-based platform and harness big data to describe phenotypes and causative variants present in a diverse cohort recruited from many LMIC settings. Over half of tested participants obtained a research-based genetic diagnosis, opening up potential benefits to patients of an accurate DNA diagnosis and demonstrating feasibility of including more diverse populations in clinical trials. We recognise there are limitations to this study. We did not seek to collect epidemiological data and our study does not allow conclusions about incidence or prevalence of NMD or genes. On the other hand this study is in a "real-world" setting reflecting current practice in each LMIC centre and shows that despite limitations, a cross-NMD solved rate of 44% to 59% can be achieved in this previously genetically untested population. This work indicates that geographical inequalities of access to an accurate DNA based diagnosis can potentially be addressed through such virtual partnerships. These have genuine bidirectional value to all partners and the wider research community. Increasing the knowledge of genetic diversity can improve reliable variant interpretation and therefore the accuracy of genetic diagnosis for patients in both low and high income settings.

Data availability

At the end of the study, participants de-identified exome and genome data will be archived in the EMBL European Bioinformatics Institute's European Genome-Phenome Archive (EMBL EBI EGA), with community access to this and selected de-identified REDCap data managed via an ICGNMD Data Access Committee.

Acknowledgements

We wish to acknowledge the kind support of Dr AK Meena (ICGNMD Lead Investigator for NIMS, Hyderabad) who sadly passed way in 2020. We thank Christine Oldfield and Zoe Scott for administrative support.

Funding

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

Funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This work was supported by an MRC strategic award to establish an International Centre for Genomic Medicine in Neuromuscular Diseases (ICGNMD) MR/S005021/1. Additional ICGNMD support including travel and subsistence costs from The National Brain Appeal (UK Charity 290173), UCL Global Engagement. Fellowships for RSSF & KN were funded by The Guarantors of Brain (UK Charity 1197319). The authors acknowledge and are grateful to the following: RSSF, SR, KN, OYK, PJT, VVY conference bursaries from World Muscle Society. SVDM and RL are members of the European Reference Network for Rare Neuromuscular Diseases (ERN EURO-MND). MPK: National Institute of Neurologic Disorders & Stroke (1K23NS112463), American Academy of Neuromuscular & Electrodiagnostic Medicine Development Award and Allen Foundation, Inc. DB: National Institute of Neurologic Disorders & Stroke (K23NS117310) and support from Biogen for KCTN1 Natural History Study. GMR: UCL/UCLH Biomedical Research Centre funding, Health Education England, and UCLH NHS Trust Innovation Fund. RMF, RWT and KP: Wellcome core support (203105/Z/16/Z). RMF additional support from The Lily Foundation and Leigh Syndrome International Consortium. AT: EU Horizon 2020 research and innovation Solve-RD project, No. 779257. FHW, MS, MB, and AV: South African Medical Research Council award "The genetics of Neuromuscular Diseases in South African patient populations: the ICGNMD study". KT is funded by JC Bose Fellowship, SERB Department of Science and Technology, India. PG is supported by the Centre for DNA Fingerprinting and Diagnostics (CDFD)-Core Research Grant, DBT, Government of India. RH: Wellcome award 109915/Z/15/Z, UK Medical Research Council award MR/N025431/1, The Lily Foundation, Evelyn Trust Research Grant (Ref 19/14), Action for AT and UKRI Newton Fund (MR/NO27302/1). PFC: Wellcome awards 212219/Z/18/Z and 224486/Z/21/Z, UK Medical Research Council awards MC_PC_21046, MR/S035699/1 and MR/S01165X/1, LifeArc Philanthropic Fund, NIHR Translational BioResource for Common and Rare Diseases, Alzheimer's Society. BioResource - Genes & Cognition and Leverhulme Trust. RDSP: UK Medical Research Council MR/S002065/1 and MC_PC_21046 and The Lily Foundation. HH: UK Medical Research Council, Wellcome, NIHR University College London Hospitals Biomedical Research Centre (NIHR-BRC), RosetreesTrust, and SOLVE-RD. MMR: Wellcome Grant G104817, National Institutes of Neurological Diseases and Stroke and office of Rare Diseases grants U54NS065712 and 1UOINS109403-01) and Muscular Dystrophy Association grant MDA510281.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Deenen JC, Horlings CG, Verschuuren JJ, Verbeek AL, van Engelen BG. The Epidemiology of Neuromuscular Disorders: A Comprehensive Overview of the Literature. *J Neuromuscul Dis* 2015; 2(1): 73-85.

2. Mercuri E, Pera MC, Scoto M, Finkel R, Muntoni F. Spinal muscular atrophy — insights and challenges in the treatment era. *Nature Reviews Neurology* 2020; 16(12): 706-15.

3. Verhaart IEC, Aartsma-Rus A. Therapeutic developments for Duchenne muscular dystrophy. *Nature Reviews Neurology* 2019; 15(7): 373-86.

4. Sirugo G, Williams SM, Tishkoff SA. The Missing Diversity in Human Genetic Studies. *Cell* 2019; 177(4): 1080.

5. Pereira L, Mutesa L, Tindana P, Ramsay M. African genetic diversity and adaptation inform a precision medicine agenda. *Nat Rev Genet* 2021; 22(5): 284-306.

6. Raga SV, Wilmshurst JM, Smuts I, et al. A case for genomic medicine in South African paediatric patients with neuromuscular disease. *Front Pediatr* 2022; 10: 1033299.

7. D'Angelo CS, Hermes A, McMaster CR, et al. Barriers and Considerations for Diagnosing Rare Diseases in Indigenous Populations. *Front Pediatr* 2020; 8: 579924.

8. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform* 2019; 95: 103208.

9. Kohler S, Carmody L, Vasilevsky N, et al. Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. *Nucleic Acids Res* 2019; 47(D1): D1018-D27.

10. Smedley D, Smith KR, Martin A, et al. 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care - Preliminary Report. *N Engl J Med* 2021; 385(20): 1868-80.

11. Van der Auwera GA, O'Connor BD. Genomics in the Cloud: Using Docker, GATK, and WDL in Terra: O'Reilly Media, Incorporated; 2020.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

12. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biology* 2016; 17(1): 122.

13. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 2020; 581(7809): 434-43.

14. Naslavsky MS, Scliar MO, Yamamoto GL, et al. Whole-genome sequencing of 1,171 elderly admixed individuals from the largest Latin American metropolis (São Paulo, Brazil). *bioRxiv* 2020.

15. Choudhury A, Ramsay M, Hazelhurst S, et al. Whole-genome sequencing for an enhanced understanding of genetic variation among South Africans. *Nature Communications* 2017; 8(1): 2062.

16. Kausthubham N, Shukla A, Gupta N, et al. A data set of variants derived from 1455 clinical and research exomes is efficient in variant prioritization for early-onset monogenic disorders in Indians. *Human Mutation* 2021; 42(4): e15-e61.

17. Jain A, Bhoyar RC, Pandhare K, et al. IndiGenomes: a comprehensive resource of genetic variants from over 1000 Indian genomes. *Nucleic Acids Res* 2021; 49(D1): D1225-D32.

18. Naslavsky MS, Scliar MO, Yamamoto GL, et al. Whole-genome sequencing of 1,171 elderly admixed individuals from Sao Paulo, Brazil. *Nat Commun* 2022; 13(1): 1004.

19. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015; 17(5): 405-24.

20. van Engelen BG, van Veenendaal H, van Doorn PA, et al. The Dutch neuromuscular database CRAMP (Computer Registry of All Myopathies and Polyneuropathies): development and preliminary data. *Neuromuscul Disord* 2007; 17(1): 33-7.

21. Kapoor M, Agrawal D, Ravi S, Roy A, Subramanian SV, Guleria R. Missing female patients: an observational analysis of sex ratio among outpatients in a referral tertiary care public hospital in India. *BMJ Open* 2019; 9(8): e026850.

22. Gess B, Schirmacher A, Boentert M, Young P. Charcot-Marie-Tooth disease: frequency of genetic subtypes in a German neuromuscular center population. *Neuromuscul Disord* 2013; 23(8): 647-51.

23. Cortese A, Wilcox JE, Polke JM, et al. Targeted next-generation sequencing panels in the diagnosis of Charcot-Marie-Tooth disease. *Neurology* 2020; 94(1): e51-e61.

24. Fridman V, Bundy B, Reilly MM, et al. CMT subtypes and disease burden in patients enrolled in the Inherited Neuropathies Consortium natural history study: a cross-sectional analysis. *J Neurol Neurosurg Psychiatry* 2015; 86(8): 873-8.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

25. Allamand V, Brinas L, Richard P, Stojkovic T, Quijano-Roy S, Bonne G. ColVI myopathies: where do we stand, where do we go? *Skelet Muscle* 2011; 1: 30.

26. Fan Y, Liu A, Wei C, et al. Genetic and clinical findings in a Chinese cohort of patients with collagen VI-related myopathies. *Clin Genet* 2018; 93(6): 1159-71.

27. Clement EM, Feng L, Mein R, et al. Relative frequency of congenital muscular dystrophy subtypes: analysis of the UK diagnostic service 2001-2008. *Neuromuscul Disord* 2012; 22(6): 522-7.

28. Natera-de Benito D, Ortez C, Jou C, et al. The Phenotype and Genotype of Congenital Myopathies Based on a Large Pediatric Cohort. *Pediatr Neurol* 2021; 115: 50-65.

29. Topf A, Johnson K, Bates A, et al. Sequential targeted exome sequencing of 1001 patients affected by unexplained limb-girdle weakness. *Genet Med* 2020; 22(9): 1478-88.

30. Viggiano E, Picillo E, Passamano L, et al. Spectrum of Genetic Variants in the Dystrophin Gene: A Single Centre Retrospective Analysis of 750 Duchenne and Becker Patients from Southern Italy. *Genes (Basel)* 2023; 14(1).

31. Selvatici R, Rossi R, Fortunato F, et al. Ethnicity-related DMD Genotype Landscapes in European and Non-European Countries. *Neurol Genet* 2021; 7(1): e536.

32. Esterhuizen AI, Greenberg LJ, Ballo R, Goliath RG, Wilmshurst JM. Duchenne muscular dystrophy in the Western Cape, South Africa: Where do we come from and where are we going? *S Afr Med J* 2016; 106(6 Suppl 1): S67-71.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

Figure legends

Figure 1 ICGNMD workflow with key nodes for international discussion at genetic analysis decision and results review. ICGNMD Fellows' training spans this pathway. MLPA: multiplex ligation dependent probe analysis; WES: whole exome sequencing; WGS: whole genome sequencing; SMA: spinal muscular atrophy, DMD: Duchenne muscular dystrophy; FSHD: facioscapulohumeral muscular dystrophy.

Figure 2 Recruitment across clinical diagnostic categories in REDCap by country of recruitment. Blue = Brazil; Red = South Africa; Grey = Turkey; Yellow = India; Light Blue = Zambia. Abbreviations: MD muscular dystrophy, DMD Duchenne Muscular Dystrophy, FSHD Facioscapulohumeral Muscular Dystrophy.

Figure 3 Genetic analysis outcomes. (A) Outcome in PN, LGMD, CM/CMD and DMD/BMD cohorts, **(B)** Genetic composition of (**i**) genetic peripheral neuropathies (PN); (**ii**) limb girdle muscular dystrophy (LGMD) and (**iii**) congenital myopathy/congenital muscular dystrophy (CM/CMD) "solved" cohorts showing the number of probands with variants detected in each named gene. *PMP22* gene deletion (*PMP22* del), *PMP22 duplication (PMP22* dup), *PMP22* point mutation (*PMP22* pm).

Figure 4 Variants detected in DMD participants with "solved" outcomes. Blue = deletion; grey = duplication; orange = splice; red = nonsense.

Figure 5 Distribution of DMD variants. (A) Exon distribution of causative *DMD* genetic variants for whole cohort; **(B)** Exon distribution of causative *DMD* genetic variants in Indian and South African cohorts. Blue = deletion; grey = duplication; orange = splice; red = nonsense.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

Table 1 Novel genetic variants identified in the Genetic Peripheral Neuropathy (Neuropathy) cohort

Disease category	Gene Symbol	Variant	ACMG Classification
Neuropathy	*SBF2*	ENST00000256190.13:c.2100+1G>A	Pathogenic
Neuropathy	*NEFL*	ENST00000610854.2:c.796G>T	Pathogenic
Neuropathy	*HSPB1*	ENST00000248553.7:c.504del	LP
Neuropathy	*SH3TC2*	ENST00000504517.5:c.321G>A	LP
Neuropathy	*IGHMBP2*	ENST00000255078.8:c.449+2T>A	LP
Neuropathy	*GAN*	ENST00000648994.2:c.280C>T	LP
Neuropathy	*MPZ*	ENST00000463290.5:c.620_623dup	LP
Neuropathy	*PMP22*	ENST00000312280.9:c.448_449delGGinsTT	LP
Neuropathy	*NEFL*	ENST00000610854.2:c.65_68delCCCGinsT	LP
Neuropathy	*VRK1*	ENST00000216639.8:c.1012A>T	LP
Neuropathy	*MPV17*	ENST00000233545.6:c.176_181del	LP
Neuropathy	*PRX*	ENST00000324001.8:c.1560_1562del	LP
Neuropathy	*PMP2*	ENST00000256103.3:c.19G>C	LP
Neuropathy	*MPZ*	ENST00000463290.5:c.212A>G	VUS
Neuropathy	*MPZ*	ENST00000672602.2:c.772C>G	VUS
Neuropathy	*ATP7A*	ENST00000341514.11:c.2083C>T	VUS
Neuropathy	*SCN11A*	ENST00000302328.9:c.1101G>T	VUS
Neuropathy	*AIFM1*	ENST00000287295.8:c.512T>C	VUS
Neuropathy	*ATL1*	ENST00000358385.12:c.1208G>C	VUS
Neuropathy	*KIF1A*	ENST00000648047.1:c.368A>G	VUS
Neuropathy	*KMT2C*	ENST00000262189.11:c.1013C>T	VUS
Neuropathy	*MPZ*	ENST00000672602.2:c.772C>G	VUS

Table 2 Novel genetic variants identified in the Congenital Myopathy/Congenital Muscular Dystrophy (CM/CMD) cohorts

Disease category	Gene Symbol	Variant	ACMG Classification
CM/CMD	*NEB*	ENST00000397345.8:c.17502_17510dup	LP
CM/CMD	*LAMA2*	ENST00000421865.3:c.4127T>A	LP
CM/CMD	*RYR1*	ENST00000355481.8:c.6175_6187del	LP
CM/CMD	*MSTO1*	ENST00000245564.2:c.1678C>T	LP
CM/CMD	*PIEZ02*	ENST00000302079.10:c.1345C>T	LP
CM/CMD	*PIEZ02*	ENST00000302079.10:c.5082+2T>C	LP
CM/CMD	*CHCHD10*	ENST00000484558.2:c.262-1_262dup	LP
CM/CMD	*MMF*	ENST00000304593.14:c.744+1G>A	LP
CM/CMD	*NEB*	ENST00000397337.6:c.736dup	LP
CM/CMD	*NEB*	ENST00000397345.8:c.23310del	LP
CM/CMD	*LAMA2*	ENST00000421865.3:c.1170C>A	LP
CM/CMD	*TPM3*	ENST00000271850.11:c.734G>C	LP
CM/CMD	*RYR1*	ENST00000355481.8:c.12323G>A	VUS
CM/CMD	*BICD2*	ENST00000356884.11:c.1993_1998dup	VUS
CM/CMD	*ACTA1*	ENST00000366683.3:c.182A>G	VUS
CM/CMD	*MYH2*	ENST00000245503.10:c.4809G>A	VUS
CM/CMD	*GBE1*	ENST00000429644.7:c.602A>G	VUS
CM/CMD	*RYR1*	ENST00000355481.8:c.9678G>T	VUS
CM/CMD	*MSTO1*	ENST00000245564.8:c.49G>C	VUS
CM/CMD	*FKRP*	ENST00000318584.10:c.1034G>T	VUS
CM/CMD	*PLOD1*	ENST00000196061.5:c.1285G>C	VUS

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

Table 3 Novel genetic variants identified in the Limb-Girdle Muscular Dystrophy (LGMD) and Duchenne/Becker Muscular Dystrophy (DMD/BMD) cohorts

Disease category	Gene Symbol	Variant	ACMG Classification
LGMD	*GNE*	ENST00000396594.8:c.1057C>T	Pathogenic
LGMD	*DYSF*	ENST00000258104.7:c.4558del	Pathogenic
LGMD	*DYSF*	ENST00000258104.7:c.3496_3508del	LP
LGMD	*GNE*	ENST00000396594.8:c.2196G>C	LP
LGMD	*GNE*	ENST00000396594.8:c.1000dup	LP
LGMD	*CAV3*	ENST00000343849.3:c.262T>G	LP
LGMD	*DYSF*	ENST00000258104.7:c.856-1G>A	LP
LGMD	*HSPG2*[a]	ENST00000374676.4:c.14C>T[a]	VUS
LGMD	*SYNE2*	ENST00000344113.8:c.18212G>A	VUS
LGMD	*DYSF*	ENST00000258104.7:c.1781T>C	VUS
LGMD	*DYSF*	ENST00000258104.7:c.5388dup	VUS
LGMD	*KIF5A*	ENST00000286452.5:c.839G>T	VUS
LGMD	*MYH3*	ENST00000583535.6:c.3131A>T	VUS
LGMD	*RYR1*	ENST00000355481.8: c.2321 G>A	VUS
LGMD	*DMD*	ENST00000343523.7:c.1859A>T	VUS
DMD/BMD	*DMD*	ENST00000357033.9:c.2381-1G>C	LP
DMD/BMD	*DMD*	ENST00000357033.9:c.4575_4579del	LP

LP = likely pathogenic, VUS = variant of uncertain significance, ACMG = American College of Medical Genetics.
[a]The *HSPG2* c.14C>T variant was identified in two unrelated participants.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

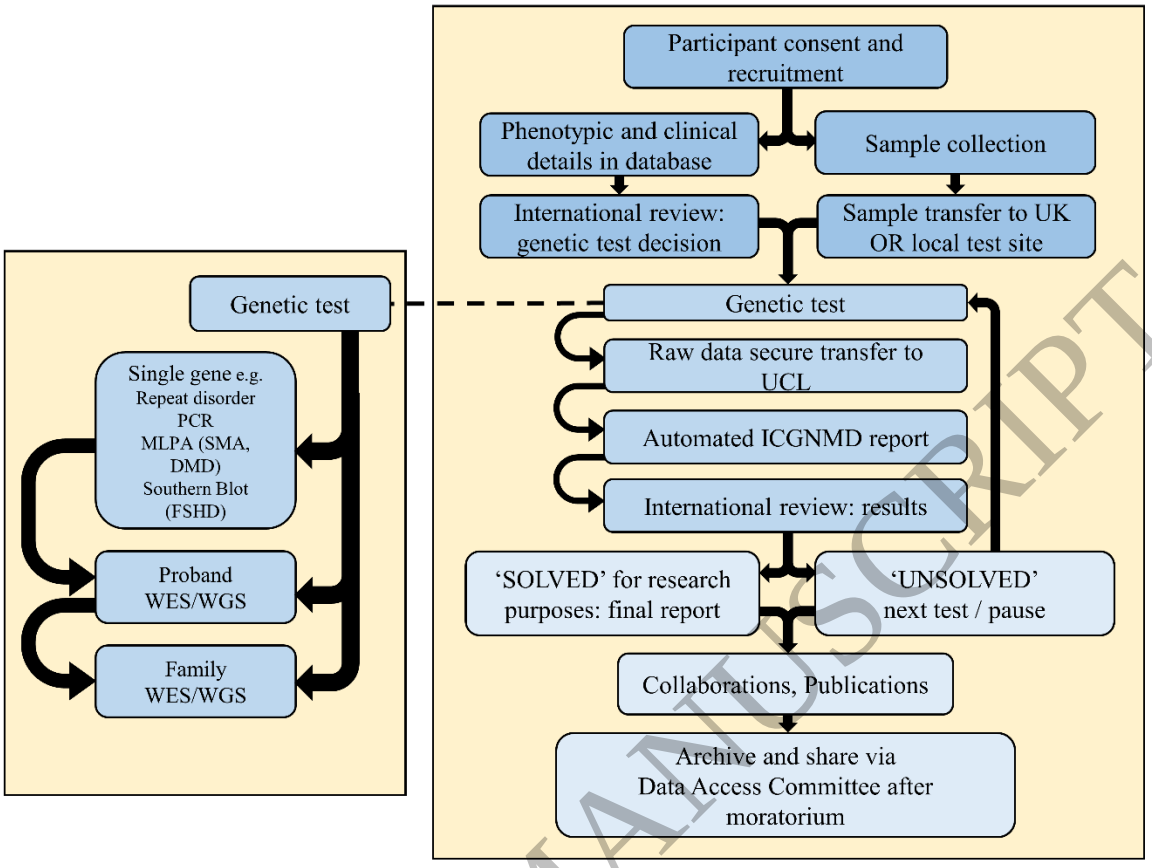

Figure 1
202x155 mm (x DPI)

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

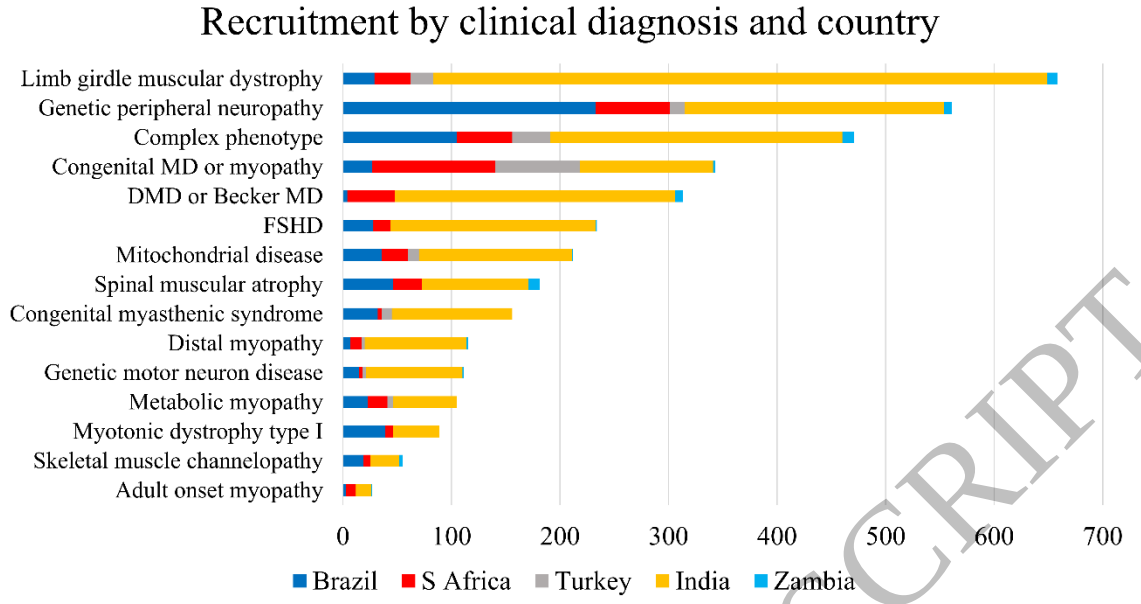

Figure 2
254x135 mm (x DPI)

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

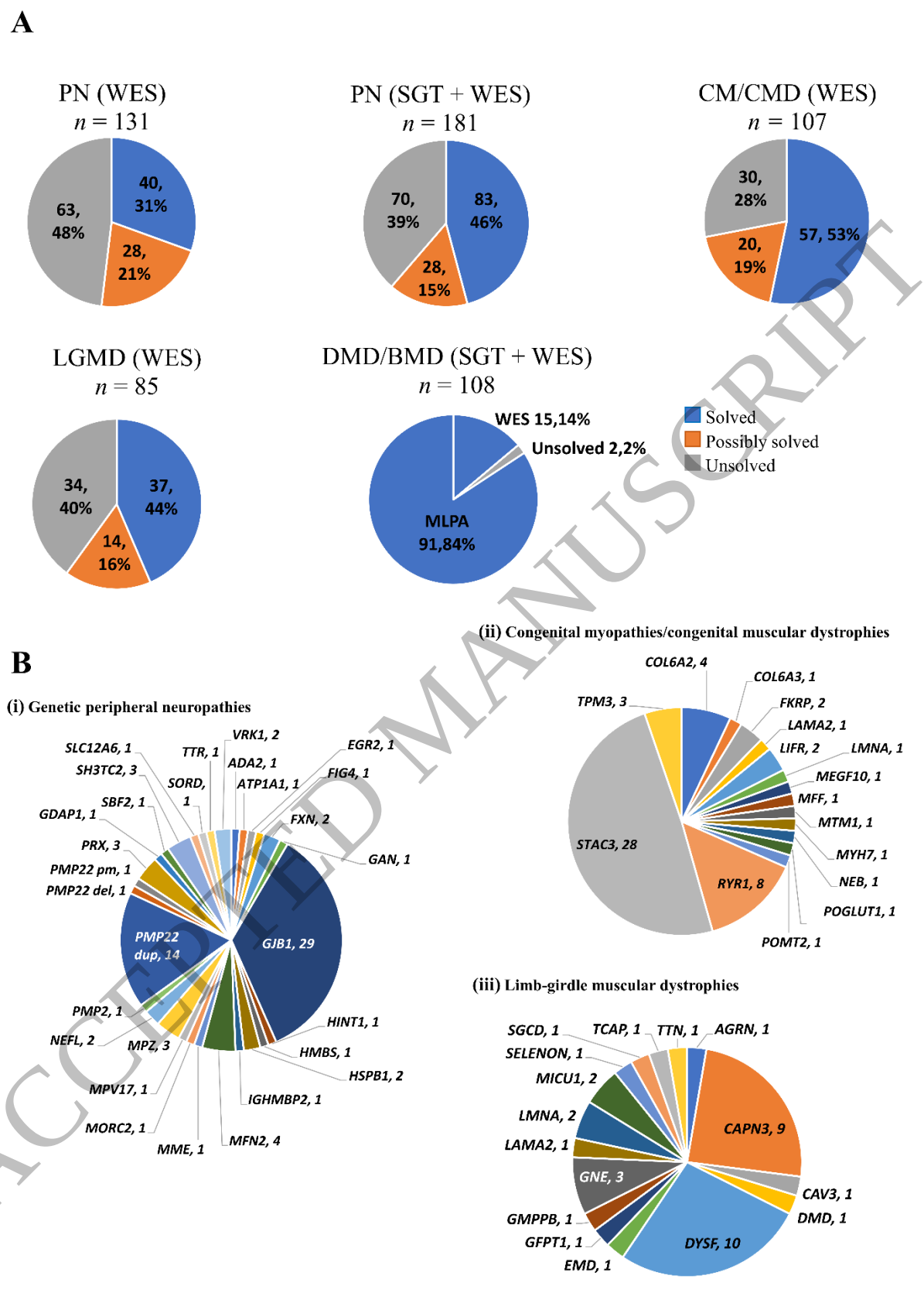

Figure 3
245x345 mm (x DPI)

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

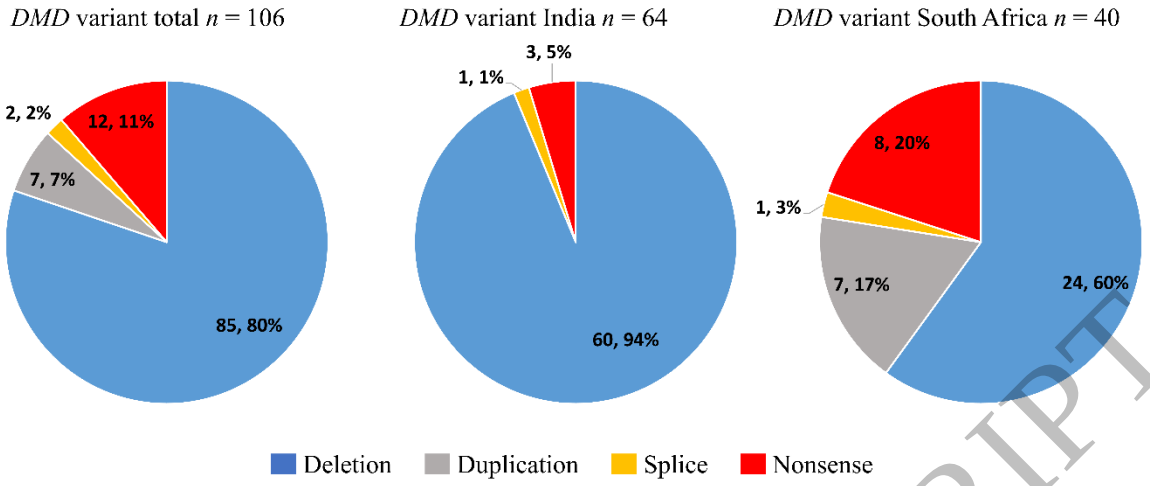

Figure 4
330x135 mm (x DPI)

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023

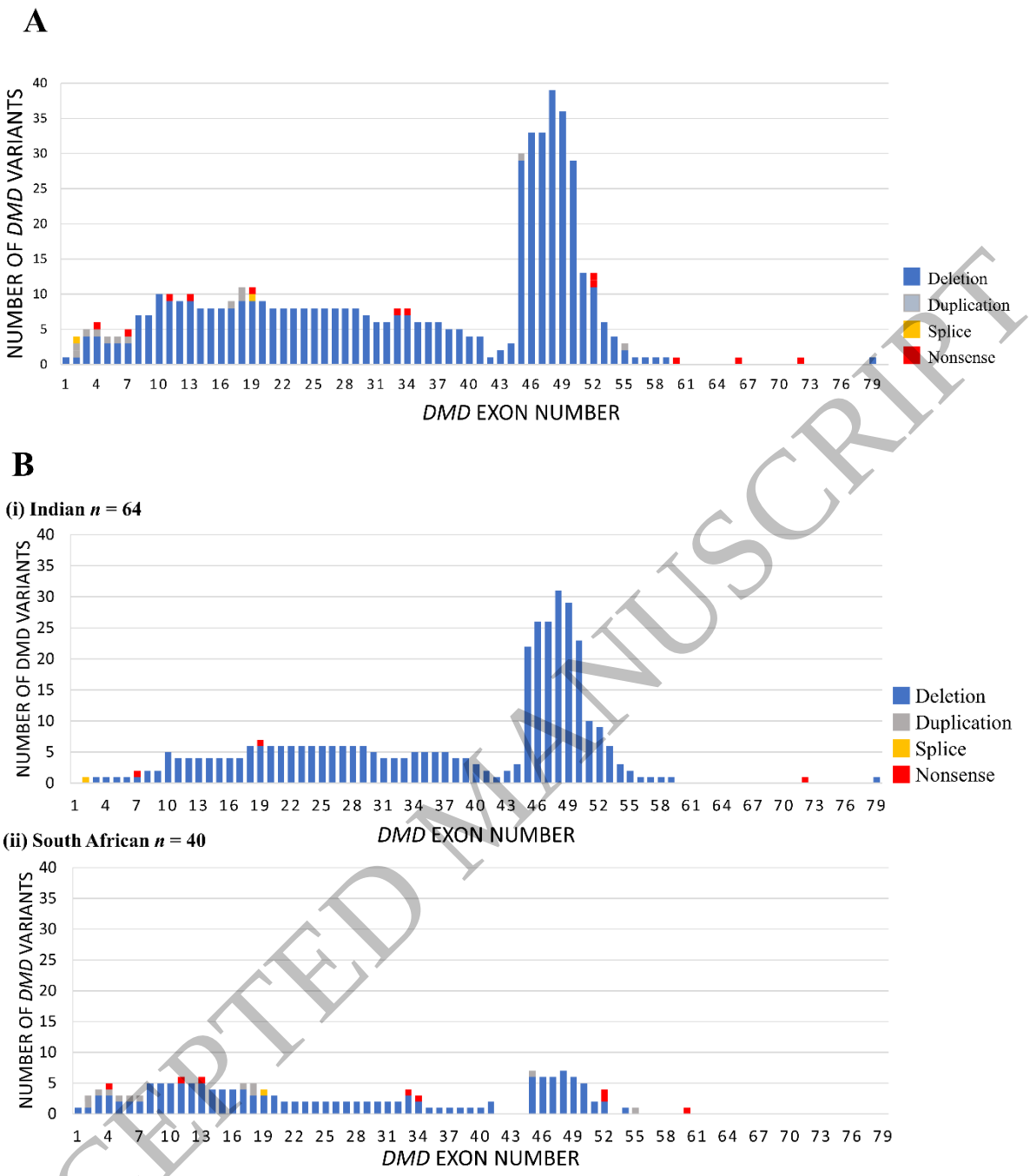

Figure 5
254x290 mm (x DPI)

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awad254/7233759 by Dokuz Eylul University user on 26 August 2023