Genetics of Hearing Impairment and Peripheral Neuropathy in Mali

by:

Abdoulaye YALCOUYE

YCLABD001

Submitted to the University of Cape Town

In fulfilment of the requirements of the PhD degree in Human Genetics

University of Cape Town, Faculty of Health Sciences, Department of Pathology, Division of Human Genetics

Supervisor: Prof. Ambroise WONKAM

Co-supervisor: Dr. Emile CHIMUSA

Date of submission: 16th August 2022

Acknowledgments

Words cannot express my gratitude to:

Allah, the almighty for the strength and guidance over this journey

My lovely spouse and daughter for their love and patience

My parents for their constant supports and best wishes

The National Institute of Health (NIH, USA), the Wellcome Trust, the University of Cape Town, and the HI-GENES Africa project

My supervisor and PI of this project, Professor Ambroise Wonkam for the coaching, guidance, training, and endless motivation and constant support. My eternal and endless gratitude to you

My co-supervisor and co-PI of HI-GENES Africa project, A/Professor Emile Chimusa for his valuable contribution to this project

My co-supervisor and co-PI of HI-GENES Africa project in Mali, Doctor Guida Landouré for the mentorship, opportunity, guidance, endless support, and encouragements. Your belief on my abilities helped me to keep my spirit and motivation so high to overcome this journey. My eternal and endless gratitude to you

Professor Cheick Oumar Guinto, Thank so much for your unconditional support.

Dr Carmen De Kock, the manager of HI-GENES Africa project, for her academic, administrative support. You are a great person

Dr Edmond Wonkam Tingang, my dear friend and colleague. Thank you so much for the precious assistance at any steps of this journey

Dr Oluwafemi Oluwole, Dr Samuel Mawuli Adadey, Dr Khuthala Mnika, Dr Karen Kengne, Mr Kevin Esoh, Mr Elvis Twumasi Aboagye and Mrs Noluthando Manyisa, for their support, lab training and assistance

The whole Neurogenetics and GeneMap teams for their encouragements and contributions

To my colleagues in the Division of Human Genetics, and international collaborators for their inputs to this work

All the patients and families who participated in this research project

Declaration

I, Abdoulaye Yalcouyé, hereby declare that this dissertation/thesis is based on my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been or is being submitted for another degree in this or any other university.

I give permission to the university to reproduce, for the purpose of research, either the whole or portion of the content, in any manner whatsoever.

Signature: Signed by candidate **Date:** 16th August 2022

Student Name: Abdoulaye Yalcouyé **Student Number**: YLCABD001

List of publications included in this thesis

I confirm that I have been granted permission by the University of Cape Town's Doctoral Degrees Board to include the following publications in my PhD thesis, and where co-authorships are involved, my co-authors have agreed that I may include these publications:

1. **Abdoulaye Yalcouyé**, Oumou Traoré, Abdoulaye Taméga, Alassane B. Maïga, Fousseyni Kané, Oluwafemi G. Oluwole, Cheick Oumar Guinto, Mohamed Kéita, Samba Karim Timbo, Carmen DeKock, Guida Landouré and Ambroise Wonkam. (2021) Etiologies of Childhood Hearing Impairment in Schools for the Deaf in Mali. Front. Pediatr. 9:726776. doi: 10.3389/fped.2021.726776. (status: published)

2. **Abdoulaye Yalcouyé**, Oumou Traoré, Salimata Diarra, Isabelle Schrauwen, Kevin Esoh, Magda Kamila Kadlubowska, Thashi Bharadwaj, Samuel Mawuli Adadey, Mohamed Kéita, Cheick O Guinto, Suzanne M Leal, Guida Landouré, Ambroise Wonkam. A monoallelic variant in *EYA1* is associated with Branchio-Otic syndrome in a Malian family. Mol Genet Genomic Med. 2022 Jun 14;e1995. doi: 10.1002/mgg3.1995. (status: published)

3. **Abdoulaye Yalcouyé**, Kevin Esoh, Landouré Guida, Ambroise Wonkam. Current profile of Charcot-Marie-Tooth disease in Africa: A systematic review. J Peripher Nerv Syst. 2022 Apr 5. doi: 10.1111/jns.12489. (status: published)

4. **Abdoulaye Yalcouyé**, Seybou H. Diallo, Thomas Coulibaly, Lassana Cissé, Salimata Diallo, Oumar Samassékou, Salimata Diarra, Dramane Coulibaly, Mohamed Keita, Cheick O. Guinto, Kenneth Fischbeck, Guida Landouré. A novel mutation in the *GARS* gene in a Malian family with Charcot-Marie-Tooth disease. Mol Genet Genomic Med. 2019;00:e782. doi.org/10.1002/mgg3.782. (status: published)

5. **Abdoulaye Yalcouyé**, Seybou H Diallo, Lassana Cissé, Mamadou Karembé, Salimata Diallo, Thomas Coulibaly, Salimata Diarra, Dramane Coulibaly, Mohamed Keita, Cheick O Guinto, Kenneth H Fischbeck, Ambroise Wonkam, Guida Landouré. *GJB1* variants in Charcot-Marie-Tooth disease X-linked type 1 in Mali. J Peripher Nerv Syst. 2022 Apr 5. doi: 10.1111/jns.12486 (status: published)

Signature: Signed by candidate **Date:** 16th August 2022

Student Name: Abdoulaye Yalcouyé **Student Number**: YLCABD001

List of publications the candidate contributed to (excluded)

1. Hotchkiss J, Manyisa N, Adadey SM, Oluwole OG, Wonkam E, Mnika K, **Yalcouye A**, Nembaware V, Haendel M, Vasilevsky N, Mulder NJ, Jupp S, Wonkam A, Mazandu GK. The Hearing Impairment Ontology: A Tool for Unifying Hearing Impairment Knowledge to Enhance Collaborative Research. Genes (Basel). 2019 Nov 21;10(12):960. doi: 10.3390/genes10120960. (status: published)

2. Manyisa N, Adadey SM, Wonkam-Tingang E, **Yalcouye A**, Wonkam A. Hearing Impairment in South Africa and the Lessons Learned for Planetary Health Genomics: A Systematic Review. OMICS. 2022 Jan;26(1):2-18. doi: 10.1089/omi.2021.0181. (status: published)

3. Oluwafemi Oluwole, Chandré Oosterwyk, Dominique Anderson, Samuel Adadey, Khuthala Mnika, Noluthando Manyisa, **Abdoulaye Yalcouye**, Edmond Wonkam, Elvis Aboagye, Yacouba Dia, Esther Uwibambe, Mario Jonas, Kalinka Popel, Carmen Dekock, Victoria Nembaware, Ambroise Wonkam. The Implementation of Laboratory Information Management System in Multi-Site Genetic Study in Africa. BMC Research Notes 2022. (status: under review)

Preface

Hearing impairment (HI) is one of the most common and disabling sensory disorders with a higher incidence in developing countries. In 2021, the World Health Organization (WHO) reported that over 5% of the world's population of which 34 million are children require rehabilitation of their disabling HI. It is estimated that 700 million people will suffer from HI by 2050, with a higher burden in sub-Saharan Africa (SSA) than in developed countries.

Although, connexin genes (*GJB2* and *GJB6*) have been reported to be the common causes of congenital hereditary HI in Caucasians, their contribution is nonsignificant in most SSA populations.

Inherited peripheral neuropathy (IPN) represents a heterogeneous group of diseases that affect the peripheral nerves. Charcot-Marie-Tooth disease (CMT) is the most common IPN worldwide with over 50% of all cases caused by mutations in the *PMP22* gene. However, the genetic profile of these diseases has not been extensively studied in most African countries, particularly in Mali where the high consanguinity and fertility rate offer a unique opportunity for gene discovery.

Research concept

The study concept and all the experiments were conceived by the candidate in conjunction with the principal investigator and primary supervisor Professor Ambroise Wonkam (UCT) and the co-supervisors A/Professor Emile Chimusa (UCT) and Doctor Guida Landouré (USTTB, Mali). The research proposal including methods for patient recruitment, sample and data collection, experiments, and analysis was designed by the candidate

with the advice and guidance of the supervisors. Besides, drafting the full manuscripts and incorporating revisions from co-authors and journals reviewers were performed by the candidate.

Data collection

All activities that were necessary prior to identifying our target population, including community engagement activities, visiting schools and hospitals were done in full by the candidate. Contacting and recruiting participants, obtaining informed consent, and physical examination of patients were done in collaboration with other specialists where necessary. Biological materials (blood samples) and patients' photographs were obtained by the candidate where needed. Also, data collection and entry were completed by the candidate.

Experimentation and data analysis

DNA extraction, polymerase chain reactions, and Sanger sequencing experiments were performed in full by the candidate. The primary analysis of data and *in silico* prediction of the pathogenicity of variants were performed by the candidate. When additional bioinformatics analysis was necessary, the involvement of collaborators was requested by the principal investigator, and with the close implication of the candidate. The work and contribution of these collaborators to the study is indicated in each publication where applicable.

Publications

Synthesis of all the results and drafting of manuscripts of the publications included in this thesis were generated by the candidate, after which revisions from all co-authors were similarly incorporated before submission to the

journal by the principal investigator or the candidate. After review, all reviewers' comments were addressed by the candidate under the supervision of supervisors. The detailed role of the candidate in the included publications is stated in each chapter.

To understand the genetic spectrum of HI and inherited peripheral neuropathy in Malian population, this research work undertaken in this thesis was written and published during the period of my registration at University of Cape Town as PhD candidate since 2019. Thus, we have chosen to include the following publications in the thesis, and hope this will highly contribute the specific field of the genetic of HI and PN not only in Mali but all over Africa. The seven chapters of this thesis are presented in the traditional format and include an introduction, aim and objectives, materials and methods, results, discussion, and conclusion sub-sections.

The inclusion of published work aimed to 1) highlight the candidate's proficiency of new knowledge in the relevant field, which will also strengthen his research and academic career; 2) enhance his candidacy for continued financial support from bodies that support this research work i.e. Hearing Impairment Genetics Studies in Africa (HI-GENES Africa), Genetic Medicine of African Populations (GeneMAP), Neurogenetic diseases study in Mali, and the University of Cape Town (UCT); 3) enable participation and attendance at academic conferences and training workshops, necessary to acquired knowledge to strengthen his capacity network with peers. The candidate successfully attended several specialized national and international conferences including the prestigious American Society of Human Genetics annual congress with travel award, Human Genome meeting in Australia with travel award, African Society of Human Genetics in

Rwanda, Peripheral Nerve Society annual meetings in Miami with travel award, Human Heredity and Health in Africa consortium meetings with one best poster prize and travel grants obtained; lastly, 4) these publications constitute key parts of a consistent body of research that fulfil the Faculty of Health Sciences (UCT) policies. These policies promote the publication of thesis research work as much as possible to disseminate knowledge generated and further improve the profile of the institution and the candidate.

This thesis is presented in four cohesive parts: 1) the current profile Charcot-Marie-Tooth disease in Africa; 2) application of target gene testing to solve CMT cases, highlights of the CMTX1 and CMT2D; 3) clinical description of HI in Mali; 4) WES and bioinformatics tools to identify causative variants in both SHI and NSHI in Mali.

The candidate has met all requirements and approval of UCT's Doctoral Degrees Board, under Rules GP6.7 as follows:

1. The candidate's proposal to include publications in the current thesis was approved by the UCT Faculty of Health Sciences Doctoral Degree Board.
2. The thesis contains an adequate summary, introduction, a chapter on the aims and objectives, a comprehensive academic discussion of the results , forming the basis of the conclusions and perspectives drawn from this research.
3. Each result chapter with publications included is preceded by a synopsis of how the publications directly tie to the aims and objectives of the project, as well as to the thesis with a consistent formatting throughout the thesis.
4. All included publications were written and published during the candidate's registration as a PhD student since 2019.

The candidate

Abdoulaye Yalcouyé

Abstract

Background 1

Hearing impairment (HI), the most common sensory disturbance, affects about 1 in 1000 living newborns globally. Its incidence is reported higher in sub-Saharan African (SSA) populations. HI is caused by environmental and genetics factors. In many developing countries, environmental factors are reported to be the most prevalent etiologies while genetic causes are predominant in the developed countries. Genetic HI epidemiology varies from one region to another. Over 50% of congenital HI has a genetic origin and, according to their clinical expression, can be classified as syndromic hearing impairment (SHI) and non-syndromic hearing impairment (NSHI). SHI refers to a HI that is associated to other organ symptoms, while NSHI refers to HI that only affect the hearing without other symptoms. Syndromic HI accounts for up to 30% of hereditary hearing impairment (HHI) and more than 400 HI-associated syndromes have been described including Waardenburg syndrome (OMIM: 193500), Branchio-oto-renal syndrome (OMIM: 113650), Usher syndrome (OMIM: 276900), Pendred syndrome (OMIM: 274600), Keratitis-Ichthyosis Deafness syndrome (OMIM: 148210) and Alport syndrome (OMIM: 301050). Among HHI, 70% are non-syndromic and all mendelian modes of inheritance are seen, with autosomal recessive being the most common pattern. Regarding the mode of inheritance, NSHI is divided into autosomal dominant deafness (DFNA, 15 to 20%), autosomal recessive deafness (DFNB, 80%), X-linked deafness (DFNX, 1%), and mitochondrial deafness which represents 1%. Autosomal dominant NSHI is often post-lingual and progressive, whereas recessive NSHI is prelingual. To date, more than 120 genes are associated with NSHI. Despite this large

number of identified genes, only *GJB2* (OMIM: 121011) and *GJB6* (OMIM: 604418) are systematically studied in SSA population for which the prevalence of NSHI-causal variants is insignificant. Yet, no case of HI has been genetically investigated in Mali where consanguineous marriage is a common practice that may increase recessive diseases.

Background 2

Peripheral neuropathies (PNs) are a group of conditions that affect the peripheral nerves and are caused by either genetic or non-genetic etiologies. While symptomatic and curative treatments for PNs are possible via addressing the underlying etiology, with subsequent nerve cell regeneration and resolution of the condition, curative treatments for PNs of genetic origin have been elusive. Charcot-Marie-Tooth disease (CMT), also known as hereditary motor and sensory neuropathy (HMSN), is the most common PN group with a high clinical and genetic heterogeneity with more than 100 genes identified to date, mostly in populations of Caucasian ancestry. Population-based studies have reported variable prevalence, with a crude global estimate of 1/2500. Yet, despite being described more than 130 years ago, and the genetic cause identified about 30 years ago, there remains a paucity of information on its global prevalence and genetic epidemiology due largely to challenges in diagnosis, especially in countries with limited resources. Phenotypically, the classic CMT manifest with distal muscle weakness and wasting, sensory loss, steppage gait and skeletal deformities (*pes cavus* and *planus*). CMT exhibits high phenotype variability with some additional symptoms associated such pyramidal sign, vocal paralysis, scoliosis, lordosis, hearing impairment. HI, a common audiological symptom

associated with CMT is caused by several genes including *GJB1*. CMT is inherited in autosomal dominant, recessive and X-linked manner. However, the genetic epidemiology of peripheral neuropathy is largely unknown in Africa, and have not been investigated in Mali.

Aim and methods

This study aimed to investigate the genetic aetiology of HI and PN in Mali. To achieve this goal, we performed 1) a literature review on the current profile Charcot-Marie-Tooth disease in Africa, 2) a clinical characterization and target gene testing to solve CMT cases, emphasizing on CMTX1 and CMT2D, 3) a clinical description of HI in Mali complemented with a literature review, 4) whole exome sequencing (WES) technique and bioinformatics tools to identify causative genes and variants involved in Branchio-otic syndrome and non-syndromic HI in Mali.

Results

Literature review on CMT in Africa

In this study, we screened articles that reported CMT cases in Africa. A total of 107 families totaling 185 patients were described. Most studies were reported in North African patients (n=22). The demyelinating form of CMT was the commonest subtype, and the phenotype varied greatly between families, and only one patient (1%) with CMT associated HI. The inheritance pattern was autosomal recessive in 91.2% (n = 97/107) of families. CMT-associated variants were reported in 11 genes: *LMNA* (OMIM: 150330), *GDAP1* (OMIM: 606598), *GJB1* (OMIM: 304040), *MPZ* (OMIM: 159440), *MTMR13* (OMIM: 607697), *MTMR2* (OMIM: 603557), *PRX* (OMIM: 605725), *FGD4/FRABIN* (OMIM: 611104), *PMP22* (OMIM: 601097), *SH3TC2* (OMIM:

608206), and *GARS* (OMIM: 600287). The most common genes reported are *LMNA*, *GDAP1* and *SH3TC2* and have been found mostly in Northern African populations. In addition, the most common CMT-associated gene (*PMP22*) did not contribute significantly to CMT in Africa (3,7%). This study reveals that CMT is not rare in Africa and describes the current clinical and genetic profile. The review emphasized the urgent need to invest in genetic research to inform counselling, prevention, and care for CMT in numerous settings on the continent.

Target gene testing in CMT in Mali

***GJB1* related neuropathy**

In this first case, three unrelated families with CMTX1 (OMIM: 118220) features were seen and the DNA samples were submitted to the Medical Neurogenetics, Atlanta, GA for CMT genes panel (50 genes, PMP22dup and mtDNA). A total of 15 patients (six males and nine females with a sex ratio of 0.67) from three unrelated families were found to be affected. The disease segregation in the family pedigrees was consistent with an X-linked dominant inheritance pattern. The mean age at onset was 23.6 years, ranging from 8 to 51 years, and the mean age at diagnosis was 27.3 years, ranging from 15 to 63 years. The predominant starting symptom was tingling in distal limbs, and the chief complaint was gait difficulty. Neurological examination found a distal muscle weakness and atrophy and sensory loss, skeletal deformities, decreased or absent reflexes and steppage gait. Nerve conduction studies (NCS) showed no response in most of the tested nerves. A severe sensorineural hearing impairment and a focal epileptic seizure were observed in two patients. A high intra and inter-familial clinical variability was observed. Genetic testing found three different pathogenic missense

variants in *GJB1* gene in the respective families (Val91Met; Arg15Trp and Phe235Cys) and were predicted to be deleterious by several bioinformatic tools.

GARS related neuropathy

In the second case, two individuals, a male, and a female, were found to be affected. Symptoms started in their teenage years with muscle weakness and atrophy in hands. Later, distal involvement of the lower limbs was noticed. Patients complained of minor sensory impairment. NCS showed no response in the upper as well as the lower limbs. DNA from the proband was used for CMT gene panel testing (including 50 genes, *PMP22* duplication and mtDNA). Surprisingly, a novel heterozygous missense variant (c.794C>A; p.Ser265Tyr) was identified in the *GARS* gene associated with CMT2D. This variant segregated with the disease in the family and was also seen in the mother who presented no symptoms. This is the first report of a genetically confirmed CMT2D case in Africa, expanding its genetic epidemiology.

Clinical description of HI in Mali

In this study, a total of 117 individuals from the two schools for the deaf in Bamako (capital city of Mali) were included with male predominance (sex-ratio 1.3; 65/52). HI was prelingual in 82.2% (n = 96), and the median of age at diagnosis was 12.7 years. The etiologies were environmental in 59.4% (n = 70), led by meningitis in 39.7% (n =28), followed by likely genetics in 29.3% (n = 34), while 11.3% (n = 13) were unknown. Amongst potential genetically suspicious cases, three cases were syndromic including two cases of Waardenburg syndrome, and 15 individuals featured non-syndromic HI.

Autosomal recessive inheritance pattern was observed in 83.9% of the families (n=15), and consanguinity was reported in 58.1% (n=10) of those cases. Environmental factors are the leading causes of HI in Mali. However, genetic causes should be investigated, particularly in the context of a population with a high consanguinity rate. Following a review of the literature, we identified eight articles that met our selection criteria. All the studies were performed in Bamako. Only one was conducted at the school for deaf, while four others were performed in hospital settings. We found three case reports, and five cross-sectional case series. None of the studies reported the prevalence or incidence of HI in Mali, nor the genetic contribution in HI in Mali.

WES and bioinformatics tools to identify causative variants in both SHI and NSHI in Mali

We submitted DNA samples of two individuals for whole exome sequencing (WES) at Omega Bioservices (Norcross, GA, USA). In parallel to WES , we performed Sanger sequencing for all the available family members using primers designed in our laboratory to target coding exons and splice sites of the *EYA1*, *SIX1* and *SIX5* genes. We identified a monoallelic pathogenic variant in *EYA1* [c.1286A>G: p.Asp429Gly; OMIM: 601653] segregating with Branchio-Otic (BO) syndrome in eight individuals from a large non-consanguineous Malian family with autosomal dominant inheritance. The ages at diagnosis ranged from 8 to 54 years. A high phenotypic variability was noted among the four male and four female affected individuals, with a post-lingual and mixed type of HI (n = 4/8). One individual had conductive HI while three had normal hearing but presented other BO features, namely, branchial fistulae and preauricular sinus. Renal screening, including serum

creatinine levels and ultrasonography, was normal in three affected individuals available for this testing. This is the third genetically confirmed case of BO syndrome in the sub-Saharan African population, expanding its genetic spectrum.

In this second part, we submitted DNA samples of eight multiplex families segregating HI to Omega Bioservices (Norcross, GA, USA) for WES. Consanguinity was reported in six families and the inheritance pattern was suggestive of autosomal recessive in all families. WES performed in probands identified known homozygous variants in three families. Two missense variants were located in the *MYO15* gene [c.A6331T; p.(N2111Y) and c.G8158A; p.(D2720N), OMIM: 602666] and one nonsense variant in *MYO7A* gene [c.C3978A; p.(C1326X), OMIM: 276903]. In addition, five novel variants in HI-associated genes were identified in four families. Compound heterozygous variants in *OTOGL* gene [c.209-9C>G and c.Cys5685Ala; p.Asp1895Glu); OMIM: 614925], a missense variant in *CDH23* gene [c.C646G; p.Leu216Val; OMIM: 605516] and PJVK gene [c.T461G; p.Val154Gly; OMIM: 610219], and homozygous splicing variant in *TMC1* gene [c.2003+2T>C; OMIM: 606706)]. Moreover, in one family, WES revealed a missense variant [c.58G>A; p.Glu20Lys] in *UBFD1* gene which was not a known to be associated with HI and could be a candidate novel gene for it. Several in silico prediction tools showed these sequence changes were damaging. Furthermore, Sanger sequencing was performed to confirm the segregation in the respective families. This study reports the first genetically confirmed cases of NSHI in Mali, and functional studies are underway to confirm the pathogenicity of variant in the novel gene.

Conclusion

This study revealed a different CMT genetic epidemiology with the most common CMT-related genes not contributing much to Africa. Genetic testing in the Malian population confirmed CMT cases in four unrelated families with known variants in the *GJB1* gene in three families and a novel variant in the *GARS* gene. These CMT cases are the first cases in the sub-Saharan African population. As previously reported, HI was found to be associated with CMT and recessive cases of CMT were found to be the most common type in Africa particularly in the northern African population. In addition, this study revealed a high contribution of the environmental factors in the occurrence of HI among school-aged children in Mali, with meningitis being the most common factor identified. Nevertheless, genetic cases were identified with novel findings in five families including a putative novel gene. Moreover, it showed that *GJB2* and *GJB6* genes do not contribute significantly to HI in the Malian population. Our project highlighted the need of conducting large scale genetic studies of HI and PN in population with African ancestry to potentially uncover hundreds of genes linked to these diseases.

Table of contents

List of figures

List of Tables

Abbreviations

aa	amino acid
AAS	African Academy of Science
ABR	Auditory Brainstem Responses
ACMG	American College of Medical Genetics and Genomics
AD	Autosomal Dominant
ADA	Adaptive boosting
ADNSHI	Autosomal Dominant Nonsyndromic Hearing Impairment
ANNOVAR	Annotate Variation
AR	Autosomal Recessive
ARNSHI	Autosomal Recessive Non-syndromic Hearing Impairment
ATP6V1B1	ATPase, H+ Transporting, Lysosomial, 56/58-KD, V1 Subunit B, Isoform1
BIAP	Bureau International d'AudioPhonologie
BLAST	Basic Local Alignment Search Tool
BLASTp	Protein-Protein Basic Local Alignment Search Tool
BO	Branchio-Otic
BORSD	Branchio-Oto-Renal Spectrum Disorders
CADD	Combined Annotation Dependent Depletion
CDH23	CADHERIN 23
CMT	Charcot-Marie-Tooth
CMT1A	Charcot-Marie-Tooth type 1A
CMT1F	Charcot-Marie-Tooth type 1F
CMT2D	Charcot-Marie-Tooth type 2D
CMT2J	Charcot-Marie-Tooth type 2J
CMT4B	Charcot-Marie-Tooth type 4B

CMTX1	Charcot-Marie-Tooth type X1
CMTX5	Charcot-Marie-Tooth type X5
CMV	Cytomegalovirus
DANN	Deleterious Annotation of genetic variants using Neural Networks
dbSNP	Single Nucleotide Polymorphism database
DFNA	Autosomal Dominant Non-syndromic Hearing Impairment
DFNB	Recessive Dominant Non-syndromic Hearing Impairment
DFNB59	Dominant Non-syndromic Hearing Impairment type 59
DFNX	X-linked Non-Syndromic Hearing Impairment
DNA	Deoxyribonucleic Acid
dSMA-V	Distal Spinal Muscular Atrophy type 5
EEG	Electroencepahlography
ENT	Ear Nose and Throat
ES	Exome Sequencing
ExAC	Exome Aggregation Consortium
EYA1	Eyes Absent 1
FATHMM-MKL	Functional Analysis Through Hidden Markov Models
FGD4	FYVE, RhoGEF, and PH Domain-Containing Protein 4
FMOS	Faculté de Médecine et d'Odontostomatologie
GA	Geargia
GARS	Glycyl-tRNA Synthetase
GATK	Genome Analysis ToolKit
GDAP1	Ganglioside-induced Differentiation-Associated Protein 1
GeneMAP	Genetic Medicine of African Populations
GERP	Genomic Evolutionary Rate Profiling
GJB1	Gap Junction Beta 1

GJB2	Gap Junction Beta 2
GJB6	Gap Junction Beta 6
gnomAD	Genome Aggregation Database
HHH	Hereditary Hearing loss Homepage
HHI	Hereditary Hearing Impairment
HI	Hearing Impairment
HI-GENES	Hearing Impairment Genetics Studies
HMSN	Hereditary Motor and Sensory Neuropathy
HREC	Human Research Ethics Committee
HUGO	Human Genome Organization
IA	Iowa
IDT	Integrated DNA Technology
Indel	Insertion/deletion
IPN	Inherited Peripheral Neuropathy
IRB	Institutional Research Board
LMNA	LAMIN A/C
LRT	Likelihood Ratio Test
MAF	Minor Allele Frequency
MD	Molecular Dynamic
MFN2	Mutofusin 2
MLPA	Multiplex Ligand-dependent Probe Amplification
MOY6	Myosin VI
MPZ	Myelin Protein Zero
mRNA	Messenger RNA
MSA	Multiple Sequence Alignment
mtDNA	Mitochondrial DNA
MTMR2	Myotubularin-Related Protein 2

MYO15	Myosin XV
MYO7A	Myosin VII
NCBI	National Centre for Biotechnology Information
NCS	Nerve Conduction Study
NCV	Nerve Conduction Velocities
NGS	Next-Generation Sequencing
NIH	National Institutes of Health
NINDS	National Institute of Neurological Disorders and Stroke
NSHI	Non-Syndromic Hearing Impairment
NSHL	Non-Syndromic Hearing Loss
OMIM	Online Mendelian Inheritance in Man
OTOGL	Otogelin-Like Protein
PCDH15	Protocadherin 15
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEPNS	PolyEndocrine-PolyNeuropathy Syndrome
phyloP	Phylogenetic *P*-values
PJVK	Pejvakin
PLP	Pathogenic or Likely Pathogenic
PMP22	Peripheral Myelin Protein 22
PN	Peripheral Neuropathy
PNs	Peripheral Neuropathies
PolyPhen	Polymorphism Phenotyping
PRISMA	Preferred Reporting Items for Systematic Review and Meta-Analysis
PROSPERO	International Prospective Register of Systematic

	Reviews
PROVEAN	Protein Variation Effect Analyzer
PRX	Periaxin
PSI-BLAST	Position-Specific Iterative Basic Local Alignment Search Tool
PTA	Pure Tone Audiometry
Ref	Reference
RefSeq	Reference Sequence
REVEL	Rare Exome Variant Ensemble Learner
RF	Random Forest
RFLP	Restriction Fragment Length Polymorphism
SCs	Schwann Cells
SD	Standard Deviation
SH3TC2	SH3 Domain and Tetratricopeptide Repeat Domain 2
SHI	Syndromic Hearing Impairment
SIFT	Sorting Tolerant from Intolerant
Siphy	Site-specific phylogenetic analysis
SIX1	Six Homeobox 1
SIX5	Six Homeobox 5
SNP	Single Nucleotide Polymorphism
SNV	Single Nucleotide Variant
SSA	Sub-Saharan Africa
SSCP	Single-Strand Conformational Polymorphism
TMC1	Transmembrane Channel-like Protein 1
UBFD1	Ubiquitin Family Domain Containing 1
UCT	University of Cape Town
USA	United States of America

USH1G	USHER Syndrome, Type 1G
USH2A	USHER Syndrome, Type 2A
USTTB	Université des Sciences, Techniques et Technologies de Bamako
WES	Whole-Exome Sequencing
WGS	Whole-Genome Sequencing
WHO	World Health Organization
WS	Waardenburg Syndrome

Chapter 1: Introduction

Hearing impairment (HI) is a partial or total loss of hearing. About 1.5 billion of the world's population have some degree of HI and 12.6 million of them suffer from profound HI [1]. Although its prevalence varies greatly across different countries, 80% of HI cases live in developing countries [1]. In the United States, 1 in 1000 living newborns has severe to profound sensorineural HI, and an additional 1 to 2 per 1000 are born with mild to moderate, clinically significant bilateral or unilateral HI [2, 3]. Sub-Saharan Africa (SSA) is one of the most burdened areas, and HI incidence is six times higher than developed countries [4, 5]. According to the World Health Organization (WHO), at least 5% of the world's population have disabling HI, which affect the language acquisition and psychosocial development [1]. Furthermore, children presenting with hearing loss early in life (either permanent or temporary) are likely to miss critical developmental periods essential for the connection of higher-level auditory processes [6]. Hearing Impairment can be classified by age at onset, severity or other audiometric characteristics, and the presence of other physical or medical characteristics. Causes include genetic and environmental factors acting independently, as well as genetic susceptibility and environmental exposures acting together [7, 8]. Among environmental causes, meningitis, chronic otitis media, chronic disease and other infectious diseases are the most prevalent in developing countries [1, 9, 10]. Other environmental causes include birth injury, neonatal anoxia and jaundice [1]. Genetic HI is the common type in developed countries with at least 50% of congenital HI being from genetic origin [11]. However, certain environmental causes, such as congenital cytomegalovirus (CMV) infection, continue to play a major

aetiological role in late-onset childhood HI [7]. Hereditary HI (HHI) are categorized into two groups, non-syndromic (NSHI) and syndromic (SHI) [12]. NSHI refers to those with isolated HI and represents 70% of all genetic cases while SHI, in addition to the HI phenotype, harbors other clinical manifestations [13, 14]. To date, more than 400 syndromes associating HI are known including Waardenburg, Norrie, Alport Syndrome, Branchio-Oto-Renal Syndrome, CHARGE Syndrome, Jervell and Lange-Nielsen, Pendred Syndrome, Perrault Syndrome, Stickler Syndrome, Treacher Collins Syndrome, Usher Syndrome [8, 15]. Advances in genetics and genomics in the past few decades have enabled major progress in the understanding of pathobiology of hearing process. To date more than 170 loci and over 124 genes are known to be associated with HI [15].

Despite the large number of genes implicated in NSHI, a particular gap junction beta 2 protein coded by *GJB2* gene (OMIM: 121011), located at 13q12.11, accounts for nearly 50% of all autosomal recessive HL and 15% of HL in all deaf individuals [16, 17]. *GJB2* and Gap junction beta 6 protein coded by *GJB6* (OMIM: 604418) are the most prevalent genes implicated in over 50% of all HHI. Genetic HI shows great heterogeneity across different populations and the systematic screening of *GJB2* was shown to be less contributive in many SSA populations [10, 18-20]. Even though, there is lack of countrywide studies in many if not all countries, a recent study on the genetic profile of HI found that *GJB2* was the most investigated gene in Africa, with the *GJB2* p.Arg143Trp founder variant only reported in Ghana, while *GJB2* c.35delG was common in North African countries [18]. Variants in *MYO15A* (OMIM: 602666) were the second most frequently reported in both North and Central Africa, followed by *ATP6V1B1* (OMIM: 192132) only

reported in North Africa. Usher syndrome was the main syndromic HI molecularly investigated, with variants reported in five genes: *USH2A* (OMIM: 608400), *USH1G* (OMIM: 607696), *USH1C* (OMIM: 605242), *MYO7A* (OMIM: 276903), and *PCDH15* (OMIM: 605514). The *MYO7A* variant [p.Pro1780Ser] was reported as the common Usher syndrome variant among Black South Africans [18, 21, 22]. Another study from Senegal reported that *GJB2* variants accounted for 34.1% of NSHI with the variant c.94C>T (p.Arg32Cys) accounting for 25% of familial cases [23]. Other genes including *MYO15A*, *TMC1* (OMIM: 606706), and *CDH23* (OMIM: 605516) are reportedly prevalent in some populations. In contrast to NSHI, etiological and phenotypic diversity make the identification of syndromic cases more challenging for both the diagnosis and management [8].

Inherited peripheral neuropathies (IPNs) are associated with variable degree of HI [24]. Charcot-Marie-Tooth (CMT) disease is the most common type of IPN reported in 1 per 2500 [25, 26]. CMT, also known as hereditary motor and sensory neuropathy (HMSN), presents a high clinical and genetic heterogeneity with more than 100 genes identified to date, mostly in populations with Caucasian ancestry [27]. *PMP22* (OMIM: 601097), *GJB1* (OMIM: 304040), *MFN2* (OMIM: 608507), *MPZ* (OMIM: 159440) genes are the commonest genes involved in CMT [28-30]. Phenotypically, CMT manifests classically with distal muscle weakness and wasting and sensory loss, steppage gait and skeletal deformities (*pes cavus* and *planus*). CMT exhibits a high phenotype variability with additional symptoms including HI, vocal paralysis, scoliosis, and lordosis [31]. Hereditary HI and IPN are both monogenic conditions transmitted following all mendelian inheritance

patterns (autosomal dominant, autosomal recessive, or sex-linked) and mitochondrial [31, 32].

HI is the most prevalent audiological symptom seen in a wide range of CMTs including CMT1A, CMT2A, CMTX1, CMTX5, CMT1F, CMT2J among others [24]. There are nearly forty genes that are known to be involved in both HI and IPN [24]. Interestingly these genes include the most common CMT-associated genes, namely *PMP22*, *GJB1*, *MFN2*, and *MPZ*. However, in a French cohort of 3,412 CMT-patients, HI was reported in 1.3% only. The exact mechanism underlying HI in patients is not well established. A recent study showed that *PMP22* mutation alone is not enough to explain hearing loss in patients suffering from IPN and can be due to cochlear impairment and/or auditory nerve dysfunction [24]. Although, HHI and PN are not life-threatening diseases, they cause serious disability to patients because no curative treatment is available. Evidence-based diagnosis and management of patients with HHI and IPN is made by a multidisciplinary team including ENT specialists, audiologists, neurologists, ophthalmologists, dermatologists, and pediatricians [27, 33]. This is particularly important in the cases of syndromic cases where a thorough clinical examination is required for better counselling [34].

The increasing access to sequencing offers a great opportunity to decipher their genetics underlying defects. The use of targeted sequencing has been shown to be less appropriate in research setting mainly due to its high cost and time consuming, especially in sub-Saharan African population [35]. Furthermore, targeted sequencing (including 180 known and candidate genes) applied on familial NSHI cases showed that known variants accounted for only 4% in Nigerian and South African patients while it was

57% in Iranian families [36]. Next generation sequencing (NGS) technique has been used efficiently to identify genetic variants causing IPN and HI [27, 37-39] and was shown as an alternative to the traditional sanger sequencing.

Like in most developing countries, genetic studies of monogenic diseases were less regarded in the past in Mali [20]. The Malian population is highly diverse with a long history of consanguinity practice reaching 29% in the general population [40]. In addition, a high fertility rate is seen with about 6.5 live births per woman. Common and rare genetic variants associated with HI and PN were not reported in Mali, though novel genetic findings were previously identified in several other neurological diseases [41-43].

Chapter 2: Aim and objectives

2.1 Aim

This study aimed to clinically characterize patients with HI and IPN in Mali and to identify their underlying genetic defects using different approach from target gene testing to whole exome sequencing techniques.

2.2 Objectives

1- Systematic literature search

1-1 To determine the profile of CMT in Africa

2- Panel gene testing in CMT cases

3-1 To identify the genetic cause of CMTX1, using target gene testing

3-2 To identify the genetic cause of CMT2D, using target gene testing

3- Clinical description of HI in Mali

3-1 To determine the clinical presentation of HI in Mali

4- WES and bioinformatics tools to identify causative variants in HI

4-1 To identify variants causing Branchio-otic syndrome using WES

4-2 To identify known and candidate genes in NSHI cases using WES

Chapter 3: Results- Literature review

Synopsis: This chapter presents the most recent and complete scientific knowledge on Charcot-Marie-Tooth disease in Africa. It provides the prevalence, incidence, clinical description and the genetic features of CMT. This chapter include one peer reviewed systematic review article.

3.1 **Abdoulaye Yalcouye**, Kevin Esoh, Guida Landouré, Ambroise Wonkam. Current profile of Charcot-Marie-Tooth disease in Africa: A systematic review. J Peripher Nerv Syst. 2022 Jun;27(2):100-112. doi: 10.1111/jns.12489. Epub 2022 Apr 5.

Nature of publication: review article

Publisher: Journal of the Peripheral Nervous System : JPNS. ISSN: 1085-9489, peer review journal

Candidate Contribution: Development of the protocol of the review. Literature search, selection of studies, extraction, analysis, and interpretation of data. Issuance of the first draft of the manuscript and incorporation of subsequent revisions.

Co-authors contribution:

GL: Conceptualization of the study, funding acquisition and critical revision of the successive versions

AW: Conceptualization of the study, funding acquisition and critical revision of the successive versions

KE: Contributed to development of the methodology

Current profile of Charcot-Marie-Tooth disease in Africa: A systematic review

Abdoulaye Yalcouyé[1,2], Kevin Esoh[2], Guida Landouré[1,3,4], Ambroise Wonkam[2,5]*

[1]Faculté de Médecine et d'Odontostomatologie, USTTB, Bamako, Mali

[2]Division of Human Genetics, Department of Pathology, Faculty of Health Sciences, University of Cape Town, 7925, South Africa

[3]Neurogenetics Branch, National Institute of Neurological Disorders and Stroke, Bethesda, Maryland

[4]Service de Neurologie, Centre Hospitalier Universitaire du Point "G", Bamako, Mali

[5]McKusick-Nathans Institute, and Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

*Corresponding author: ambroise.wonkam@uct.ac.za; Tel: +27-21-4066-307

Received: 21 October 2021, Revised: 17 January 2022, Accepted: 25 February 2022

Abstract

Background and aims: Charcot-Marie-Tooth disease (CMT) is the most common inherited peripheral neuropathy characterized by a high clinical and genetic heterogeneity. While most cases were described in populations with Caucasian ancestry, genetic research on CMT in Africa is scant. Only a few

cases of CMT have been reported, mainly from North Africa. The current study aimed to summarize available data on CMT in Africa, with emphasis on the epidemiological, clinical, and genetic features. **Methods:** We searched PubMed, Scopus, Web of Sciences, and the African Journal Online for articles published from the database inception until April 2021 using specific keywords. A total of 398 articles were screened, and 28 fulfilled our selection criteria. Results: A total of 107 families totaling 185 patients were reported. Most studies were reported from North Africa (n=22). The demyelinating form of CMT was the commonest subtype, and the phenotype varied greatly between families, and one family (1%) of CMT associated with hearing impairment was reported. The inheritance pattern was autosomal recessive in 91.2% (n=97/107) of families. CMT-associated variants were reported in 11 genes: *LMNA*, *GDAP1*, *GJB1*, *MPZ*, *MTMR13*, *MTMR2*, *PRX*, *FGD4/FRABIN*, *PMP22*, *SH3TC2*, and *GARS*. The most common genes reported are *LMNA*, *GDAP1*, and *SH3TC2* have been found mostly in Northern African populations. **Interpretation:** This study reveals that CMT is not rare in Africa and describes the current clinical and genetic profile. The review emphasized the urgent need to invest in genetic research to inform counselling, prevention, and care for CMT in numerous settings on the continent.

Keywords: Africa, Charcot-Marie-Tooth disease, clinical, epidemiology, genetic

3.1.1 Introduction

An estimated 2-7% of the world population suffers from a peripheral neuropathy (PN) [44]. Although rarely life-threatening, PN can be severely disabling, leading to wheelchair dependence. PN can be of either genetic or non-genetic aetiology. While symptomatic and curative treatments for acquired PNs are possible via addressing the underlying aetiology, with subsequent nerve cell regeneration and resolution of the condition, curative treatments for those of genetic origin have been elusive [45].

CMT, the most common IPN is highly heterogeneous. Population-based studies have reported variable prevalence [46], with a crude global estimate of 1/2500 [25]. Yet, despite being described more than 130 years ago, and the genetic cause identified about 30 years ago [47, 48], there remains a paucity of information on its global prevalence and genetic epidemiology due largely to challenges in diagnosis, especially in countries with limited resources. Studies of CMT in Africa, particularly the genetic epidemiology are notably scarce [49], likely due to the limited access to neurologists and diagnostic tools [50].

Classically, CMT is divided into two main types: type 1 (CMT1) when the disease is primarily demyelinating with the median motor nerve conduction velocities (MNCVs) < 38m/s, and type 2 (CMT2) when the disease is axonal with MNCVs > 38m/s [9]. An intermediate type is suggested when the MNCVs are between 25 m/s and 45 m/s. Other sub-classifications are based on the inheritance pattern, which can be autosomal dominant (AD), autosomal recessive (AR), or X-linked (CMTX) [27, 51].

Over 100 genes have been associated with CMT [27], and it is reported that over 90% of all genetically diagnosed cases are due to mutations in four genes: *PMP22*, *GJB1*, *MFN2*, and *MPZ* [27, 51], the 1.4 Mb duplication on the chromosome 17 (17p) accounting for over 60% of all genetically diagnosed cases of CMT in Europe and America [27]. This region contains nine genes including the peripheral myelin protein 22 gene (PMP22), that is emendable to therapeutic manipulation of CMT1A, mainly aiming at reducing PMP22 transcription [45]. Still, no curative treatment exists for these CMT subtypes, although several clinical trials are ongoing [52-54].

Interestingly, *PMP22* has been associated with CMT in only four families in Africa [55-57]. This may be due to the limited studies on CMT in the continent or, alternatively, the genetic diversity of the African population probably due to the genetic drift that happened thousands of years ago. The difference in the genetic profile of CMT amongst people of African and non-African ancestries is seen in other CMT cases, as demonstrated with higher proportion of CMT4B in Tunisia [56, 57]. This is sustained by the substantial genetic architectural differences that have been extensively documented between people of African and non-African ancestries in general [55, 58, 59]. However, the highest genetic diversity among Africans could be due to the limited number of genetic studies. Extensive studies in the African population will offer the opportunity to uncover novel genes or variant, as shown in congenital hearing impairment research [39, 60]. Hence, current therapeutic strategies under clinical trial may not be beneficial to Africans, unless the relevant genetic variants for these populations are fully identified [61].

Given the extensive genetic diversity [62], the high consanguinity and fertility rates [40, 63] [64], Africa presents a unique opportunity to discover novel

disease genetic variants [64] that could lead to the better understanding of the pathophysiological mechanisms of CMT.

In this review, we report the scarcity of research on CMT in Africa, the current clinical profiles, the specificity in the pattern of inheritance and available genetic data.

3.1.2 Methods

The present review was performed in accordance with the guidelines for transparent reporting of systematic reviews and meta-analyses (PRISMA statement 2020).

3.1.2.1 Search Strategy

We searched four databases for articles reporting CMT in Africa that fitted with the aim of this study. These databases included PubMed (https://pubmed.ncbi.nlm.nih.gov), Scopus (https://www.scpus.com), African Journals Online (https://ajol.info), and Web of Sciences (https://clarivate.com/products/web-of-science/). We used the following keywords: (“Charcot Marie Tooth disease” OR CMT OR “hereditary motor and sensory neuropathy”) AND Africa. The structured search strategy was designed to identify any published article that reports epidemiological, clinical, and genetic studies of CMT in Africa. Articles published in both English and French were included.

3.1.2.2 Selection Criteria

We included observational studies published from database inception until April 2021 that report data on the epidemiology, clinical, and genetic features of CMT in Africa. In case of duplicate studies, we selected the most recent

or more informative studies. We excluded qualitative studies, letters to the editor, reviews, and commentaries. Also, studies with unavailable full text or missing key data were removed from this review. In the case of articles reporting on patients from Africa and outside of Africa, we extracted the relevant data of interest.

3.1.2.3 Selection of Studies

All titles, abstracts, and full-text articles were independently screened by two reviewers (AY and KE). Selected articles were physically downloaded and imported into Endnote version X9.1 (Bld 12691). One author (AY) analyzed the articles before submitting to the second author (KE) to cross check the accuracy. Any disagreements between the two authors were solved by consensus and discussion.

3.1.2.4 Data Extraction Process and Assessment of Methodological Quality

One researcher (AY) extracted data from the studies included in this review. A second researcher (KE) checked the accuracy of the extraction process. Any discrepancy was resolved through discussion and consensus. The extracted data included the last name of first author, the year of publication, the country of origin of the patients, the prevalence, the incidence, the study setting, the study design, sex ratio, age ranges, sample size, number of affected individuals, type of CMT, age onset, starting symptoms, major neurological signs, inheritance pattern, technique used to identify genes, identified gene or/and variants, and the reporting journal. We also extracted data of the available histological studies.

The two investigators (KE and AY) assessed the risk of bias and the quality of included studies using the quality of genetic studies (Q-Genie) tool developed by Sohani et al. [65] for genetic studies and the risk of bias assessment tool for prevalence studies developed by Hoy et al. [66] for the other studies. Discrepancies were solved by discussion and consensus.

3.1.3 Results

Initially, 395 records were identified. We removed 364 articles after screening for titles and abstracts. The remaining 31 records were considered for full review, after which we removed six records for reasons of missing key data, letter to the editor, or review articles. In addition, three articles were found through other sources. Finally, a total of 28 articles fulfilled our selection criteria and were included in the review (Figure 1).

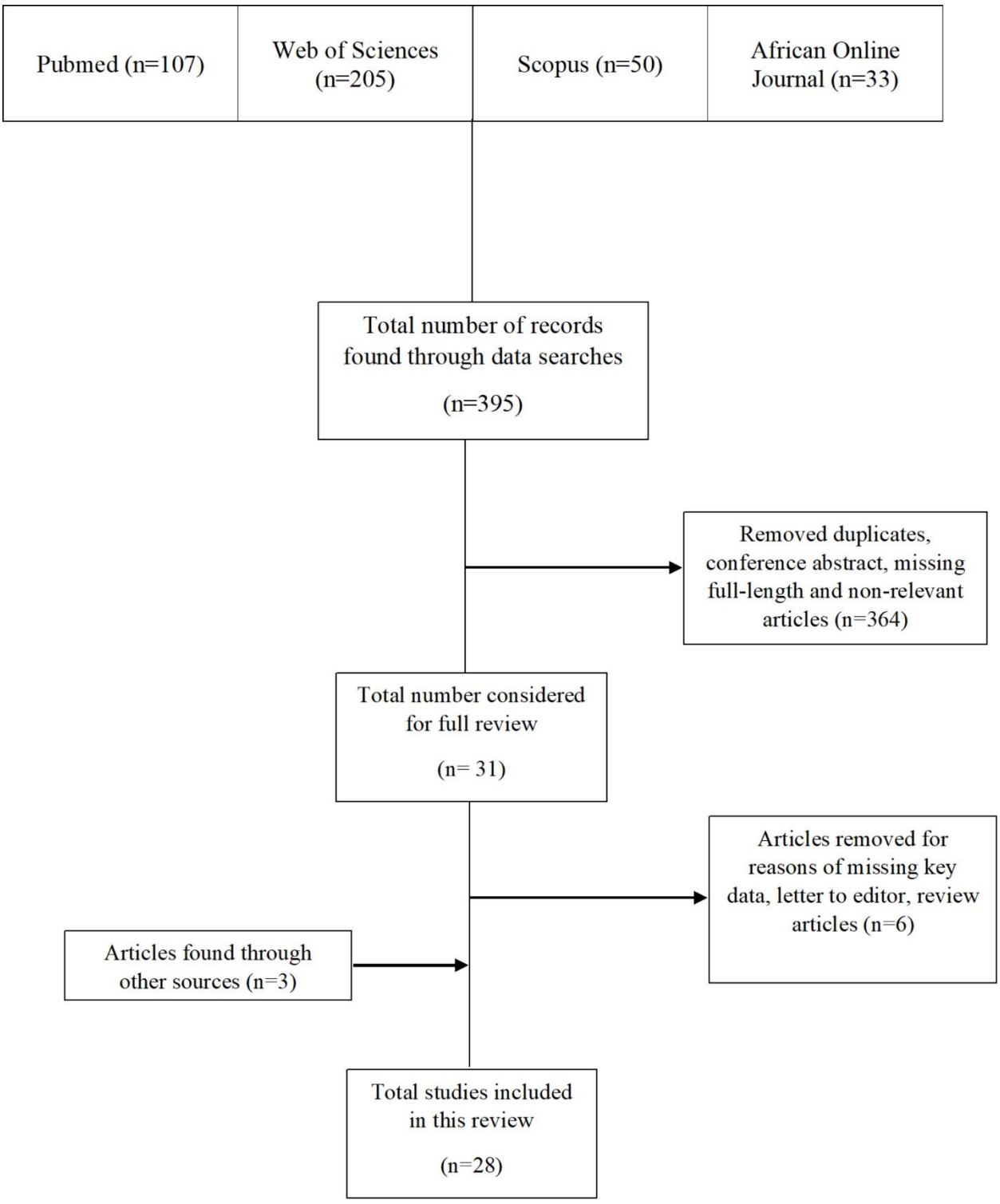


Figure 1: Flow chart of study selection

3.1.3.1 Epidemiology of CMT

Of the 28 studies included in this review, only one was population-based, namely a community-based study from Egypt that included 42,223 individuals, five patients were found with CMT phenotypes, representing an estimated prevalence of 12/100.000 [67]. One hospital-based and cross-sectional study in Nigeria reported a frequency of 0.15/100.000 among 2.1 million patients seen for neurodegenerative diseases [68]. Most of the studies were case reports, and only seven were cross-sectional studies. The total number of affected individuals per study varied from 1 to 42. The age at diagnosis ranged from 4 to 70 years (22/28 studies), and not specified in six studies. Most of the studies (78.6%; n =22/28) were reported from Northern Africa, including Algeria, Tunisia, Morocco, and Egypt, and only 21.4% (n =28) in sub-Saharan Africa (SSA) including one from Mali, two from South Africa and three from Nigeria. All the descriptive features are summarized in Table 1 and Figure 2.

Table 1: Epidemiological aspects of the studies included in this review

First author's name/Publication year	Study design	Study setting	Incidence	Prevalence	Sample size	Number of affected	Age range	Proportion of male %	Reporting Journal
Aiyesimoju, 1984[68]	Cross sectional study	Hospital	NR	0.15/100000^{F}	2.1^{M}	3	28-43	66.7 (n=3)	Neurology
LeGuern, 1996[69]	Case report	Hospital	NR	NR	2*	11	NR	54.5 (n=11)	Human Molecular Genetics
Kessali, 1997[70]	Case report	Hospital	NR	NR	25	12	11-28	58.3 (n=12)	Neurology
Meggouh, 1998[71]	Case report	Hospital	NR	NR	6	1	17	0	Journal of Medical Genetics
Bouhouche, 1999[72]	Case report	Hospital	NR	NR	17	9	15-49	89 (n=9)	American Journal of Human Genetics
Othmane, 1999[57]	Case report	Hospital	NR	NR	26	9	NR	NR	Genomics
Barhoumi, 2001[73]	Case report	Hospital	NR	NR	24	13	19-70	7.7 (n=13)	Neuromuscular Disorders
Baxter, 2001[74]	Case report	hospital	NR	NR	4*	8	NR	NR	Nature Genetics
Sandre-Giovannoli, 2002[75]	Cross sectional study	Hospital	NR	NR	23*	NR	NR	NR	American Journal of Human Genetics
Kakar, 2003[76]	Case report	Hospital	NR	NR	7	1	72	100 (n=1)	Muscle and Nerve
Azzedine, 2003[77]	Case report	Hospital	NR	NR	30	7	NR	14.3 (n=7)	American Journal of Human Genetics
Chaouch, 2003[78]	Case report	Hospital	NR	NR	4*	8	16-30	62.5 (n=8)	Neuromuscular Disorders
Birouk, 2003[63]	Case report	Hospital	NR	NR	17	4	15-20	0	Arch Neurology
Tazir, 2004[79]	Cross sectional study	Hospital	NR	NR	62	21	12-45	62 (n=21)	Brain
Azzedine, 2006[80]	Cross sectional study	NS	NR	NR	4*	NS	NS	NS	Neurology
Bösenberg, 2006[81]	Case report	Hospital	NR	NR	2	2	14-19	100 (n=2)	Southern African Journal of Anaesthesia & Analgesia

Onwuekwe, 2007[82]	Case report	Hospital	NR	NR	1	1	31	100 (n=1)	Journal of college of Medicine
Bouhouche, 2007[83]	Case report	Hospital	NR	NR	11	6	4-19	50 (n=6)	Canadian Journal of Neurological Sciences
Delague, 2007[84]	Case report	Hospital	NR	NR	7	3	NR	100 (n=3)	American Journal of Human Genetics
Bouhouche, 2007[85]	Cross sectional study	Hospital	NR	NR	95	31	4-49	42 (n=31)	Brain
Hamadouche, 2008[86]	Cross sectional study	Hospital	NR	NR	25*	42	NR	48 (n=42)	Annals of Human Genetics
Nouioua, 2011[87]	Cross sectional study	Hospital	NR	NR	2*	7	9-22	85.7 (n=7)	Neuromuscular Disorders
Baudot, 2012[88]	Case report	Hospital	NR	NR	1*	1	16	100 (n=1)	Journal of the Peripheral Nervous System
Kandil, 2012[67]	Cross sectional study	Community	NR	1.2/10.000	42.223	5	NR	80 (n=4)	Neurological Research
Boubaker, 2013[89]	Case report	Hospital	NR	NR	8	3	6-22	33.3 (n=3)	Annals of Human Genetics
Mathis, 2014[90]	Case report	Hospital	NR	NR	1	1	10	0	Neuromuscular Disorders
Yalcouyé, 2019[49]	Case report	Hospital	NR	NR	4	3	37 (19-58)	33 (n=1)	Molecular Genetics and Genomic Medicine
Manyeruke, 2020[91]	Case report	Hospital	NR	NR	1	1	11	0	South African Ophthalmology Journal

F: hospital frequency; M: million, NR: not reported, NS: not specified; *: number of families

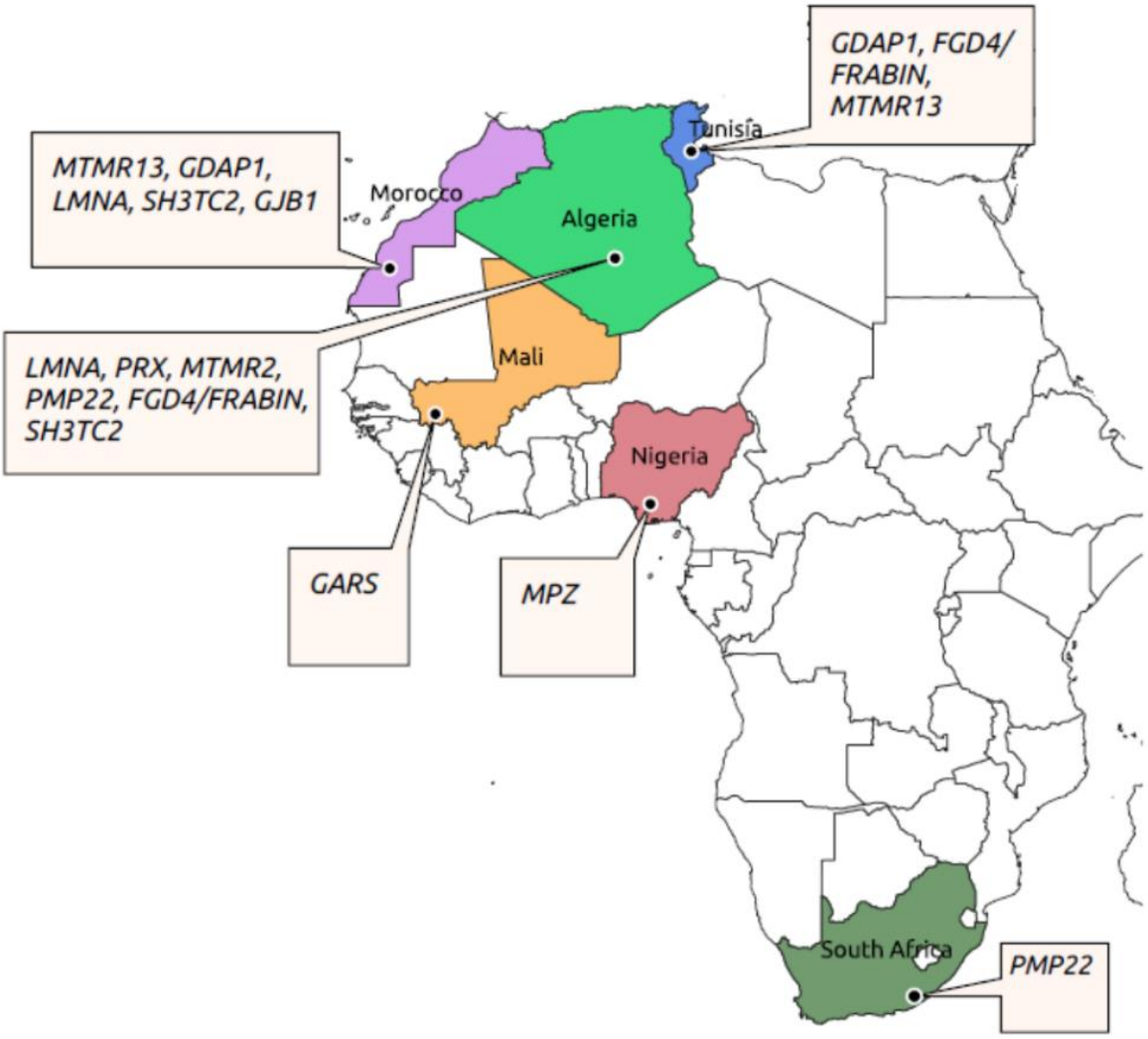


Figure 2: Genes reported in the respective African countries

3.1.3.2 Clinical expression

A total of 185 patients were described in the studies reviewed here, and the sex ratio was 1.2 (99 males vs 86 females). The demyelinating type (CMT1) was reported in 58.3 % of the studies [68, 74, 76, 77, 80, 87-90], followed by the axonal type (CMT2) in 37.5% (n =28) [63, 72, 73, 79, 83, 85] and the intermediate form in 4.2% [43]. The disease started mostly in the first two

decades of life, but cases with later onset were also reported [76]. Almost all studies reported muscle weakness predominantly in the lower limbs as the starting symptoms, and only few cases reported sensory impairment as presenting symptoms [49, 63, 77, 79, 83, 89, 90]. The major neurological signs included muscle weakness and wasting, predominantly in distal limbs but proximal involvement was reported in some studies [72, 78]. In addition, other neurological signs such steppage gait, skeletal deformities (*pes cavus* and *pes planus*, hammer toes, claw hands, scoliosis, and kyphosis) (Figure 3C, D, E, F), and sensory impairment were reported [89]. Reflexes were reduced to absent in most of the cases. However, a case with upper motor neuron involvement with brisk reflexes was reported [73]. A severe case associated with marked stridor during inspiration causing dyspnea and abdominal respiration, and a vocal cord paralysis was reported in a family with three affected siblings [87] (Figure 3A, B, C).

A rare case of autosomal recessive demyelinating form of CMT associated with early-onset glaucoma was reported from Tunisia [31]. Moreover, an unusual phenotype associating CMT1A with macular oedema was first reported in a South African girl [91]. Only one sporadic case of intermediate motor NCV in a female from Morocco was reported with a mutation in the *GJB1* gene [71]. The phenotype was more severe in the recessive cases with early onset, and patients were wheelchair-dependent by the time of diagnosis [63, 72, 74, 87, 89]. Similar to other reports, intra-and interfamilial phenotype variability was seen in the African patients as the cases reported from Algeria [71, 90]. Also, a case of CMT associated with hearing impairment was reported in an Algerian family [70]. The disease course was slowly progressive in most of the cases, but a rapid progressive case was

found in an Algerian family [70]. CMT1 was the most reported sub-type (Table II). Nerve biopsy was performed in a select of cases and showed the classic aspect of “onion bulbs” [70, 78, 85, 89]. In addition, some axonal cases were reported with an important loss of large, myelinated fibers and a few clusters of regeneration [63, 92]. The main characteristics of the clinical profile are highlighted in Table 2.

Table 2: Main clinical and genetic characteristics of studies included in this review

Studies	Age of onset	Starting symptoms	Major neurological signs	Type of CMT	Genes	Variants	Methods of diagnosis
Aiyesimoju, 1984[68]	NR	NR	NR	NA	NA	NA	NP
LeGuern, 1996[69]	NR	NR	NR	ARCMT1	NI	NI	Homozygosity mapping
Kessali, 1997[70]	1st decade	Foot and spine deformities	Distal muscle weakness in UL and LL, areflexia, foot deformities, kyphoscoliosis, hypoacusis and facial weakness	ARCMT	NI	NI	Linkage analysis
Meggouh, 1998[71]	2nd decade	Distal LL muscle weakness and wasting	Distal muscle weakness predominantly in LL, tendon areflexia, pes cavus and kyphoscoliosis	CMTX	*Cx32/GJB1*	del499G	Sanger sequencing
Bouhouche, 1999[72]	2nd decade	NR	muscles weakness and wasting of the distal limbs, and areflexia predominantly in the lower limbs. Involvement of the proximal muscles in few patients. Pes cavus and severe kyphoscoliosis.	ARCMT2	NI	NI	Linkage analysis, physical mapping and direct sequencing
Othmane, 1999[57]	1st/2nd decade	NR	Atrophy and weakness of intrinsic foot muscles, peronei, and anterior tibial muscles. Pes cavus and hammer toes	CMT4B	NI	NI	Homozygosity mapping and linkage analysis
Barhoumi, 2001[73]	1st decade	Walking difficulty	Severe distal muscle wasting, and atrophy of legs and of small muscles of hands. Steppage gait with bilateral foot drop, brisk tendon reflexes in UL and knee, and absent ankle reflexes. Distal sensory loss in LL including sense of touch, pain, proprioception and pallesthesia	ARCMT2	NI	NI	Homozygosity mapping and linkage analysis
Baxter, 2001[74]	1st decade	Muscle weakness	Weakness and atrophy of the feet and hands (clawhands). wheelchair-dependent and/or develop kyphosis. Mild sensory loss, proprioception and vibration senses	CMT4	*GDAP1*	c.G92A; p.W31X c.G482A; p.R161H	Direct sequencing

Sandre-Giovannoli, 2002[75]	1st decade	Muscle weakness	NR	ARCMT2	*LMNA*	c.C892T, p.R298C	Direct sequencing
Kakar, 2003[76]	5th decade	Bilateral numbness and tingling in feet	Severe atrophy and weakness of the distal arm and legs. Tendon areflexic, with flexor plantar responses. There was sensory loss of all modalities in a glove and stocking distribution. Gait was abnormal with bilateral foot drop. He had pes cavus	CMT1B	*MPZ*	c.C234G, p.S78W	Direct sequencing
Azzedine, 2003[77]	1st/2nd decade	Muscle weakness	Motor and sensory loss, areflexia, foot deformities and scoliosis	CMT4B2	*MTMR13*	c.C2875T, p.Gln956Stop* c.C3586T; p.Arg1196Stop*	Sanger sequencing
Chaouch, 2003[78]	1st/2nd/3rd decade	Muscle weakness	Weakness and amyotrophy of proximal muscles of pelvic girdle. Variable distal sensory disturbances with a glove and stock distribution	ARCMT2	*LMNA*	c.C892T; p.R298C	Sanger sequencing
Birouk, 2003[63]	1st decade	Foot deformities and muscle weakness	Distal muscle weakness and wasting of legs, predominantly in peroneal muscles, with severe foot deformities of the pes equinovarus type. Total areflexia, and loss of proprioception in the lower limbs	ARCMT2	*GDAP1*	S194X	Sanger sequencing
Tazir, 2004[79]	1st/2nd/3rd decade	Difficulty to running and walking	Distal and proximal muscle weakness, sensory loss, amyotrophy and areflexia. Foot deformities with pes cavus, scoliosis	ARCMT2	*LMNA*	c.C892T, p.R298C	Sanger sequencing
Azzedine, 2006[80]	1st decade	Scoliosis and kyphoscoliosis	Scoliosis or kyphoscoliosis and foot deformities	CMT4C	*SH3TC2*	del GCTGCTCGGAG; A74_77 indel fsX128* IVS10–1G/A* c. 2190delC ; p.E731fsX750* c.C 2710T; p.R904X* c. C2860T; p.R954X*	Direct sequencing
Bösenberg, 2006[81]	1st/2nd decade	Running difficulty and peroneal spasm	wasting of the thenar, eminence and the interossei of both hands and feet. Deep tendon reflexes were absent, slight sensory loss in his hands and feet. Feet deformities	CMT1A	*PMP22*	*PMP22* duplication	NR

Onwuekwe, 2007[82]	2nd decade	Paraesthesia	Distal quadriparesis, spontaneous fasciculations, hyporeflexia and loss of proprioception	CMT1	NA	NA	NP
Bouhouche, 2007[85]	1st decade	Hypotonia at birth and walking delay	Predominantly distal motor deficit and atrophy of both UL and LL. Atrophy and weakness of proximal muscles, distal sensory impairment involving particularly proprioception in the LL	CMT4A	*GDAP1*	c.C233T; p. P78L	Linkage analysis and direct sequencing
Delague, 2007[84]	1st decade	Delayed walking	Muscle weakness and amyotrophy in the distal extremities, marked feet abnormalities (pes cavus), absent tendon reflexes in the four limbs, ataxia, and a waddling gait	CMT4H	*FGD4/FRABIN*	c.T893C, p.Met298Thr	Direct sequencing
Bouhouche, 2007[85]	1st decade	hypotonia at birth and delayed first motor acquisition	Distal muscle weakness, foot deformities and claw fingers, areflexia, sensory loss, wheelchair bound	CMT2B1	*LMNA*	c.892C>T; p.Arg298Cys*	Microsatellite markers and direct sequencing
				CMT4A	*GDAP1*	c.C581G; p.S194X*	
Hamadouche, 2008[86]	1st/2nd/3rd decade	NR	NR	ARCMT2	*LMNA*	c.892C>T; p.Arg298Cys	Sanger sequencing
Nouioua, 2011[87]	1st/2nd decade	Spine deformities and gait instability	Predominantly motor neuropathy with a steppage gait and distal limb weakness and wasting, claw hands and sensory loss, stridor and breathing difficulties	CMT4B1	*MTMR2*	c.331dupA; p.Arg111LysfsX24*	Sanger sequencing
				CMT4F	*PRX*	c.1090C>T ; p.Arg364X*	
Baudot, 2012[88]	NR	NR	NR	CMT4H	*FGD4/FRABIN*	c.1325G>A; p.Arg442His	Sanger Sequencing
Kandil, 2012[67]	NR	NR	NR	NR	NA	NA	NP
Boubaker, 2013[89]	1st decade	Gait disturbance	Amyotrophy and muscle weakness in the UL and LL. Muscle tone was low and deep tendon reflexes were absent. Walking on her tip toes, pes cavus and mild scoliosis	CMT4H	*FGD4/FRABIN*	c.514_515insG; p.Ala172Glyfs∗27	Sanger sequencing
Mathis, 2014[90]	1st decade	Walking difficulties	Weak deep tendon reflexes in all four limbs and kyphoscoliosis	CMT1A	*PMP22*	*PMP22* duplication	MLPA and direct sequencing
Yalcouyé, 2019[49]	2nd decade	UL muscle weakness	Distal muscle and sensory loss, muscle weakness and steppage gait	CMT2D	*GARS*	c.794C>A; Ser265Tyr	NGS (CMT gene panel)
Manyeruke, 2020[91]	1st decade	NR	NR	CMT1A	*PMP22*	*PMP22* duplication	NR

NA: not applicable, NI: not identified, NP: not performed, NR: not reported, UL: upper limbs, LL: lower limbers, *: different families, MLPA: Multiplex Ligation Probe Amplification; NGS: next generation sequencing

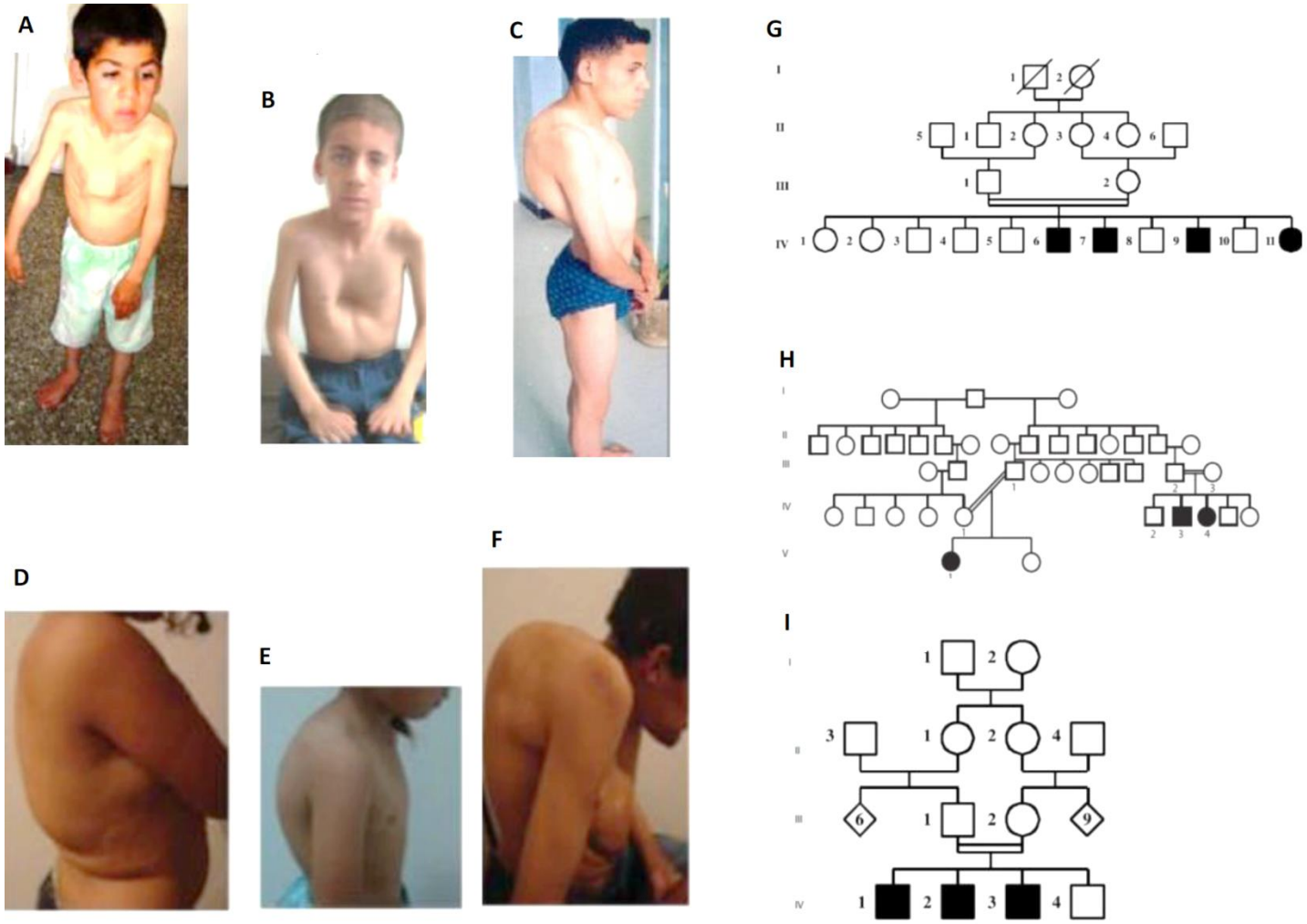


Figure 3: Illustration of some phenotypes of CMTs. (A, B, C) Images of CMT patients with mutation in MTMR2 genes showing the chest deformities, dyspnoea, and severe scoliosis. D, E, F) Images of CMT patients with mutation in FGD4/FRABIN gene showing a severe kyphoscoliosis. G, H, I) Pedigrees of some families showing autosomal recessive transmission manner with the consanguinity. (Images extracted from the articles by Yalcouye et al) [87, 89].

3.1.3.3 Pedigree analysis

The pattern of inheritance was autosomal recessive in most cases (91.2%, n=97) (Figure 3G, H, I) while autosomal dominant represented 4.9%, and X-linked and unknown patterns were seen in 3.9 % each. Consanguinity was reported in 62% (n=66) of families [87, 89].

3.1.3.4 Genetic analysis

Several techniques were used to identify the causative genes associated with CMT over the time (Table 2). While the recent studies use NGS methods and targeted CMT genes panel, in the past decades, Multiplex Ligation Probe Amplification (MLPA), homozygosity mapping, and direct sequencing were used to identify the causative genes in CMT. The latter methods were mostly used in the studies reported here and allowed the identification of four genetic loci and 22 variants in African families, representing more than half of the cases (Table II). However, the genes and variants in the mapped loci were not identified. In a consanguineous family from Morocco, Othmane et al mapped the first CMT-associated (CMT4B) locus (11p15) in Africa [57]. Of note, none of the studies reviewed here used WES or whole genome sequencing (WGS) to diagnose CMT cases. To date, only eleven genes including *LMNA*, *GDAP1*, *PMP22*, *MTMR2*, *MTMR13*, *Cx32/GJB1*, *PRX*, *MPZ*, *FGD4/FRABIN*, *SH3TC2* and *GARS* have been associated with CMT in Africa. The most common genes were *LMNA*, *GDAP1* and *SH3TC2,* representing more than 80% (n=65) of the molecularly diagnosed CMT cases in Africa. The specific genes and the respective countries are shown in Figure 2. Variants in *PMP22* gene were reported in four families only [81, 90, 91]. Interestingly, a variant (c.C892T, p.Arg298Cys) in the *LMNA* gene was found to have a founder effect in North-Western Africa (Algeria and

Morocco) [86]. In other studies, most of the variants reported were novel [49, 71, 74, 76, 87, 88] and no genes have been identified in a few studies [57, 70-72, 82]. This is not surprising since most of these studies were conducted more than two decades ago, and the techniques used were less efficient compared to NGS. NGS with a CMT gene panel testing was performed in only one study [49]. All the genes and variants reported are summarised in the Table 2.

3.1.4 Discussion

To the best of our knowledge, this review is the most comprehensive and complete report on the epidemiological, clinical, and genetic features of CMT in Africa. It revealed the lack of data from most African countries, especially from SSA. The review has also allowed us to identify the genetic profile of CMT in Africa and suggests a difference from what is reported to date in high-income countries. In fact, a lower contribution of *PMP22*-associated variants in Africa and a higher rate of novel and founder variants in known genes likely related to higher consanguinity rates were seen in Africa. In contrast to the high-income countries, the prevalence or incidence of CMT in Africa is still largely unknown. Two clinical studies published in the 1980s have reported prevalence rates of 8/100,000 and 10/100,000 in Libya [93] and in Nigeria [94], respectively. More recently, an estimated prevalence of 12/100.000 was reported in Egypt [66]. Yet, most of the studies included in this review were case reports on familial or isolated cases, illustrating a widely variable regional epidemiological description of CMT in Africa [66, 92].

Nevertheless, the clinical description in the patients included in this review wasn't different to that reported in other populations with different ethnic background. In most patients, symptoms appear during the first or second

decade of life with an insidious onset and a slowly progressive weakness that starts in lower extremities, later involving upper extremities [95-97]. Diverse phenotypes were reported in Africa including asymptomatic, mild, moderate, and severe forms of CMT [49, 63, 85, 87, 89]. These data confirm its clinical heterogeneity reported in other populations [27, 97]. However, the clinical presentation seems more severe in the African families [72] than reported elsewhere [98]. This could be due to other genetic modifiers, environmental factors, the differences in the care, or the high frequency of recessive cases, which is known to be more severe. The distribution of muscle weakness is mainly in the distal part but can also be proximal as reported in some studies [78, 99]. In addition to muscle weakness, other motor signs include decreased or absent tendon reflexes, amyotrophy and walking difficulties with steppage gait. Similar to other reports, some rare cases can cause respiratory failure or breathing difficulties like the case reported by Nouhouia et al [87, 100, 101]. Sensory impairment is typically associated with the phenotype, affecting generally the distal part in “gloves and socks” pattern [57, 60]. The disease course is slowly progressive in most cases, but in exceptional cases, it can progress rapidly [70]. In this study, most families segregated CMT disease in its recessive form associated with high consanguinity rate. This is different from the dominant manner inheritance pattern that is the most common reported worldwide [79], but underreported cases in Africa could explain this difference.

Histological studies have a role in identifying underlying genetic aetiology in sporadic cases, and it helps distinguishing CMT from acquired peripheral neuropathies [99]. Nerve biopsy may also support a functional association when the genetic tests detect "variants of uncertain significance" or a novel

variant [99]. However, these were not performed as standard procedures in the reviewed studies [49, 68, 72, 88, 90]. Nerve conduction studies are an important step in the algorithm of CMT diagnosis, and allow the classification of different CMT types [99]. This testing was performed in almost all the studies, though not all patients in each study were screened [68, 82].

The present review highlighted some regional specificity with *LMNA* and *GDAP1* genes mostly found in North Africa where most of the studies were reported from, a region known for its high consanguinity rates as confirmed in this review [102]. Consanguinity rate was also high in Mali, Morocco, Algeria, and Egypt [40, 63, 64, 72, 78, 85], a population profile that will favour genes discovery in future. Indeed, the data also showed a limited use of NGS to investigate CMT in Africa, and no study used WES or WGS. In fact, WES/WGS are highly likely to identify novel genes and variants in known genes, particularly in the understudied and highly genetically diverse populations of Africa. Despite the limited number of reports of CMT in Africa when compared to those from Europe and North America, this study could indicate the genetic heterogeneity of CMT disease in Africa in line with the global knowledge [27, 51]. All Mendelian models of inheritance were seen, but the dominant pattern is the most commonly reported worldwide [79] while recessive cases were shown to be predominant in the cases described in Africa; likely associated with high consanguinity rate of most cases reported from North Africa [86]. Most demyelinating CMT types result from mutations in genes expressed by SCs, whereas axonal types result from mutations expressed by neurons and their axons [103]. Recent studies reported that the increasing insights into the molecular-genetic mechanisms have revealed potential therapeutic targets [103]. These will enable the

development of novel therapeutics for genetic neuropathies that remain, in their majority, with no effective treatment [9, 27].

The techniques used in Africa in the past were mostly MLPA, targeted sequencing and homozygosity mapping. These have allowed the identification of genetic loci and known and novel variants for CMT cases. Despite the rapid evolvement of the genetic diagnosis of CMT in recent years with the advent of the NGS technology [9, 27, 61, 104], only one study in Africa has used it to diagnose a CMT case [49]. NGS technology allows multiple parallel sequencing of either targeted genes, only the protein coding sequences (WES) or the whole genome (WGS) [27, 51, 104]. The challenge is how best to use these in clinical practice. To answer this question, Claudia et al. performed WES on individuals with CMT, and reported a diagnostic rate of 45% [105]. Recent studies have confirmed the efficiency of NGS in diagnosing CMT cases [27]. In a cohort of pre-excluded *PMP22*dup/del from Japan, authors identified the causative genes in 30% of the cases, and the most common genes were *GJB1*, *MFN2* and *MPZ* [97]. The overall diagnosis rate is higher in demyelinating CMT compared to the axonal type [50, 97].

The molecular profile of CMT is sparse but globally *PMP22*, *GJB1*, *MFN2*, *MPZ* genes explain at least 90% of CMT cases [27, 46, 51]. This epidemiology described above may not necessarily be extrapolated to other populations with different ethnic background, most notably those from the African continent which remains understudied [61, 106] and is under-represented in large population genetic databases [107]. In fact, the genetic epidemiology profile from the studies reviewed here does not reflect what was reported in other populations. While CMT1A represents more than 60% of all CMTs [27], to date, the commonest CMT gene (*PMP22*) has been

reported in only four families in Africa [81, 90, 91]. Moreover, an autosomal recessive CMT case (CMT4B) was first mapped in a family from Africa, before the gene was identified in subsequent studies done abroad [57]; suggesting that the African population harbours specific gene variants for CMT but the limited access to diagnosis tools may delay the molecular diagnosis confirmation.

The scarcity of the most common CMT genes in Africa might be associated with the limited number of studies as it is expected that prevalence of CMT1A (and Hereditary Neuropathy with pressure palsies) could be similar in all populations worldwide. It is possible that CMT1A might be under reported because the phenotype is already well known, and only the most severely affected patients come to the medical attention in many African regions, owing to the limited access to diagnostic tools, and to scarce neurology specialists. It is also possible that the findings of the current review may be due to the genetic diversity of African populations, the population structure, consanguinity rates, or the genetic drift. Therefore, this stresses the need for more studies on the genetics of CMT in Africa using NGS, with the potential of uncovering novel genes or variants important for the function of the peripheral nerve system. CMT is a disabling condition that does not have a cure, but the advances in the understanding of its pathophysiology have advanced research in the identification of therapeutic targets in human and animal models [52, 97, 108, 109]. The extension of such studies to Africa could be especially beneficial and equitable.

3.1.5 Strengths and limitations

To the best of our knowledge, this review provides the most comprehensive and complete data on CMT in Africa. It summarized the available data on the

epidemiological, clinical, and genetic profiles of CMT in Africa. It identified the enormous gaps in the knowledge of CMT in Africa compared to developed countries and highlighted the necessity to undertake large scale genetic studies on CMT in Africa to further our understanding to its global epidemiology and perhaps identify other therapeutic perspectives. Therefore, this review may be the first step for future perspectives in the research of CMT in Africa. However, this study has some limitations. First, the absence of nationwide studies in Africa, and most of the studies included herein were case reports which are obviously limited with regards to epidemiological data. Second, the keywords we used for searches may have missed some articles that do not include those words. Third, the language restriction to English and French may have also missed some articles reported in other languages. Fourth, many African researchers do not have access to indexed journals and may have published in journals that our selection criteria do not catch.

3.1.6 Conclusion

This study reveals that CMT is not rare, and likely underreported in Africa and describes the current clinical and genetic profile. Large and multicentric cohort studies in Africa would not only inform in the genetic epidemiology of CMT in this region but could also lead to new discoveries important to the global research effort for therapeutic perspectives. The increasing access to NGS technologies offers to African scientists a unique opportunity to fully describe relevant variants in known genes, and to discover novel CMT-associated genes that may improve our understanding and care of this condition in Africa.

Conflict of interest: The authors declare no conflicts of interest.

Author contributions

Abdoulaye Yalcouyé: Developed the methodology; database search; analysed and interpreted the data, wrote the first draft; read and agreed to the published version of the manuscript. **Kevin Esoh**: contributed to developing the methodology and critically revised successive drafts of the manuscript; read and agreed to the published version of the manuscript, **Guida Landouré**: Conceived the study; critically revised successive drafts of the manuscript; supervised the project; read and agreed to the published version of the manuscript. **Ambroise Wonkam**: Conceived the study; critically revised successive drafts of the manuscript; supervised the project; read and agreed to the published version of the manuscript.

Funding information: The study was funded by the NIH/NINDS grant number U01HG007044 to G.L.; NIH, USA, grant number U01HG009716 to A.W.; and the African Academy of Science/Wellcome Trust, grant number H3A/18/001to A.W. The funders had no role in data synthesis and analysis, and decision to publish this manuscript.

Data availability statement: The data that support the findings of this study are available from the corresponding author, [AW], upon reasonable request.

Chapter 4: Results - Target gene testing to solve CMT cases

Synopsis: This chapter presents the use of CMT- gene panel testing to solve CMT cases. The first study reports investigated the clinical and genetic profile of CMTX1 cases while the second focused on the clinical and genetic description of CMT2D in Mali.

4.1 **Abdoulaye Yalcouyé**, Seybou H. Diallo, Lassana Cissé, Mamadou Karembé, Salimata Diallo, Thomas Coulibaly, Salimata Diarra, Dramane Coulibaly, Mohamed Keita, Cheick O. Guinto, Kenneth H. Fischbeck, Ambroise Wonkam, Guida Landouré, The H3Africa Consortium. GJB1 variants in Charcot-Marie-Tooth disease X-linked type1 in Mali. J Peripher Nerv Syst. 2022 Jun;27(2):113-119. doi: 10.1111/jns.12486.

Nature of publication: Case series article

Journal Publisher: Journal of the peripheral nervous system: JPNS- ISSN: 1085-9489, peer review journal

Candidate contribution: Performed clinical evaluation, genetic analysis, drafting the first version of the manuscript and subsequent revisions

Co-authors contribution:

GL: conceptualization, supervision and funding acquisition

SHD: Contributed to clinical evaluation

MK: Contributed to clinical evaluation

SD: Contributed to clinical evaluation

TC: Contributed to clinical evaluation

LC: Contributed to clinical evaluation and EMG

SD: Contributed to clinical evaluation

DC: Contributed to clinical evaluation

COG: Conceptualization and supervision of the study

KF and **AW** and **GL**: Critically revised the manuscript

All agreed the final version of the manuscript

GJB1 variants in Charcot-Marie-Tooth disease X-linked type 1 in Mali

Abdoulaye Yalcouyé[1,2], Seybou H Diallo[2,3], Lassana Cissé[4], Mamadou Karembé[4], Salimata Diallo[3], Thomas Coulibaly[2,4], Salimata Diarra[2,5], Dramane Coulibaly[6], Mohamed Keita[2,7], Cheick O Guinto[2,4], Kenneth Fischbeck[5], Ambroise Wonkam[1], Guida Landouré[2,4,5], The H3Africa Consortium

Short running title: *GJB1* variants in Mali

[1]Division of Human Genetics, Department of Pathology, University of Cape Town, Cape Town, South Africa,

[2]Faculté de Médecine et d'Odontostomatologie, USTTB, Bamako, Mali,
[3]Service de Neurologie, Centre Hospitalier Universitaire Gabriel Touré, Bamako, Mali

[4]Service de Neurologie, Centre Hospitalier Universitaire du Point "G", Bamako, Mali

[5]Neurogenetics Branch, National Institutes of Neurological Disorders and Stroke, Bethesda, MD

[6]Service de Médecine, Centre Hospitalier Universitaire Mère-Enfant le "Luxembourg", Bamako, Mali

[7]Service d'ORL, Centre Hospitalier Universitaire Gabriel Touré, Bamako, Mali

Corresponding author: Guida Landouré, Faculté de Médecine et d'Odontostomatologie, USTTB, Bamako, Mali, BP: 1805, Tel: +223 76 36 34 68, Fax: +223 20 22 96 58, email: guida@icermali.org

Abstract

Background and aims: X-linked Charcot-Marie-Tooth type 1 (CMTX1) disease is one of the most common subtypes of inherited neuropathies and is caused by mutations in the *GJB1* gene. To date, more than 400 mutations have been reported in *GJB1* worldwide but none in sub-Saharan Africa (SSA). We aimed to clinically characterize patients with CMTX1 and identify the genetic defects.

Methods: All patients were examined thoroughly, and Nerve Conduction Studies (NCS) were done. Electroencephalography (EEG) and pure tone audiometry (PTA) were also done in select individuals having additional symptoms. DNA was extracted for CMT gene panel testing (50 genes + mtDNA and PMP22 duplication), and putative variants were screened in available relatives.

Results: The predominant starting symptom was tingling, and the chief complaint was gait difficulty. Neurological examination found a distal muscle weakness and atrophy, and sensory loss, skeletal deformities, decreased or absent reflexes and steppage gait. The inheritance pattern was consistent with dominant X-linked. NCS showed no response in most of the tested nerves in lower limbs, and normal or reduced amplitudes in upper limbs. A severe sensorineural hearing impairment and a focal epileptic seizure were observed in one patient each. A high intra and inter-familial clinical variability was observed. Genetic testing found three pathogenic missense variants in *GJB1*, one in each of the families (Val91Met; Arg15Trp and Phe235Cys).

Interpretation: This is the first report of genetically confirmed cases of CMTX1 in SSA, and confirms its clinical and genetic heterogeneity.

Keywords: Inherited neuropathies, CMTX1, GJB1, Mali, Africa.

4.1.1 Introduction

Charcot-Marie-tooth disease (CMT, OMIM: 118220), also known as progressive peroneal atrophy, is the most common inherited peripheral neuropathy with a global prevalence of 1/2500 [110]. X-linked Charcot-Marie-Tooth type 1 (CMTX1, OMIM: 302800) is dominantly inherited, and is one of the most common subtypes of CMT caused by mutations in the *GJB1* gene (connexin 32, OMIM: 304040), representing up to 6.7 to 10% of all types of CMT [30, 46, 111]. However, its genetic epidemiology varies from one region to another. CMTX1 is a highly disabling condition with an important clinical and genetic variability. Its clinical symptoms include distal muscle weakness and atrophy and sensory loss, skeletal deformities, decreased or absent tendon reflexes, and steppage gait [112]. Consistent with X-linked inheritance, male patients are more severely affected when compared to the female patients. They often show “split hand syndrome” in which atrophy and weakness are more prominent in specific muscle such abductor pollicis brevis [113]. Electrophysiological patterns differ in males when compared to females as they usually show intermediate motor NCVs (25–35 m/s) while females have mild slow motor NCVs (>35 ms) [30]. Both myelin and axon can be involved in the pathogenic process in CMTX1. Hence, CMTX1 patients may present both demyelination and axonal involvement, although previous electrophysiological and pathological studies suggested a more prominent axonal involvement [28, 114]. CMTX1 is an intermediate neuropathy where patients have typically intermediate NCV [115]. More than 400 mutations in GJB1 gene have been reported [116]. Despite being a well-known condition elsewhere, no case has been reported

in SSA, to date. We report here the first cases of CMTX1 in this region (Mali) confirmed at the molecular level.

4.1.2 Material and Methods

This study was conducted in the full compliance of the declaration of Helsinki. The ethical approval (N°2018/182/CE/FMOS) and (HREC REF 691/2020) have been obtained from the Ethics Committee of the Faculty of Medicine and Dentistry of the University of Sciences, Techniques and Technologies of Bamako, Mali, and the University of Cape Town, Cape Town, South Africa. Informed consent and assent were obtained from all participants involved in this study before their enrolment including approval to publish photographs.

All patients were carefully evaluated by a multidisciplinary team including medical geneticist, neurologists, ENT specialists and ophthalmologists. Blood chemistries including peripheral blood ions, vitamin B12 and blood glucose levels were checked to exclude common causes of acquired polyneuropathies. NCS were done at least in the proband of each family. DNA was extracted from peripheral blood in all available participants for genetic analysis. *PMP22* duplication/deletion analysis was first done and a Next-Generation Sequencing (NGS) CMT gene panel testing including 50 genes and mtDNA (Medical Neurogenetics, Atlanta, GA) was performed in the probands. The putative variants were verified in SNP databases to exclude rare variants and were screened in all available family members for segregation analysis. In silico analysis using several prediction tools was done to check for deleteriousness, following the ACMG criteria [117].

4.1.3 Results

A total of 15 patients (six males and nine females with a sex ratio of 0.67) from three unrelated families were found to be affected. The disease segregation in the family pedigrees were consistent with a dominant inheritance pattern (Figure 4A, B and C). The mean age at onset was 23.6 years, ranging from 8 to 51 years, and the mean age at diagnosis was 27.3 years, ranging from 15 to 63 years. However, when taken separately, the age of onset in males and females are 17.3 years and 28.3 years, respectively (ranging from 8 to 24 years, and from 16 to 51 years), and the mean age at diagnosis was 22.2 and 42.1 years, respectively (ranging from 15 to 33 years, and from 17 to 63 years). Symptoms started with tingling in distal legs, and later they presented with muscle weakness, gait difficulty and cramps in legs that worsened over the time. In addition, two patients reported other symptoms, one recurring epileptic seizures and the other a hearing impairment. Neurological examination found a high phenotype variability within and between the families. A distal muscle weakness and atrophy and sensory loss ranging from mild to severe and decreased to absent reflexes throughout were seen in all patients. Seven patients presented with a steppage gait. Skeletal deformities such as *pes cavus*, hammer toes and claw hands were also observed (Figure 5A). All symptoms were bilateral and symmetrical with no proximal involvement. In addition, a severe sensorineural hearing impairment in the proband of Family 2 and a focal epileptic seizure in the proband of Family 3 were confirmed by a pure tone audiometry and an EEG, respectively. As expected, symptoms were more severe among the men when compared to the women with children being more affected than their mothers or older sisters (Figure 5A and B). However, in one case, a diabetic female patient has presented a faster disease progression than seen in the other female patients. The NCS performed in

select patients including probands showed no response in almost all of the tested nerves in the lower limbs and normal or reduced amplitudes in the upper limbs (details are in Table 3).

Genetic testing in the probands identified a hemizygous missense variant in the GJB1 gene at position c.271G>A (p.Val91Met) in Family 1 (Figure 4D), a heterozygote missense variant c.704T>G (p.Phe235Cys) in Family 2 (Figure 4E) and a hemizygous missense variant c.43C>T (p.Arg15Trp) in Family 3 (Figure 4F). All the variants were previously described in other populations, and segregate with the disease status in the respective families. The Val91, Phe235 and Arg15 residues are well conserved across a wide range of species and are located in highly conserved domains of the protein (Figure 4G, H and I). These variants are predicted to be deleterious by several in silico analysis tools (Table S1). All the phenotype and genotype characteristics are summarized in Table 3.

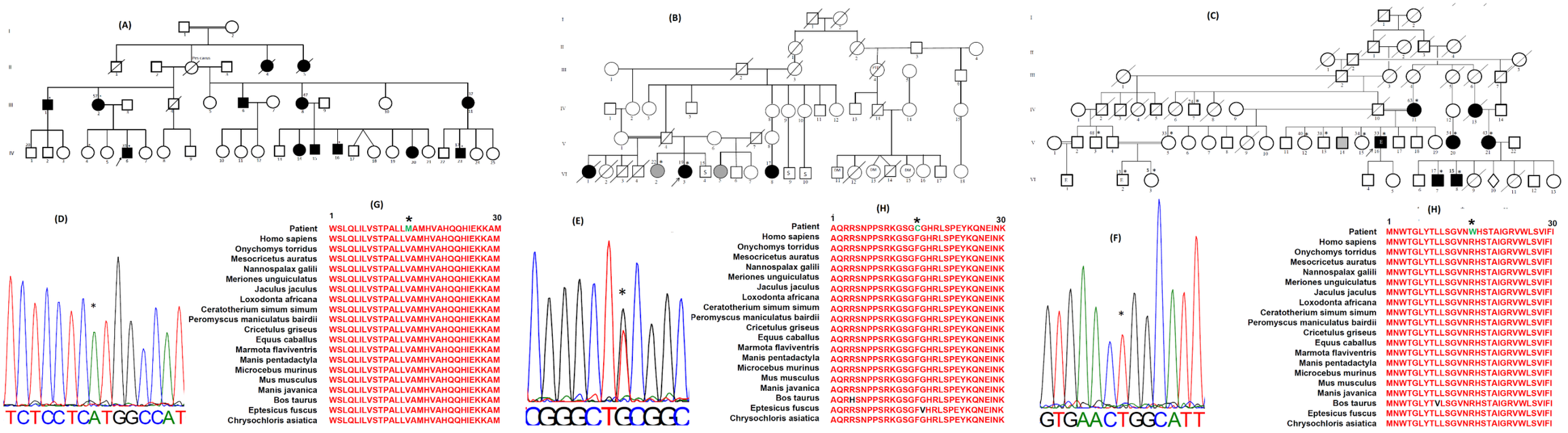


Figure 4: Pedigrees of families with CMTX1 and genetic data

A: Pedigree of Family 1. Asterisks represent individuals seen in clinic, numbers on left are ages at diagnosis, and the arrow shows the proband. B: Pedigree of Family 2. Asterisks represent individuals seen in clinic, numbers on left are ages at diagnosis, and the arrow shows the proband. The symbol “S” stands for seizure and “DM” for deaf-mute. C: Pedigree of the Family 3. Asterisks represent individuals seen in clinic, numbers on left are ages at diagnosis, and the arrow shows the proband. The symbol “E” stands for epilepsy. D: represents the chromatogram showing the “G” to “A” change (asterisk), E: represents the chromatogram showing the “G” to “T” change (asterisk), F: represents the chromatogram showing the “C” to “T” change (asterisk), G: a portion of the highly conserved core catalytic domain (red) where the mutated valine residue is highlighted and conserved among wide range of species (asterisk), H: a portion of the highly conserved core catalytic domain (red) where the mutated phenylalanine residue is highlighted and conserved among wide range of species (asterisk), I: a portion of the highly conserved core catalytic domain (red) where the mutated arginine residue is highlighted and conserved among wide range of species (asterisk).

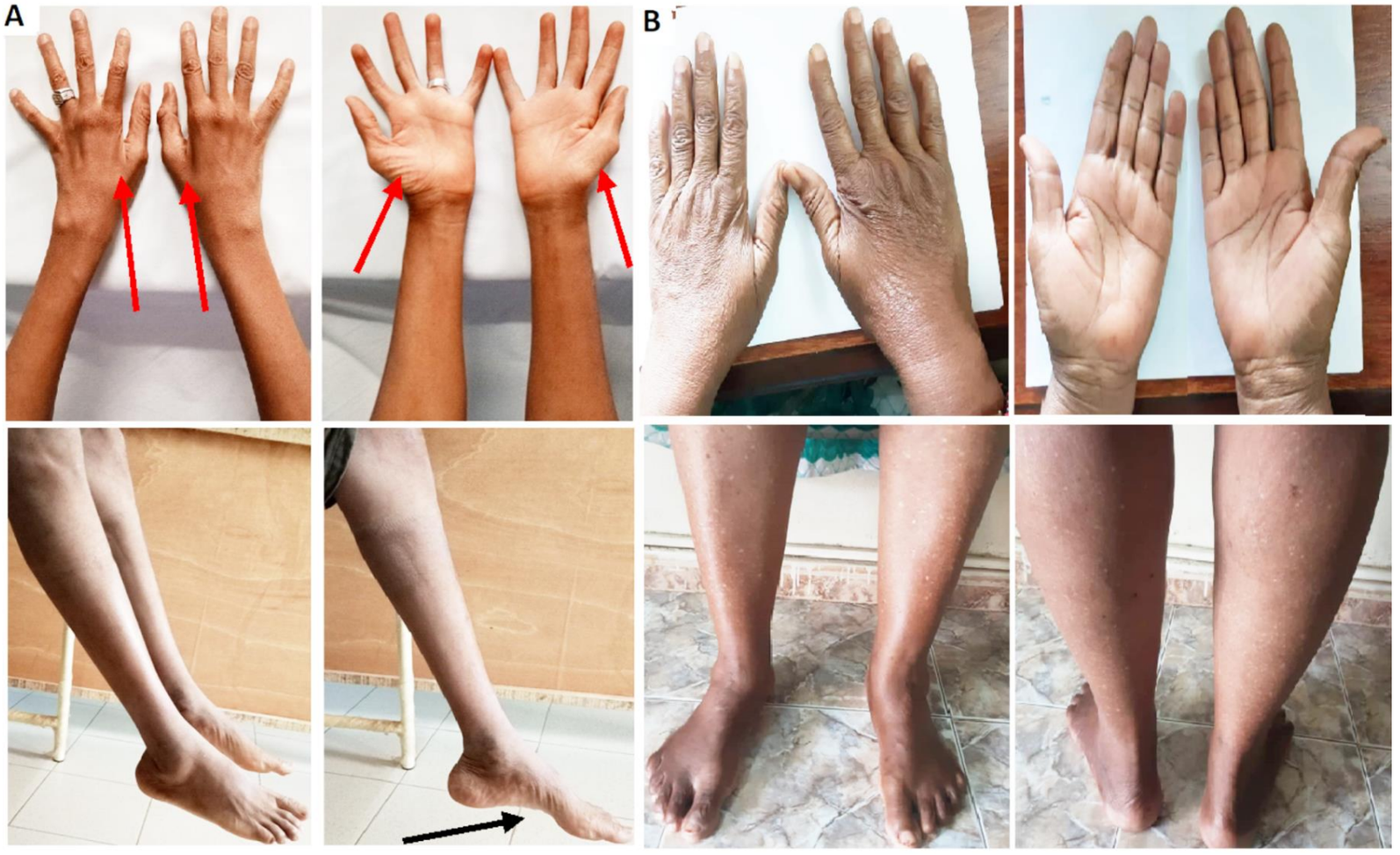


Figure 5: Images of patients with CMTX1: A: Photo of the proband of the Family 1 (IV.6) showing a severe distal muscle atrophy in the four limbs (red arrow) and a mild pes cavus (black arrow), B: Photo of the mother of the proband (IV.6) without muscle atrophy in the four limbs

Table 3: Phenotypic and genetic findings in patients with CMTX1

				Clinical examination findings							Laboratory findings								Genetic findings
					Arm weakness		Leg weakness		Sensory loss	Hearing loss	Nerve Conduction Studies								
											Median			Ulnar			Sural	Peroneal	*GJB1* gene
Patient	Age (yr)	Sex	Age of onset (yr)	First symptom	Proximal	Distal	Proximal	Distal			SNAP Amp	CMAP Amp	CV m/s	SNAP amp	CMAP amp	CV (m/s)	SNAP amp	CMAP Amp	
F1 IV.6	31	M	24	Walking difficutly	None	Severe	None	Severe	Severe	No	A	3.4	37	A	3.5	43	A	A	c.271G>A, p.Val91Met
F1 III.2	57	F	35	Muscle cramps	None	Mild	None	Mild	Mild	No	A	9.8	52	A	12.7	57	ND	ND	c.271G>A, p.Val91Met
F1 V.15	20	M	19	Hand deformities	None	Moderate	None	Moderate	Mild	No	ND	ND	ND	ND	ND	ND	ND	ND	c.271G>A, p.Val91Met
F1 III.8	47	F	35	Muscle cramps	None	Mild	None	Mild	Mild	No	ND	ND	ND	ND	ND	ND	ND	ND	c.271G>A, p.Val91Met
F1 IV.23	17	M	17	Walking difficulty	None	None	None	Mild	Minimal	No	ND	ND	ND	ND	ND	ND	ND	ND	c.271G>A, p.Val91Met
F1 III.11	37	F	20	Hand weakness	None	Mild	None	Mild	Mild	No	ND	ND	ND	ND	ND	ND	ND	ND	c.271G>A, p.Val91Met
F2 VI.3	19	F	16	Paresthesia	None	Severe	None	Severe	Severe	Yes	A	8.7	50	A	9.8	55	A	A	c.704T>G, p.Phe235Cys
F2 VI.2	24	F	NA	NA	None	NA	None	NA	NA	No	ND	ND	ND	ND	ND	ND	ND	ND	c.704T>G, p.Phe235Cys
F2 VI.8	17	F	16	Walking difficulty	None	Mild	None	Mild	Mild	No	36	11.5	53	12.2	11	67	30	3.7	ND
F3 V.16	33	M	23	Walking diffiulty	None	Severe	None	Severe	Severe	No	A	4.1	40	ND	4.1	49	A	A	c.43C>T, p.Arg15Try
F3 V.20	54	F	51	Foot weakness	None	Moderate	None	Moderate	Mild	No	ND	ND	ND	ND	ND	ND	ND	ND	c.43C>T, p.Arg15Try
F3 VI.7	17	M	13	Walking difficulty	None	Severe	None	Severe	Moderate	No	A	10.4	38	A	13.5	39	A	A	c.43C>T, p.Arg15Try
F3 VI.8	15	M	8	Pain in legs	None	Severe	None	Severe	Moderate	No	ND	ND	ND	ND	ND	ND	ND	ND	c.43C>T, p.Arg15Try
F3 IV.11	63	F	33	Muscle cramps	None	Moderate	None	Moderate	Mild	No	ND	ND	ND	ND	ND	ND	ND	ND	c.43C>T, p.Arg15Try
F3 V.21	43	F	20	Walking difficulty	None	Moderate	None	Severe	Moderate	No	A	3.8	30	A	4.2	53	A	A	c.43C>T, p.Arg15Try

Amp: amplitude; **SNAP:** sensory nerve action potential; **CMAP:** compound motor action potential; **CV**: conduction velocity; NA: not applicable; A: absent **ND:** not done; **m/s**: meter per second; normal median, **yr**: year; CMAP>4.5 mV

4.1.4 Discussion

CMTX1 is caused by the mutation in *GJB1* gene that encodes the Connexin 32 protein that constitutes a gap junction channels for the movement of small molecules and ions between cells [118]. This connexin gene is not only expressed on Schwann cells peripheral nervous system, but also expressed by oligodendrocytes in the CNS, coupling oligodendrocytes and astrocytes [119]. Connexins belong to a multigene family encoding ~20 highly homologous proteins [120]. X-linked Charcot-Marie-Tooth disease has been widely reported throughout the world with more than 400 mutations but missense mutations are the most commonly reported in the literature [121]. The genetic epidemiology of CMTX1 is sparse globally and was reported as the second most common subtype after CMT1A [95, 122]. However, in a cohort from Malaysia, CMTX1 was reported as the most common CMT subtype [95, 122, 123]. Only one case was reported in Africa on a North African patient [71]. We report genetically confirmed cases of CMTX1 in SSA population from three unrelated Malian families. The age at onset of symptoms reported in this study is confirmed in many other studies confirming that this disease occurs mostly after the second decade of life. CMT1X is an X-linked dominant trait because it affects female carriers. In addition to the typical CMT clinical features, patients with CMTX1 can have delayed motor development, sensorineural hearing loss, tremors, dysarthria, ataxia, spasticity, hyperreflexia, extensor plantar response [124]. Also, recurrent central nervous system disturbance, or transient white matter lesions have been reported in patients with CMTX1 which are reversible [125]. As seen in other reports, the proband of Family 2 had a severe sensorineural hearing impairment [126-128]. However, another unusual

association, epilepsy and CMTX1, was noted in the Family 3. To the best of our knowledge a case of epilepsy associated with CMTX1 has not been reported in the literature. This can be another CNS manifestation in CMTX1 patients or just an intercurrent disease as symptomatic epilepsy is common in SSA. Nevertheless, the mechanism by which the mutation in GJB1 gene causes epileptic seizure needs to be elucidated in further investigations as other CNS symptoms were reported in CMTX1. NCS showed no response in most of the nerves tested in lower limbs, and normal or reduced amplitudes in upper limbs, consistent with axonal type neuropathy. Studies including animal models demonstrated that loss of Cx32 causes both demyelination and axonal involvement [116]. However, as reported here, axonal type is reported to be the most common [129]. The variants identified in this study were previously reported in other studies [130-135]. Mutations of the Val91 residue leading to different amino acid changes were reported in the literature. In addition, mutations of several contiguous residues have also been reported as causing CMTX1, indicating that this region of the protein may be a hotspot for pathogenic variants. As found in our cohort, several studies in the literature also reported a high phenotype variability [116, 123]. The most severe phenotype seen in men reported in the study corroborate what is reported in literature [116]. About half of female patients have mild signs or are asymptomatic. Affected women usually have a later onset than men, and at every age the phenotype is milder. X-inactivation is the likely explanation for reduced severity in females as demonstrated in mice [136]. However, other studies have reported a more severe neuropathy in women linked to specific variants (Phe235Cys) [137]. The female patient with this variant in our cohort has shown a more severe phenotype than the females bearing other variants. This is important of note for the clinicians

that, notably females with neuromuscular disease of X-linked inheritance usually present, if at all, with minor symptoms and signs, in female patients with severe hereditary sensorimotor neuropathy, the X linked dominant form of CMT should be considered [138].

4.1.5 Conclusion

We report here the first genetically confirmed cases of CMTX1 in the sub-Saharan African population, expanding the global epidemiology of this disease. Further studies could inform on the underlying factors leading the high clinical variability. We believe that large and multicentric cohort studies in Africa will uncover more genetic variants for CMT and will lay the ground for inclusion of these under deserved patients in ongoing clinical trials.

Funding: This work is supported by grant number U01HG007044 funded by the National Institute of Neurological Disorders and Stroke and administered by the National Human Genome Research Institute under the H3Afria initiative, NINDS intramural funds NS002974, to GL; Wellcome Trust, grant number 107755Z/15/Z to AW (co-applicants); NIH, USA, grant number U01HG009716 to AW; and the African Academy of Science/Wellcome Trust, grant number H3A/18/001 to AW.

Author Contributions: Conceptualization, GL, AW; methodology, software, **AY**; validation, GL; formal analysis, **AY**, GL; investigation, **AY**, SHD, LC, MK, SD, TC, SD, DC, MK, COG, GL; writing original draft preparation, **AY**; writing review and editing, AW, KF, COG, GL; supervision, GL, AW; funding acquisition, GL, AW. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: Our thanks go to all the patients and their family members for participating in this study and the institutions that supported this study.

Data availability statement: The data that support the findings of this study are available from the corresponding author, [GL], upon reasonable request.

Conflict of interest: Authors declare they have no conflict of interest

4.2 **Abdoulaye Yalcouyé**, Seybou H. Diallo, Thomas Coulibaly, Lassana Cissé, Salimata Diallo, Oumar Samassékou, Salimata Diarra, Dramane Coulibaly, Mohamed Keita, Cheick O. Guinto, Kenneth Fischbeck, Guida Landouré, The H3Africa Consortium. A novel mutation in the *GARS* gene in a Malian family with Charcot-Marie-Tooth disease. Mol Genet Genomic Med. 2019 Jul;7(7):e00782. doi: 10.1002/mgg3.782.

Nature of publication: Case report

Journal publisher: Molecular genetics & genomic medicine. MGGM-2324-9269, peer review journal

Candidate contribution: Performed clinical evaluation, genetic analysis, drafting the first version of the manuscript

Co-authors contributions:

GL: conceptualization, supervision and funding acquisition

SHD: Contributed to clinical evaluation

TC: Contributed to clinical evaluation

LC: Contributed to clinical evaluation and EMG

SD: Contributed to clinical evaluation

DC: Contributed to clinical evaluation

COG: Conceptualization and supervision of the study

KF: Critically revised the manuscript

A novel mutation in the *GARS* gene in a Malian family with Charcot-Marie-Tooth disease

Abdoulaye Yalcouyé[1], Seybou H Diallo[1,2], Thomas Coulibaly[1,3], Lassana Cissé[3], Salimata Diallo[2], Oumar Samassékou[1], Salimata Diarra[1,4], Dramane Coulibaly[5], Mohamed Keita[1,6], Cheick O Guinto[1,3], Kenneth Fischbeck[4], Guida Landouré[1,3,4], The H3Africa Consortium

[1]Faculté de Médecine et d'Odontostomatologie, USTTB, Bamako, Mali, [2]Service de Neurologie, Centre Hospitalier Universitaire Gabriel Touré, Bamako, Mali, [3]Service de Neurologie, Centre Hospitalier Universitaire du Point "G", Bamako, Mali, [4]Neurogenetics Branch, National Institutes of Neurological Disorders and Stroke, Bethesda, MD, [5]Service de Médecine, Centre Hospitalier Universitaire Mère-Enfant le "Luxembourg", Bamako, Mali, [6]Service d'ORL, Centre Hospitalier Universitaire Gabriel Touré, Bamako, Mali,

Corresponding author: Guida Landouré, Faculté de Médecine et d'Odontostomatologie, USTTB, Bamako, Mali, BP: 1805, Tel: +223 76 36 34 68, Fax: +223 20 22 96 58, email: guida@icermali.org

Funding information: This work was supported by National Institutes of Neurological Disorders and Stroke, H3Africa Initiative U01HG007044; National Institutes of Neurological Disorders and Stroke, Intramural funds NS00297.

Abstract

Background: Charcot-Marie-Tooth (CMT) disease is a very heterogeneous neurological condition with more than 90 reported genetic entities. It is the most common inherited peripheral neuropathy; however, cases are rarely

reported in sub-Saharan Africa. In addition, only few families, mostly of Caucasian ancestry, have been reported to have CMT2D mutations. To date no case of CMT2D was reported in Africa. We present here a consanguineous family with CMT phenotype in which a novel mutation in the GARS gene was identified.

Methods: Patients were examined thoroughly, and nerve conduction studies (NCS) were performed. DNA from the proband was used for CMT gene panel testing (including 50 genes, *PMP22* duplication and mtDNA). Putative mutations were verified in all available family members to check for segregation.

Results: Two individuals, a male and a female, were found to be affected. Symptoms started in their teenage years with muscle weakness and atrophy in hands. Later, distal involvement of the lower limbs was noticed. Patients complained of minor sensory impairment. NCS showed no response in the upper as well as the lower limbs. Genetic testing surprisingly identified a novel heterozygous missense mutation c.794C>A (p.Ser265Tyr) in the *GARS* (glycyl-tRNA synthetase) gene associated with CMT2D. This variant segregated with the disease in the family and was also seen in the mother who presented no symptoms.

Conclusion: This is the first report of a genetically confirmed CMT2D case in Africa, expanding its genetic epidemiology. Increasing access to genetic testing may reveal more novel CMT variants or genes in the African population that could be relevant to other populations and further our understanding of their mechanism

Keywords: CMT, CMT2D, *GARS*, novel mutation, Mali

4.2.1 Introduction

Charcot-Marie-Tooth (CMT) disease also called hereditary motor and sensory neuropathy is a heterogeneous group of degenerative peripheral nerve disorders. It is characterized by a progressive demyelination or axonal degeneration and cell death resulting in distal muscle weakness and atrophy and sensory loss. Charcot-Marie-Tooth disease type 2D (CMT2D) is a classic axonal peripheral sensorimotor neuropathy characterized by weakness and atrophy of more upper than lower limbs involving mostly thenar and first dorsal interosseous muscles. Earliest symptoms in many individuals include transient cramps and pain in hands and calf muscles on exposure to cold or exertion. Sensory loss may affect all modes and can be variable within and across families. The disease is caused by mutations in the glycyl-tRNA synthetase (*GARS*, NG_007942.1, #601472) gene [139] and distal spinal muscular atrophy type V (dSMA-V) which does not present sensory impairment is an allelic disorder [140].

Although described in other populations, no genetically confirmed CMT2D case has been reported in the literature in Africa , and particularly in West Africa. In this study, we report a novel mutation in GARS that causes autosomal dominant CMT2D in a Malian family.

4.2.2 Methods

4.2.2.1 Ethical compliance

This study was approved by the Ethics Committee of the Faculty of Medicine and Dentistry, University of Sciences, Techniques and Technologies of Bamako, Mali.

4.2.2.2 Clinical and genetic analysis

All individuals included in the study were evaluated by a group of neurologists after giving consent. Blood chemistries including glucose and vitamin B12 levels were done to exclude common acquired causes of polyneuropathy. Nerve conduction studies (NCS), ENT and ophthalmologic examinations were done to assess peripheral nerve, ear or ocular involvement. DNA was extracted from peripheral blood in all available family members for genetic analysis. PMP22 duplication/deletion (NG_007949.1) analysis was done first, and then a Next-generation CMT gene panel testing composed of 50 genes and the mtDNA (Medical Neurogenetics, Atlanta, GA) including *GARS* gene (NG_007942.1) was performed on the index patient. The putative mutation was verified in SNP databases to exclude rare variants and was checked in all available family members. In silico analyses with PolyPhen, I-Mutant and SIFT were performed for deleteriousness.

4.2.3 Results

The patients are from a consanguineous family of Bambara ethnicity, and two out of the five living siblings presented with distal muscle weakness and atrophy (Figure 6A). Family history showed that neither of the parents reported symptoms. Symptoms started for both patients when they were around 12 years old with upper limb muscle weakness that progressed over time and involved the thenar and interosseous muscles. Later, lower limbs were also involved. On examination, both affected individuals had distal muscle weakness and atrophy and sensory loss (pin-prick and vibration) more pronounced in the upper than the lower limbs, decreased to absent reflexes, and plantar stimulation was neutral. Overall, symptoms were worse in the older brother who had claw hands, and both had steppage gait. Past medical history is consistent with recurrent seizures in both patients around

age 12, and EEG has showed slow frontal temporal waves in the older patient, though he has no active seizures at present. NCS showed no response in any of the nerves tested. The clinical and laboratory findings are summarized in Table 4. Genetic testing of the proband identified a heterozygous missense variant in the GARS gene at position c.794C>A, leading to the amino acid change Ser265Tyr (Figure 6B). Sequencing of DNA from other family members showed that the affected sister carried the Ser265Tyr variant as well as the mother who had no obvious symptoms. Thus, what appeared to be autosomal recessive inheritance was instead autosomal dominant with variable penetrance. The Ser265 residue is conserved across a wide range of species and is located in a highly conserved domain of the protein (Figure 6C). In addition, the c.794C>A change was not found in SNP databases (ExAC Browser, ClinVar, dbSNP, 1000genome), and was shown deleterious *in silico*.

Table 4: Phenotypic characteristics of subjects with CMT2D

Patient	Clinical and demographic features							Nerve Conduction Studies						
	Age (yr)	Sex	Age of onset (yr)	First symptom	Distal upper limb weakness and atrophy	Distal lower limb weakness and atrophy	Sensory loss	Left peroneal		Sural	Median		Tibial	
								CMAP Amp (mV)	CV m/s	SNAP Amp	CMAP Amp (mV)	CV m/s	CMAP Amp (mV)	CV m/s
IV.4	35	M	12	Hand weakness	Severe	Severe	Moderate	NR	NR	NR	NR	NR	NR	NR
IV.6	19	F	10	Walking difficulty	Moderate	Moderate	Moderate	NR	NR	NR	NR	NR	NR	NR
III.7	58	F	N/A	N/A	No	No	No	ND	ND	ND	ND	ND	ND	ND

Amp: amplitude, **SNAP:** sensory nerve action potential, **CMAP:** compound motor action potential, **NR:** no response, **ND:** not done, **CV:** conduction velocity, normal median CMAP>4.5 mV (recorded at abductor pollicis brevis muscle), normal peroneal CMAP>2.5 mV (recorded at extensor digitorus brevis muscle), normal tibial CMAP>6 mV (recorded at abductor hallucis muscle), normal sural SNAP>10 □V, normal F wave<55 ms lower limbs and <32 ms upper limbs.

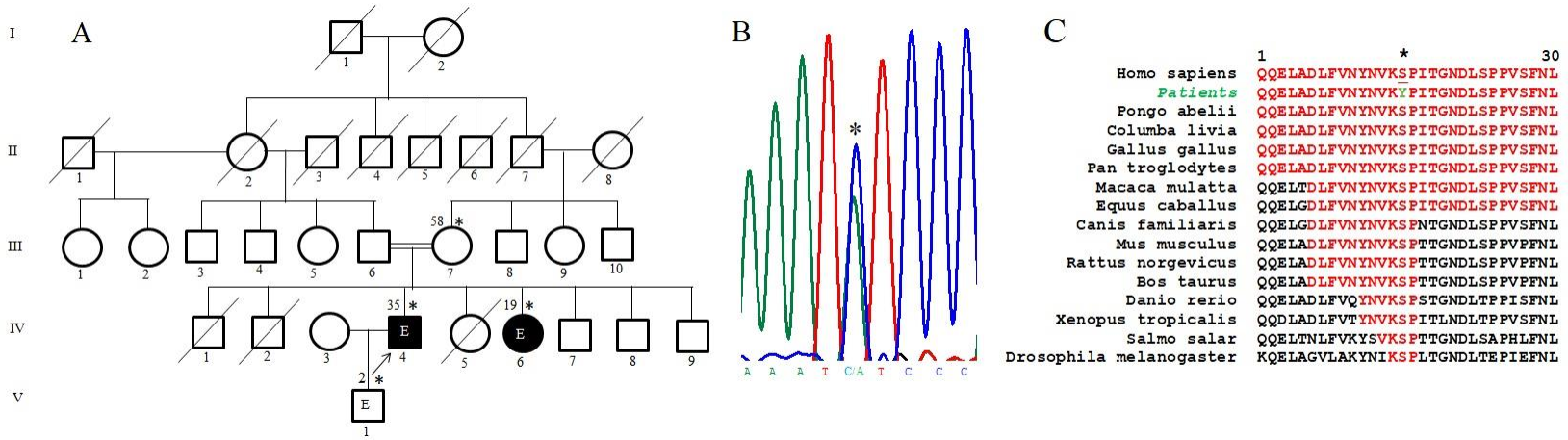


Figure 6: Pedigree of the family with CMT2D and genetic data

Pedigree of the family with CMT2D and genetic data. (a) Pedigree of the family with CMT2D. Asterisks represent individuals seen in clinic, numbers on left are ages at diagnosis, and the arrow shows the proband. The symbol “E” stands for epilepsy. (b) represents the chromatogram showing the “C” to “A” change (asterisk), and (c) a portion of the highly conserved core catalytic domain (red) where the mutated serine residue is highlighted and conserved from humans to flies (asterisk)

4.2.4 Discussion

CMT2D caused by mutations in the *GARS* gene have been found in a relatively small number of families from different populations in Europe, North America, and Asia [141]. Mutations in the *GARS* gene have also been shown to cause distal spinal muscular atrophy type V (dSMA-V), an early onset of pure motor neuropathy. While no CMT2D case was reported in Africa, dSMA-V has been reported in a family with North African origin [142].

We report here a novel mutation in *GARS* causing CMT2D in a Malian family. CMT2D is an autosomal dominant disease with high intra-family variability, and most of the patients present symptoms during the second decade [140]. The hallmark of the disease is the presenting manifestations in the hands with weakness and atrophy in the thenar and the first dorsal interosseous muscles and the sparing of the hypothenar eminence, which is involved later in the disease course. Lower limbs are involved in about half of affected individuals, and mild loss of vibration sense is observed in a third of individuals in the late stage of the disease. High intra and inter-familial variability is observed in this disease. This is reflected in the family we report here in which two siblings presented with symptoms around 12 years of age and neither of the parents complained of symptoms. The patients' mother was tested positive for the variant, but her neurological examination was normal. Unfortunately, she was not available for nerve conduction studies as it has been shown that in CMT conduction abnormalities may precede clinical ones. The adolescent onset and presence of sensory symptoms in the patients indicate that they present CMT2D rather than dSMA-V. The underlying mechanism distinguishing CMT2D from its allelic disease dSMA-V is still not well elucidated although a recent study has suggested that

CMT2D pathogenesis involves both neurodevelopmental and neurodegenerative processes [143].

Glycyl-tRNA synthetase is a member of the aminoacyl-tRNA synthetase family. The human GARS protein has three major functional domains including the core catalytic domain located at the 92nd–168th residues and the 241st–324th residues where the mutation reported in this study is located [144]. Its function is to catalyze the esterification reaction between the carboxyl group of glycine and its cognate tRNAs, resulting in aminoacylation of the tRNAs, substrates for protein synthesis in the ribosome [145]. Mutations in or close to the catalytic domain have been previously reported, and reduced GARS catalytic activity or other functional impairment of GARS have been suggested as a cause of the axonal neuropathy [140]. However, other studies have suggested that toxicity of mutant GARS may cause the neuropathy [146].

Some *GARS* mutations have weak genetic and functional evidence of causing CMT [147]. However, a mutation of the Ser265 residue leading to a different amino acid change was reported in a family with dSMA-V [127]. This, in addition to the conservation of the Ser265 throughout wide range of species, the deleteriousness of the Ser265Tyr change by in-silico predictions and its absence in SNP databases, suggests that the mutation we report is indeed pathogenic.

We report here the first CMT2D cases in Africa and a novel mutation in GARS, expanding the genetic epidemiology of this CMT sub-type. The Ser265Tyr variant, which was not previously reported elsewhere, adds to the limited number of *GARS* disease-associated variants. Although some clinical variability has been published in CMT2D families, the clinical pattern seen in

the family presented in this study has rarely been reported. This can be stochastic or due to other genetic or environmental modifiers. Larger cohort studies in Africa will allow phenotype-genotype studies to understand the phenotypic variability and the underlying mechanism distinguishing CMT2D from dSMA-V. Such studies may also uncover other new CMT variants or genes that can be studied in other populations.

Acknowledgments: This work is supported by grant number U01HG007044 funded by the National Institute of Neurological Disorders and Stroke and administered by the National Human Genome Research Institute under the H3Africa initiative, NINDS intramural funds NS002974, the Université des Sciences, des Techniques et des Technologies de Bamako and the Centre Hospitalier Universitaire du Point “G”, Bamako, Mali.

Conflict of interest: The authors declare that they have no conflict of interest.

Chapter 5: Results - Clinical description of HI in Mali

Synopsis: This chapter presents a comprehensive clinical description of both environmental and genetic aspects of HI in Mali. In addition, it includes the summary of all previous HI studies reported in the Malian population.

5.1 Abdoulaye Yalcouyé, Oumou Traoré, Abdoulaye Taméga, Alassane B Maïga, Fousseyni Kané, Oluwafemi G Oluwole, Cheick Oumar Guinto, Mohamed Kéita, Samba Karim Timbo, Carmen De Kock, Guida Landouré, Ambroise Wonkam. Etiologies of Childhood Hearing Impairment in Schools for the Deaf in Mali. Front Pediatr. 2021 Nov 29;9:726776.doi: 10.3389/fped.2021.726776. eCollection 2021.

Nature of publication: Original full article,

Journal publisher: Frontiers in pediatrics. ISSN: 2296-2360, peer review journal

Candidate's contribution: Patient recruitment, analysis and interpretation of the data, drafting first version and subsequent revisions.

Co-authors' contributions:

GL and **AW**: conceptualization, supervision, and funding acquisition.

FK, OT, AT, AM, and **GL**: contributed to developing the methodology.

OO, CD, CG, MK, ST, GL, and **AW**: editing and reviewing the manuscript. All authors have read and agreed to the published version of the manuscript.

Etiologies of Childhood Hearing Impairment in Schools for the Deaf in Mali

Abdoulaye Yalcouyé[1,2], Oumou Traoré [1], Abdoulaye Taméga[1], Alassane B Maïga[1], Fousseyni Kané[1], Oluwafemi G Oluwole[2], Cheick Oumar Guinto[1,3], Mohamed Kéita[1,4], Samba Karim Timbo[1,4], Carmen DeKock[2], Guida Landouré[1,3], Ambroise Wonkam[2,5].

[1]Faculté de Médecine et d'Odondostomatologie, Université des Sciences, Techniques et Technologies de Bamako (USTTB), Bamako, Mali.

[2]Division of Human Genetics, Department of Pathology, University of Cape Town, Cape Town, South Africa.

[3]Service de Neurologie, Centre Hospitalier Universitaire du Point "G", Bamako, Mali.

[4]Service d'ORL, Centre Hospitalier Universitaire de Gabriel Touré, Bamako, Mali.

[5]Department of Medicine, University of Cape Town, Cape Town, South Africa.

Received: 17 June 2021, **Accepted**: 01 November 2021, **Published**: 29 November 2021

Abstract

Objectives: To identify the etiologies of hearing impairment (HI) in schools for students who are deaf and to use a systematic review to summarize reports on the etiologies and clinical and genetic features of HI in Mali. **Methods:** We included individuals with HI that started before the age of 15 years old. Patients were carefully evaluated under standard practices, and

pure-tone audiometry was performed where possible. We then searched for articles published on HI in the Malian population from the databases' inception to March 30, 2020.

Results: A total of 117 individuals from two schools for the deaf were included, and a male predominance (sex ratio 1.3; 65/52) was noted. HI was pre-lingual in 82.2% (n = 117), and the median age at diagnosis was 12 years old. The etiologies were environmental in 59.4% (70/117), with meningitis being the leading cause (40%, 20/70), followed by cases with genetic suspicion (29.3%, 21/117). In 11.3% (8/117) of patients, no etiology was identified. Among cases with genetic suspicion, three were syndromic, including two cases of Waardenburg syndrome, while 15 individuals had non-syndromic HI. An autosomal recessive inheritance pattern was observed in 83.3% of families (15/18), and consanguinity was reported in 55.5% (10/18) of putative genetic cases.

Conclusion: This study concludes that environmental factors are the leading causes of HI in Mali. However, genetic causes should be investigated, particularly in the context of a population with a high consanguinity rate.

Keywords: Africa; Mali; etiology; genetics; hearing impairment

5.1.1 Introduction

Hearing impairment (HI) is a partial or total inability to hear, and constitutes a leading cause of disability in the world [148]. Its global incidence is estimated to 1/1000 live newborns in developed countries which is five to six times lower than what is reported in developing countries [4, 149]. This disparity is attributed to poor healthcare systems in developing countries that are not always adequately equipped to prevent, screen and manage causes of HI [35]. The prevalence varies widely from one region to another based on multiple factors like the prevalence of infectious conditions (meningitis, rubella, cytomegalovirus), and monitoring and assistance of pregnancy and birth [150]. It affects both adults and children but has more serious implications among children due to its potential in interfering with language acquisition and cognitive development. At least 5% of the world population experience a disabling HI which represents 466 million people, and 34 million of these are children living in middle and low-income countries [148]. HI is a public health concern negatively impacting people's well-being with functional, economic, social and emotional impact.

HI can be classified following several aspects such as the age of onset, degree of hearing loss, location of the lesion, or the aetiology. WHO reported that 60% of the childhood HI are due to preventable causes [148]. The aetiology of HI can be environmental and/or genetic, and the prevalence of each group of aetiologies varies from different regions. Environmental causes include infectious diseases (meningitis and measles), ototoxicity (medication, exposure to noise, and trauma) or other factors such as low birth weight, prematurity, and neonatal jaundice. While infectious diseases are the leading cause of HI in low and middle-income countries, the burden

of disease is lower in high-income countries [10, 151, 152]. Some studies reported that the aetiologies of almost 30% of HI remain unknown [5, 153]. In the developed countries, more than 50% of congenital HI are of genetic origin [154-156].

Hereditary hearing impairment (HHI) is divided into two main types, syndromic HI (SHI) where HI is associated with other organs' abnormalities and non-syndromic HI (NSHI) when HI is the only sign of the disease. NSHI is the most common type of HHI accounting for approximately 70% of all types of HHI [154]. HI is known to be associated with over 400 syndromes including Waardenburg syndrome, Branchio-Oto-Renal syndrome, Usher syndrome, Pendred syndrome, keratitis–ichthyosis–deafness syndrome and Alport syndrome [35, 154, 157]. The genetic cause of HI in numerous African countries such as Mali has not been properly investigated.

Only one study assessed the aetiologies of HI among 46 school-age children in Bamako (Mali), and reported meningitis as the leading cause (54,3%) while the aetiology was not known in 19,6 % [153]. Despite the high prevalence of HI in Mali [148], only a few data sources describe the aetiologies of this disease [153, 158, 159]. Therefore, we performed the present work to identify the aetiologies of HI among school-age children in Mali and to conduct a systematic review on the previously reported data emphasizing on the frequency, aetiologies, clinical and genetic features of HI in Mali.

5.1.2 Methods

Material and methods:

5.1.2.1 Patients

The study was performed at the two schools for the deaf in Bamako and the Department of Neurology of the Teaching Hospital of Point “G” (Figure 5.1) in collaboration with the Division of Human Genetics, Faculty of Health, University of Cape Town, Cape Town, South Africa. Were included in this study, individuals with HI that started before the age of 15 years. Parents and patients were gathered at the schools for the deaf to explain the study objectives, emphasizing the interest for familial or non-environmental cases. The voluntariness of participation and the possibility to withdraw with no consequences were highlighted. Families that registered were called for enrollment the following days to schedule inclusion day. After obtaining informed consent, sociodemographic data, and a detailed medical including prenatal, neonatal and postnatal and family history were obtained from all the participants.

5.1.2.2 Clinical assessment

Each patient completed a questionnaire and underwent a careful clinical examination. Then pure tone audiometry (PTA) was performed on all available patients who were lacking audiological tests with a mobile audiometer (KUDUWAVETM N° 090103564). Audiological testing results that were obtained before school admission were reviewed for some patients. The hearing level was classified according to the International Bureau of Audiophonoly (IBAP) number 02/1 (www.biap.org). For families with a suspected genetic origin of HI, a pedigree was drawn to elaborate on the patterns of inheritance. When SHI was suspected, additional testing such as serum creatinine dosage, thyroid hormone, kidney and thyroid gland ultrasounds, temporal bone CT-scan when possible and ophthalmological assessment, was later performed to refine the diagnosis.

5.1.2.3 Operational definitions

In the context of the present study, HI was defined as 1) acquired if there was a relationship between a putative environmental factor and the onset of the HI; 2) likely genetic when more than one case was reported in the family, in the case of consanguinity, and clinically well-defined syndromic cases; 3) of unknown aetiology when the cause was not established as environmental or genetic as previously reported by Wonkam et al [9].

5.1.2.4 Literature review process

5.1.2.4.1 Selection criteria for the literature review

We included studies published from database inception to March 30th 2020 that report data on the prevalence, aetiologies, clinical or genetic features of HI in the Malian population. Also, specific Malian authors' names active in the field of HI were used to complement the literature searches. There was no restriction on the reporting language of the article and accessible full-length articles were selected.

5.1.2.4.2 Method of search for relevant articles

We searched in PubMed, Google Scholar, Microsoft Academy, Scopus, Science Direct, MEDLINE, African Journals Online, AFROLIB, African Index Medicus using the keywords "hearing impairment OR hearing loss OR deafness AND Mali". As Mali is majorly a French-speaking country, the search was also done in French

5.1.3 Results

5.1.3.1 Clinical and audiological data

We enrolled 100 families totaling 117 individuals with HI, of which 65 were males and 52 were females (sex-ratio: 1.3). Most participants were of the Bambara ethnic group, and from Kayes (first region of Mali). Overall, 89.7% (N = 117) were attending primary schools, and only 4.3% (5/117) were in secondary schools. Seven patients were school attendants' siblings who dropped school at the primary level. The median of age at diagnosis was 12 years, and HI was prelingual in 82.2% (96/117). Audiometry was performed in 61 participants and confirmed sensorineural HI in 80.7% of cases (n = 49) (Figure 8A). The aetiology was environmental in most of the cases 59.4% (70/117) (Figure 8B) with meningitis being the major factor representing 40% 28/70of the cases followed by chronic otitis and low birth weight in 18.6% (13/70) and 15.7 % (11/70), respectively. Ototoxic medications, congenital infection and neonatal asphyxia were implicated in 14.3% (10/70), 8.6% (6/70) and 2.8% (2/70), respectively. The ototoxic medication was quinine hydrochloride taken for up to nine months by four women during pregnancy. In our cohort, following pedigrees analysis of 18 families, HI was observed in another sibling or a relative of the proband, suggesting a genetic origin. Among them, three were SHI including two cases of Waardenburg syndrome (WS) (Figure 8C and D), one case of congenital microtia-deafness syndrome (Figure 8E), and 15 had NSHI features. Soninke ethnic group was the most represented in the familial cases with 33.3% (n = 18). The inheritance pattern was consistent with autosomal recessive in 83.3 % (15/18) (Figure 8F), autosomal dominant in 6.5% (Figure 8G), and likely sporadic in 11.1% of the families (n = 2/18). Furthermore, parents reported consanguinity in 55.5% of the putative genetic cases (n = 10/18). More pedigrees of familial cases are provided as Supplementary material (Figure S2). The sociodemographic and phenotypic characteristics of the patients are summarized in Table 5.

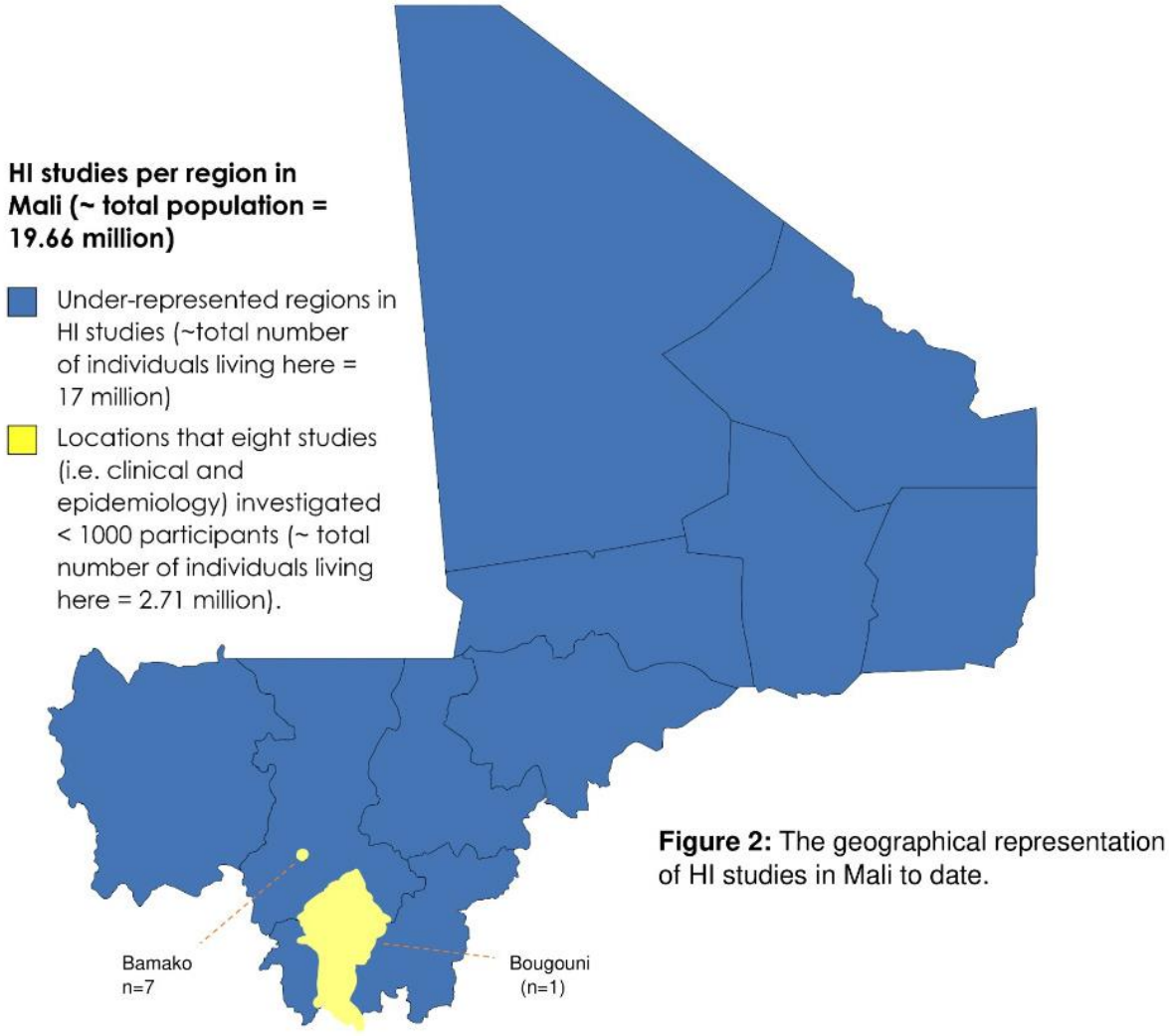


Figure 2: The geographical representation of HI studies in Mali to date.

Figure 7: Map of Mali showing the regions where the studies in this review were performed. *The geographical representation of HI studies in Mali to date.*

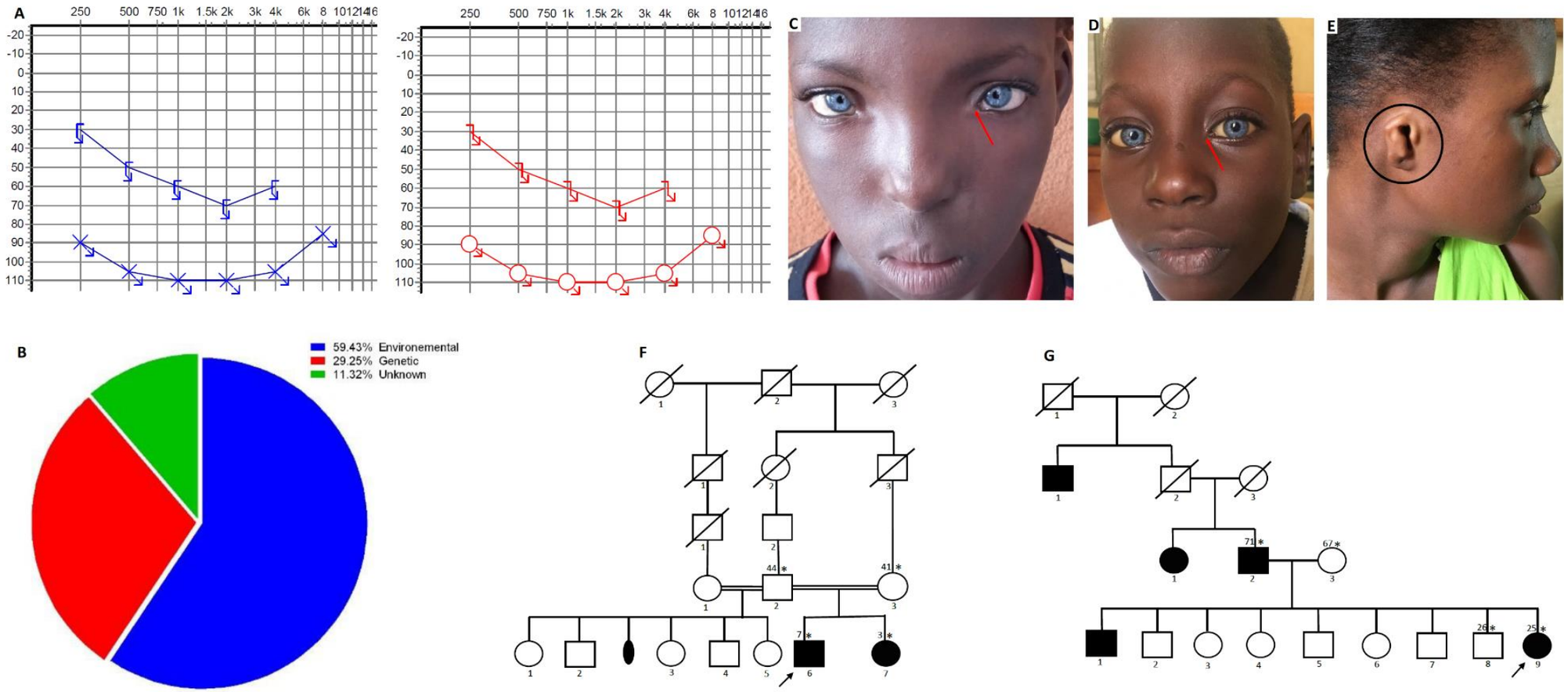


***Figure 8: Clinical profiles**:* (A) Audiogram of a patient showing a bilateral and profound sensorineural hearing impairment (blue is the left side and red is the right side). (B) Diagram of the etiologies of hearing impairment in Mali. (C,D) Photos of patients with Waardenburg syndrome type 1 with dystopia canthorum (red arrow), and type 2 without dystopia canthorum (red arrow). (E) Photo of a patient with congenital microtia-deafness syndrome showing a right side microtia (black circle). (F,G) Pedigree of the families showing an autosomal recessive and autosomal dominant pattern of inheritance (asterisks represent individuals seen in clinic, numbers on the left are ages, and the arrow shows the proband).

Table 5: Sociodemographic, phenotypic expression, and causes of HI

Sociodemographic aspects				**Phenotype characteristics of hearing impairment**															
		Sex		Onset		Type			Causes										
Median age at diagnosis	Age range	M	F	Pre N (%)	Post N (%)	Sen N (%)	Con N (%)	Mixed N (%)	Environmental (N=70)				Genetics (N=18)						Unknown
									Meningitis N (%)	Chronic otitis N (%)	Low birth weight N (%)	Other N (%)	Nonsyndromic N (%)	Syndromic N (%)	Inheritance N (%)			Consanguinity N (%)	
															Dominant	Recessive	Sporadic		
12 years	3-35 years	65	52	82.2 (n=96)	17.8 (n=21)	80.7 (n=49)	4.6 (n=3)	14.7% (n=9)	40% (n=28)	18.6% (n=13)	15.7% (n=11)	25.7% (n=18)	83.3 (n=15)	16.7 (n=3)	5.6 (n=1)	83.3 (n=15)	11.1 (n=2)	55.% (n=10)	11.3 (n=13)

HI: hearing impairment, Sen: sensorineural, Con: conductive, Pre: prelingual, Post: post lingual, *Other = Ototoxic medications, congenital infection, and neonatal asphyxia

5.1.3.2 Literature review data

Following a review of the literature, we identified eight articles that met our selection criteria (any article reporting hearing impairment cases in the Malian population). All the studies were performed in Bamako, the capital city of Mali. Only one was conducted at a school for the deaf, while four others were performed in hospital settings. We found three case reports and five cross-sectional case series. None of them reported the prevalence or incidence of HI in Mali, nor the genetic contribution in HI in Mali. However, they reported environmental factors in 80% (n = 37/46) of the participants, and meningitis contributed in HI to up to 54.3% (n = 25/46) in a study conducted in one of the schools for the deaf of Bamako [153]. The main diagnostic tool was the pure tone audiometry in almost all the studies. In a study conducted a Teaching Hospital that included 200 individuals, the audiometric profile varied from mixed type in 43.4 % (n = 87/200) to sensorineural in 32.2% (n = 64/200) of HI [158]. Two clinical cases of WS including one associated with cleft lip and another associated with dystopia Cantorum were reported [160, 161]. One case of HI of suspected ototoxic medication origin was reported [161]. The description of these studies is summarized in Table 6.

Table 6: Characteristics of studies included in the literature review

First author's name and year of publication	*Area*	*Region*	*Study Setting*	*Study Design*	*Males N (%)*	*Mean age (years)*	*Age range (years)*	*Sample Size*	*Prevalence*	*Main aetiologies*	*Diagnosis tool*
AA Mohamed, 1996[153]	Urban	Bamako	School for the deaf	Cross-sectional cases series	65.2 (*n*=30)	11	5–19	46	NR	Meningitis	PTA
Traoré H, 2011[159]	Urban	Bamako	Hospital	Case report	0	1.6	NA	1	NR	NI	PTA
Imperato PJ, 2015[160]	Rural	Bougouni	Community	Case report	100 (*n*=1)	13	NA	1	NR	NI	NA
Traoré M, 2006[161]	Urban	Bamako	Hospital	Case report	100 (*n*=1)	33	NA	1	NR	Toxic	PTA
Sacko HB, 1990[162]	Urban	Bamako	Community	Cross-sectional cases series	NR	NR	10–60	147	NR	Meningitis	PTA
Sacko HB, 2015[163]	Urban	Bamako	Hospital	Cross-sectional cases series	64 (*n*=533)	6	0–15	833	NR	Meningitis	PTA
Diarra K, 2020[158]	Urban	Bamako	Hospital	Cross-sectional cases series	60 (*n*=120)	37,18	15– 83	200	NR	NR	PTA
Sacko HB, 2016[164]	Urban	Bamako	Community	Cross-sectional cases series	100 (*n*=68)	32	30–55	68	NR	Noise	PTA
This study	Urban	Bamako	School for the deaf	Cross sectional case series	55.6 (*n*=65)	7	4-21	117	NR	Meningitis	PTA

NR: not reported, NI: not identified, NA: not applicable, PTA: pure tone audiometry

5.1.4 Discussion

5.1.4.1 Context of the research

Mali is a landlocked country located in the Centre of West Africa and is surrounded by seven other countries. With a territory of 1,220,190 sq km and a population of 15,302,000, the country has a young population; 47% are under 15 years of age [64, 165]. Mali is an ethnically and culturally diverse country, and its subpopulations have a long tradition of intra-ethnic and consanguineous marriages [40] that have particularly favoured gene identification for numerous recessive conditions [42, 43]. In a study on a sample of about 600 Malian students from different ethnic backgrounds, 27% reported consanguinity including 17% of parental first cousin marriage [40]. This results in homogeneous cluster populations with typical phenotypic characteristics and an increased prevalence of recessive disorders in some parts of the country. Moreover, while family-based genetic studies are often limited in high-income countries due to small sibships, the average fertility rate in Mali which is over 6 births per woman offers a unique opportunity to find new disease genes or variants that can then be studied in other populations [165].

There are at least 14 different major ethnic groups in Mali, each speaking a different language. However, about 80% of Malians speak the national language that is Bambara, also spoken in five neighboring countries [165]. Genetic studies have been introduced to Mali about 15 years ago leading to the identification of numerous genetics variants for rare neurogenetic diseases some of which presenting with characteristic features, favoured by the structure of its populations [42, 64, 166]. Despite that, genetic and genomic studies in Mali have been relatively limited to select families in a

specific specialty, hence an underestimation and neglect of genetic diseases and the genetic contribution in common diseases [64]. This may be due to social factors surrounding these diseases and limited resources that prevent patients from seeking care, and a lack of infrastructure and medical geneticists. The low literacy rate in Mali, as in many developing countries, may foster a low understanding of basic genetic concepts by the general population. For instance, Malians often consider genetic diseases to be a result of bad fate, which leads to their stigmatization [64]. The Malian populations had seldom been exposed to genetic testing in a routine clinical setting, and the knowledge, attitudes, and beliefs of Malian families with hereditary neurological disorders regarding genetic testing were assessed in a previous study. The results showed that in general, the majority favoured genetic testing and some gained knowledge from genetic counselling [64].

5.1.4.2 Sociodemographic, clinical, and genetic profiles

This study reports the main aetiologies of HI among school-age children based on an observational study in two schools for the deaf and reviews the literature, therefore providing comprehensive better understanding of the aetiologies of HI in Mali to date. Like reports from other sub-Saharan African (SSA) countries, our study revealed the high proportion of environmental causes [5, 9], while emphasizing the likely contribution of genetic aetiology of HI in Mali. This study also revealed a serious epidemiological gap, as no report about its incidence and prevalence were found, urging the need larger and more in-depth studies to evaluate its burden in the SSA countries [148]. Indeed, as reported in the present study and in the literature, prelingual HI is the most common type among children, with a relatively late diagnosis, which interferes strongly with language acquisition as this period is crucial for

harmonized psychosocial development [167]. It is not surprising that the Bambara ethnolinguistic group was the most represented in the cohorts, as it is the main ethnic group in the Malian population, representing 34.1% [153]. Like other reports from SSA countries [5, 9, 153, 158], males were predominant in this study. As previously reported elsewhere in Africa that, compared to female children, males are often preferentially sent to school for the deaf [9]. This is still a detrimental gender bias that needs to be addressed in society. Environmental factors were the main aetiologies of HI in Malian children as reported in many studies conducted across developing countries including in Africa [9, 10, 35, 167, 168]. This can be associated with limited access to adequately equipped healthcare centres to assist and monitor pregnancy and birth, and to provide relevant information on the environmental risk factors for HI in the community [150]. In addition, as reported in Cameroon [5], Mali is located in the African meningitis belt characterized by seasonal epidemics [5, 9, 158], resulting in a higher burden of meningitis-associated HI in Mali (Table 1) and in other African countries [10, 167]. Meningitis has been also identified as a major cause of HI in other developing countries outside Africa [5, 9, 10, 168]. Additional environmental factors such as mumps, measles, prematurity, or neonatal nuclear jaundice were also identified as aetiologies in this study but were not reported to play an important role in HI in some SSA countries [5]. However, this could be due to lack of health records or limited access to diagnosis tools.

The genetic contribution to HI varies widely among countries. In this study, we suspected its contribution in one-third of cases, similar to studies conducted in Cameroon [10]. This is lower than reported in high income countries where nearly 50% of congenital HI has a genetic origin. This is

likely due to better prevention and interventions that have contributed to reducing environmental hazard including new-born screening for HI, the availability of a comprehensive care and an easy access to molecular diagnosis facilities to confirm suspected genetic causes [169]. Moreover, the lack of proper genetic investigation facilities in Mali and in other SSA countries may lead to an underestimation of HI of genetic origin. To date, over 120 genes have been associated with HI [32], and it is estimated that 1% of human genes are involved in the hearing process [170]. From the first HI-associated gene identified in 1995 [171] to date, genes are continuously being discovered, and some cases with unknown aetiologies could probably have a genetic component as evoked by Frazer et al [172]. The higher level of consanguinity in Mali [40] should be associated with a much higher rate of congenital HI of genetic origin. This is confirmed by the high autosomal recessive inheritance patterns seen in our cohort. In fact, most of the familial cases reported in this study were seen in the Soninke ethnic group that have a high tendency for consanguineous marriage. In Mali, like in numerous understudied African populations, it is probable that numerous variants in known genes and potentially novel genes are still to be discovered [37, 173, 174]. Among cases with a suspected genetic origin, we identified two families segregating WS, the most common SHI reported in some countries [5], and associates a myriad of symptoms including dystopia Cantorum (lateral displacement of the inner canthus of each eye), pigmentary abnormalities of hair, iris, and skin (often white forelock and heterochromia iridis) and sensorineural deafness [32]. While one case was a type 1 with dystopia Cantorum, the other presented with type 2. Two other cases of WS were previously reported in the Malian population, a type 1 associated with labial cleft and a classical type 1 with no other malformations [159, 161]. Similar to

other reports in SSA [5], we also found one case of congenital microtia associating profound HI, confirming its rarity in the literature. Unsurprisingly, NSHI was the commonest subtype among putative genetic cases in our study, confirming what is already known worldwide [154].

5.1.4.3 Strengths and limitations

The data presented here has some limitations. For one, the recruitment from the schools of the deaf cannot be representative of the entire country as not all communities will have access to these schools (Figure 7). In addition, there is a lack of systematic screening of the HI at the admission of children in normal schools in Mali due to the limited availability of audiologists and trained nurses in school clinics, and the limited knowledge of the teachers and parents to early detect potential hearing-impaired children. These aspects can limit the clinical characterization of the hearing-impaired individuals even those attending the schools for the deaf, with a possible underestimation of syndromic cases. Despite these limitations, this cohort is the largest so far reported in the schools for the deaf in Mali and represents an initial step toward a proper epidemiological description, and genetic characterization of HI in the general population in Mali.

5.1.5 Conclusion

This study confirms environmental factors as the leading aetiologies of childhood HI and emphasized the high contribution of consanguinity to genetic causes of childhood HI in Mali. Policy actions must be taken in terms awareness for early treatment of otitis and reinforcing the existing immunization programs to vigorously fight these preventable factors which will decrease the occurrence of HI in developing countries. There is equally

an urgent need to undertake genetic investigations, and the increasing access to whole exome sequencing may identify variants in HI genes to favour early treatment of HI to allow for improved quality of life for those affected by HI and improve counselling of people with disease traits in recessive settings. This could also provide an opportunity for novel HI gene discovery that will further the understanding of these diseases and trigger future drug development.

Author Contributions: Conceptualization: **GL** and **AW**; Methodology: **AY**, OT, AT, ABM, GL collected the data; **AY**, FK analysed the data; Writing the first draft: **AY**; Editing and reviewing the manuscript: OGO, CDK, COG, MK, SKT, GL, AW; Supervision: GL, AW; Funding Acquisition, GL, AW. All authors have read and agreed to the published version of the manuscript.

Funding: This study is supported by the NIH (USA), grant number U01-HG-009716 to A.W, the African Academy of Science/Wellcome Trust, grant, number H3A/18/001 to A.W and grant number U01HG007044 to G.L. The funders had no role in data synthesis and analysis, and decision to publish.

Ethical considerations: This research was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Ethics Committee of the Faculty of Medicine and Dentistry of the University of Sciences, Techniques and Technologies of Bamako, Mali (N°2020/129/CE/FMOS/FAPH) and the University of Cape Town (HREF REF: 691/2020). Informed consent was obtained from all participants of this study.

Conflicts of Interest: The authors declare no conflict of interest

Acknowledgments: We would like to thank all the patients and their parents and the teachers of the “Ecole pour Déficients Auditifs (EDA) de l’Hippodrome” and “Djiguiya Kalanso de Niamakoro” of Bamako, Mali.

Supplementary material: Can be found online at:
https://www.frontiersin.org/articles/10.3389/fped.

2021.726776/full#supplementary-material

Chapter 6: Results - Whole exome sequencing and bioinformatics tools to identify causative variants in HI

Synopsis: This chapter presents the use of whole exome sequencing to identify causative variants in hereditary HI. It contains two parts, the first describes the case of Branchio-otic syndrome and the second the case of NSHI. These reports are the first in the Malian population to date.

6.1 Abdoulaye Yalcouyé, Oumou Traoré, Salimata Diarra, Isabelle Schrauwen, Kevin Esoh, Magda Kamila Kadlubowska, Thashi Bharadwaj, Samuel Mawuli Adadey, Mohamed Kéita, Cheick O Guinto, Suzanne M Leal, Guida Landouré, Ambroise Wonkam. A monoallelic variant in EYA1 is associated with Branchio-Otic syndrome in a Malian family. Mol Genet Genomic Med. 2022 Jun 14;e1995. doi: 10.1002/mgg3.1995.

Nature of publication: Original case report

Journal publisher: Molecular genetics & genomic medicine. ISSN 2324-9269. peer review journal

Candidate's contribution: Clinical and genetic data collection, genetic experiments and drafting first version of the manuscript.

GL, SML and **AW**: conceived the study

OT, SD, SMA, MK, COG: Contributed to perform clinical evaluation

IS, KE, MKK, TB: Contributed to the analysis and interpretation of the whole exome sequencing data

KE, AW, GL, IS, SML critically revised successive drafts of the manuscript

A monoallelic variant in EYA1 is associated with Branchio- Otic syndrome in a Malian family

Abdoulaye Yalcouyé[1,2], Traoré Oumou[2], Salimata Diarra[2,3], Isabelle Schrauwen[4], Kevin Esoh[1], Magda Kamila Kadlubowska[4], Thashi Bharadwaj[4], Samuel Mawuli Adadey[1], Mohamed Kéita[2,5], Cheick O. Guinto[2,6], Suzanne M. Leal[4,7], Guida Landouré[2,3,6], Ambroise Wonkam[1,8]

[1]Division of Human Genetics, Department of Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

[2]Faculté de Médecine et d'Odondostomatologie, USTTB, Bamako, Mali

[3]Neurogenetics Branch, National Institute of Neurological Disorders and Stroke, Bethesda, Maryland, USA

[4]Center for Statistical Genetics, Gertrude H. Sergievsky Center, and Department of Neurology, Columbia University Medical Center, New York, New York, USA

[5]Service d'ORL, Centre Hospitalier Universitaire de Gabriel Touré, Bamako, Mali

[6]Service de Neurologie, Centre Hospitalier Universitaire du Point “G”, Bamako, Mali

[7]Taub Institute for Alzheimer's Disease and the Aging Brain, Columbia University Medical Center, New York, New York, USA

[8]McKusick- Nathans Institute and Department of Genetic Medicine, Johns Hopkins University, Baltimore, Maryland, USA

Correspondence: Ambroise Wonkam, McKusick- Nathans Institute and Department of Genetic Medicine, Johns Hopkins University, Baltimore, MD, USA. Email: ambroise.wonkam@uct.ac.za

Funding information: National Institute of Deafness and other Communication Disorders (NIDCD), Grant/Award Number: R01 DC01165, R01 DC003594 and R01 DC016593 to SM; National Institute of Neurological Disorders and Stroke, Grant/Award Number: U01HG007044 to GL; National Institute on Deafness and Other Communication Disorders, Grant/Award Number: U01HG009716 to AW; Wellcome Trust, Grant/Award Number: 107755Z/15/Z to AW and H3A/18/001 to AW

Abstract

Background: Branchio-otic syndrome (BO) is one of the most common types of syndromic hearing impairment (HI) with an incidence of 1/40.000 globally. It is an autosomal dominant disorder typically characterized by the coexistence of branchial cysts or fistulae, malformations of the external, middle and inner ears with pre-auricular pits or tags and a variable degree of HI. Most cases of BO have been reported in populations of European ancestry. To date, only few cases have been reported in people from African descent. **Methods:** After a careful clinical examination, a pure tone audiometry was performed. DNA was extracted from peripheral blood and Whole Exome, and Sanger Sequencing were performed for genetic analysis. **Results:** Eight individuals from a large non-consanguineous Malian family, with autosomal dominant inheritance were enrolled. The ages at diagnosis ranged from 8 to 54 years. A high phenotypic variability was noted among the affected individuals. Four patients presented with a post-lingual and mixed type of HI, one individual had conductive HI while three had normal hearing but presented other BO features namely branchial fistulae and preauricular sinus. Serum creatinine level and renal ultrasonography were normal in three affected individuals who performed them. Genetic testing identified a monoallelic pathogenic variant in *EYA1* (c.1286A>G; p.Asp429Gly) segregating with BO syndrome in the family.

Conclusion: This is the first genetically confirmed case of BO syndrome caused by *EYA1* variant in the sub-Saharan African population, expanding the genetic spectrum of the condition.

Keywords: Branchio-otic syndrome, Syndromic hearing impairment, *EYA1*, Mali, Africa

6.1.1 Introduction

Branchio-oto-renal syndrome (BO) is an autosomal dominant (AD) condition that is one of the most common forms of syndromic hearing impairments (SHI) with an estimated incidence of 1/40.000 people [175]. It is a heterogeneous condition typically characterized by the coexistence of branchial cysts or fistulae, malformations of the external, middle and inner ears with pre-auricular pits or tags, diverse degrees of HI and renal symptoms (OMIM# 113650). The syndrome is defined as Branchio-Oto syndrome (BO) (OMIM# 602588) in the absence of renal abnormalities [176]. The clinical diagnosis is currently guided by the major and minor criteria of the Branchio-oto-renal spectrum disorder (BORSD) defined by Chang et al [177]. There are three genes (*EYA1*; OMIM# 601653; *SIX1*; OMIM# 601205 and *SIX5*; OMIM# 600963) that are known to cause BOR/BO syndrome. Mutations in *EYA1* are the most common causes of BOR/BO syndrome, and more than 200 pathogenic variants have been identified in various populations [175]. Despite these numerous variants reported worldwide, none has previously been identified in the sub-Saharan population with black ancestry. Here we report the first confirmed case of BO syndrome caused by a heterozygous pathogenic missense variant in *EYA1* (c.1286A>G; p.Asp429Gly) in a large Malian family.

6.1.2 Methods

6.1.2.1 Ethical compliance

The study was performed according to the guidelines of the declaration of Helsinki. This study was approved by the institutional ethics committees/institutional review board (IRB) of the Faculty of Medicine and

Dentistry of Bamako, Mali (N°2020/129/CE/FMOS/FAPH), the University of Cape Town, Cape Town, South Africa (HREC REF 691/2020) and Columbia University (IRB-AAAS2343). Informed consent and assent for minor participants were obtained prior to their enrolment in this project, including the permission to publish photographs.

6.1.2.2. Participants' recruitment

All participants included in this study were recruited at the Department of Neurology of the Teaching Hospital of Point G, Bamako, Mali. After a record of the medical and family history followed by a pedigree description, all participants were carefully evaluated by medical geneticists (AW, GL) for the description of dysmorphological signs, and ENT specialists (MK) for an otological examination including otoscopy, Pure Tone Audiometry (PTA) for air and bone conduction. HI of common acquired causes was ruled out based on the medical history and otoscopy results. The degree of HI was classified according to the recommendation number 02/1 of the Bureau International d'Audiophonologie (BIAP), Belgium (www.biap.org). Moreover, kidney morphology and function were investigated through renal ultrasound and serum creatinine levels. The Branchio-Oto-Renal spectrum Disorders (BORSD) criteria were used for the clinical diagnosis [177].

6.1.2.3 Molecular analysis

Genomic DNA was extracted from peripheral blood using the QIAGEN Gentra Puregene Blood DNA Kit C (Germantown, MD), following the manufacturer protocol, in the laboratory of neurogenetics of the teaching hospital of Point G, Bamako, Mali.

6.1.2.4 Whole Exome Sequencing

We initially sent DNA samples of individuals II.5 and II.6 (Figure 1A) for whole exome sequencing at Omega Bioservices (Norcross, GA, USA). The library preparation was performed with an Illumina Nextera Rapid Capture Exome Kit® (Illumina, San Diego, CA, USA) following the manufacturer's instructions, and the resulting libraries were hybridized with a 37 Mb probe pool to enrich exome sequences [173]. Sequencing was performed on an Illumina HiSeq 2500 sequencer using the pair-end 150 bp run format. Sequence data were processed using the Illumina DRAGEN Germline Pipeline v3.2.8. Briefly, high-quality reads were aligned to the human reference genome GRCh37/hg19 using the DRAGEN software version 05.021.408.3.4.12 and, after sorting and duplicate marking using Picard, variants were called, and individual genomic variant call format (gvcf) files were generated using the genome analysis toolkit (GATK) software v4.0.6.0 [178]. Joint single nucleotide variant (SNV) and Insertion/Deletion (Indel) variant calling was also performed using GATK. The sex of the two individuals undergoing exome sequencing were verified using plinkv1.9 [179]. Familial relationships for these two family members were verified via Identity-by-Descent sharing (plinkv1.9) and the Kinship-based INference for Gwas (KING) algorithm [179, 180].

6.1.2.5 Annotation and Filtering Strategy

ANNOVAR was used for variants filtering and annotation (https://annovar.openbioinformatics.org/) and custom scripts. Variants were first prioritized based on the inheritance model, considering an AD mode of inheritance. Subsequently, rare variants with a minor allele frequency (MAF) <0.0005 in all populations of the genome aggregation database (gnomAD)

were retained. Known pathogenic HI variants listed in ClinVar were also retained, regardless of their frequencies. dbNSFP v3.0 was used to evaluate missense variants, with 17 bioinformatic tools predicting the deleterious effects of the identified variants. We used several coding variants prediction tools including Sorting Intolerant from Tolerant (SIFT), polymorphism phenotyping v2 (PolyPhen-2) × 2, Mutation Assessor, the likelihood ratio test (LRT), Mendelian clinically applicable pathogenicity (M-CAP) score, Rare Exome Variant Ensemble Learner (REVEL), MutPred, protein variation effect analyzer (PROVEAN), MetaSVM, and MetaLR, while MutationTaster, Eigen, Eigen-PC, functional analysis through Hidden Markov models (FATHMM-MKL), combined annotation dependent depletion (CADD) score, and deleterious annotation of genetic variants using neural networks (DANN) score were used to annotate variants [181].

Adaptive boosting (ADA) and random forest (RF) scores derived from dbscSNV v1.1 were used to predict the deleterious effect of variants within splicing consensus regions (−3 to +8 at the 5′ splice site and −12 to +2 at the 3′ splice site). We used phyloP, Genomic Evolutionary Rate Profiling (GERP), SiPhy, and phastCons scores to estimate the evolutionary conservation of the nucleotides and amino acid residues at which the variants occurred [182]. The hereditary hearing loss homepage (HHL), online Mendelian inheritance in man (OMIM), human phenotype ontology (HPO), and ClinVar databases were used to determine if there were any existing associations between the identified variants and genes and HI. Candidate variants were considered when: (1) they occurred in known HI genes (and genes expressed in the inner ear); (2) they had a predicted effect on protein function or pre-mRNA splicing (nonsense, missense, start-loss, frameshift,

splicing, start-loss, etc.); and (3) they co-segregated with the phenotype within the family [173].

6.1.2.6 Sanger Sequencing

While exome sequencing was being conducted, Sanger sequencing was performed for all the available family members II.4, II.6, II.7, II.8, III.1, III.2, III.3, III.4, III.5 (Figure 1A). Primers to target coding exons including splice sites of the EYA1, SIX1 and SIX5 genes (Table S3) were validated using NCBI BLAST and ordered through Integrated DNA Technologies (IDT DNA, Coralville, IA, USA). The optimal annealing and extension temperatures for the PCR were 58°C and 70 °C for 30 s and 40 s. PCR-amplified DNA products were Sanger sequenced using a BigDyeTM Terminator v3.1 Cycle Sequencing Kit and an ABI 3130XL Genetic Analyzer® (Applied Biosystems, Foster City, CA, USA) in the Division of Human Genetics, University of Cape Town, South Africa and the Neurogenetics Branch, NINDS, NIH (Bethesda, MD). Sequencing chromatograms were manually examined using FinchTV v1.4.0, and aligned in UGENE v34.0 to the EYA1, SIX1 and SIX5 reference sequences (ENSG: ENST00000340726.8; ENST00000645694.3 and ENST00000317578.7; retrieved from Ensembl browser), respectively.

6.1.2.7 Evolutionary Conservation of Amino Acids and Secondary Structure Analysis

We performed a multiple sequence alignment (MSA) of human EYA1 gene with non-human similar proteins to provide more evidence on the evolutionary conservation of the amino acid residue at which our candidate missense variant occurred. A PSI-BLAST search against the non-redundant protein database of EYA1 was performed. Non-redundant, non-synthetic

EYA1 proteins from all the different species in the 500 BLAST hits were manually retrieved as FASTA files. The MSA was performed using CLUSTAL Omega v1.2.4 [183] and the MSA file was visualized using Jalview v2.10.5 [184]. Furthermore, PSIPRED v4.0 [185] and Swiss-Model [186] were used to assess the secondary structural features of both protein forms. Additionally, the InterPro database was queried via the InterProScan web service [187] to identify domains and potential domain changes for the protein.

6.1.2.8 Protein Modelling

The EYA1 homolog 1 isoform 1 (NM_001370335.1) protein sequence (NP_001357264.1) was retrieved from the NCBI GenePept database in FASTA format and secondary structures were predicted for the wild type as well as mutant sequences using PSIPRED workbench (http://bioinf.cs.ucl.ac.uk/psipred). The three-dimensional (3D) structure of EYA1 G429 mutant protein was modelled using Modeller v10.1 based on the EYA1 wild type protein 3D structure that was retrieved from the AlphaFold protein structure database (https://alphafold.ebi.ac.uk/entry/Q99502). The 3D structure of the mutant protein was then refined using the Seok lab's Galaxy Refine algorithm [188]. The PyMol software [189] was used for hydrogen bond analysis and structure visualization, while the ExPASy Protparam web service (https://web.expasy.org/protparam/) was used to investigate the effect of the mutation on EYA1 physicochemical properties.

6.1.3 Case description

6.1.3.1 Participants Phenotypes

We enrolled eight affected individuals (four males and four females) and one unaffected from a single family. Affected family members presented with HI and other ear deformities with variable expression segregating in an autosomal dominant pattern (Figure 9A). The mean age at diagnosis was 25.8 years (ranging from 8 to 54 years). All the patients had at least one major criteria, fulfilling the BORSD clinical diagnosis criteria. Four of them presented with a mixed type of HI from mild to profound, one had a conductive HI and three had a normal hearing. The onset of the HI was post-lingual in five individuals (n=5/8) and was symmetrical and asymmetrical in two and in three patients, respectively. A branchial fistula was observed in seven patients (87.5%) and absent in one patient (12.5%) (Figure 1B and C). It was bilateral in three patients (42.9%) (n=3/7) and unilateral in four (57.1%) (n=4/7). A preauricular sinus was seen in seven patients (87.5%) and was bilateral in six patients (Figure 9C and D). Only one patient (III.2) had a minor criterion, a preauricular tag in the right side (Figure 9C). Renal ultrasonography and creatinine levels were performed in three patients (II.4, II.6, and III.2), and were all normal. The clinical and laboratory findings are summarized in Table 7.

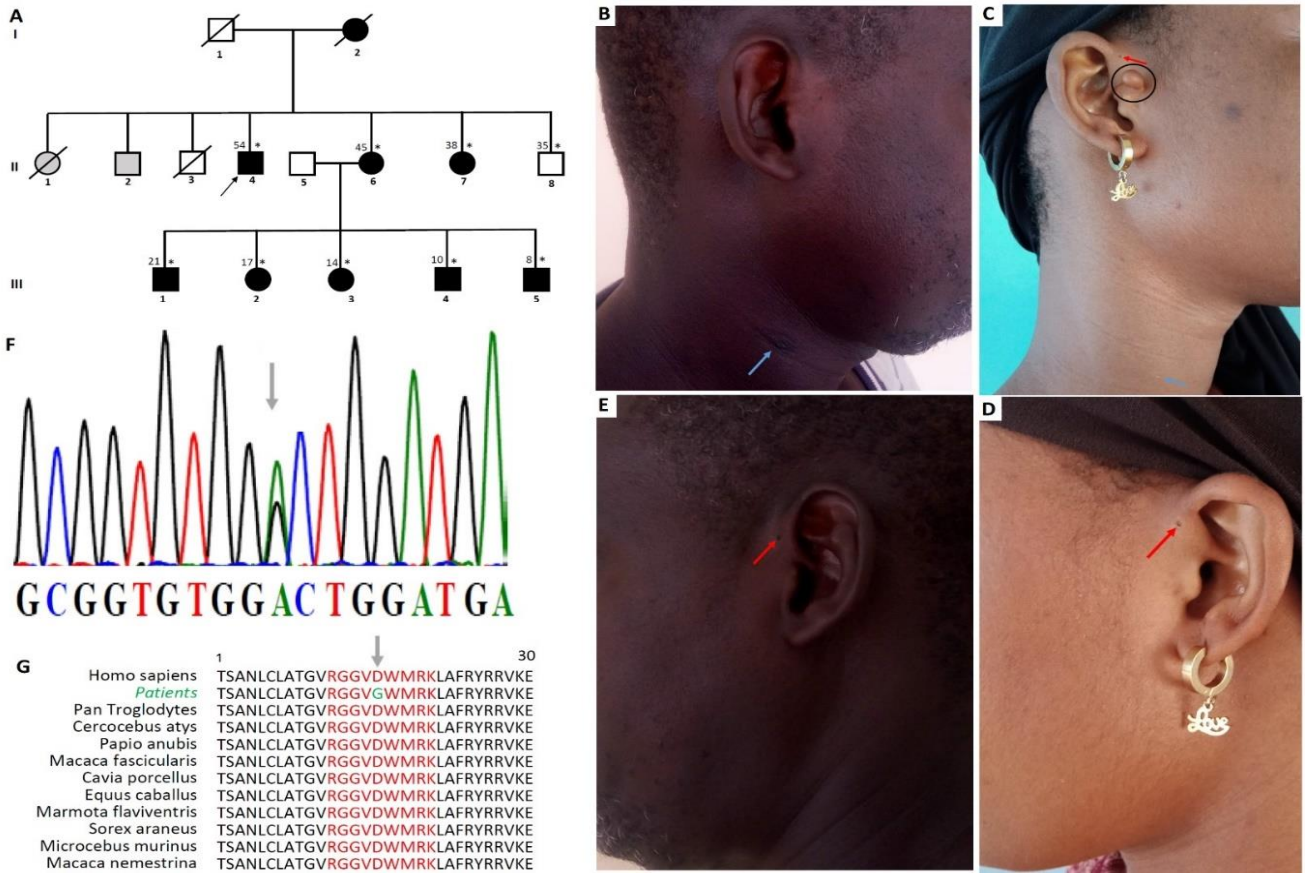


Figure 9: Phenotypic and genotypic features of the family with BO syndrome

(A) pedigree of the family showing an autosomal dominant pattern of inheritance, asterisks indicating those seen in clinic, the black arrow indicates the proband, and shaded individuals are reportedly affected but not seen in clinic, (B and E) images of the patient II.4, with the blue arrow indicating the right branchial fistula and the red arrow indicating preauricular sinus, (C and D) images of the patient III.2, with the red arrow indicating bilateral periauricular sinus, the black circle indicating the preauricular tag and the blue arrow showing the right branchial fistula, (F) chromatogram displaying the c.1286A >G variant indicated by the gray arrow, (G) the Asp429 residue conservation among different species indicated by gray arrow.

Table 7: Clinical and laboratory findings in the patients with BO syndrome

Patients	Age/Sex	Physical signs	Creatinine level	Pure tone audiometry	Renal ultrasonography
II.4	54/M	Left pre-auricular sinus, right branchial fistula	Normal	Moderate symmetrical and mixed HI	Normal
II.6	45/F	Bilateral pre-auricular sinus and branchial fistula	Normal	Severe asymmetrical and mixed HI	Normal
II.7	38/F	Right branchial fistula	NP	Slight asymmetrical conductive HI	NP
III.1	21/M	Bilateral branchial fistula and pre-auricular sinus	NP	Normal	NP
III.2	17/F	Right branchial fistula, bilateral pre-auricular sinus, right ear tag	Normal	Moderate to severe asymmetrical mixed HI	Normal
III.3	14/F	Bilateral pre-auricular sinus	NP	Normal	NP
III.4	10/M	Bilateral branchial fistula and pre-auricular sinus	NP	Slight symmetrical mixed HI	NP
III.5	8/M	Bilateral branchial fistula and pre-auricular sinus	NP	Normal	NP

Note: Age in years. Abbreviations: F, female; HI, hearing impairment; M, male; NP, not performed.

6.1.3.2 Sanger Sequencing Confirmation of Variants

Sequencing of the *SIX1* and *SIX5* genes did not reveal any pathogenic variant. However, sequencing of the *EYA1* gene identified a heterozygous missense variant at position c.1286A>G (NM_000503.6), leading to the amino acid change p.Asp429Gly (Figure 9F). Sequencing of other family members showed that all affected individuals but not the unaffected family member (II.8) carried the variant.

6.1.3.3 Exome Sequencing and Confirmation of Candidate Gene Variant

The average target region coverage was about 225X, with 96.30% of the target region being covered to a depth of 10 X or more. After applying our various filtering criteria described in the methods section, the candidate variant identified through Sanger sequencing was found (EYA1; OMIM# 601653, c.1286A>G (NM_000503.6), p.Asp429Gly). The variant was predicted to be damaging by most of the in-silico tools, including MutationTaster, FATHMM-MKL, Eigen-PC, CADD, and DANN (Table S2). The variant was predicted to occur in conserved region of the genome and was absent from the gnomAD, UK10K, Greater Middle East (GME) variome project databases, as well as the SNP Database (dbSNP) (Table S3). Based on the American College of Medical Genetics' (ACMG) guidelines for the interpretation of sequence variants, the variant was classified as pathogenic (PM1, PS1, PM2, PP2, PP3) [17, 18].

6.1.3.4 Evolutionary Conservation of Amino Acids

A PSI-BLAST search of EYA1 (NP_000494.2) against the non-redundant protein database found the Asp429 residue to be highly conserved across all non-human species retrieved in the top 500 BLAST hits (Figure 9G). As

expected, there was substantial conservation across an extensive amino acid block (on which the variant resides) which forms the thioredoxin/Genetic Diversity Statistics (GST)–C-terminal binding domain. This was consistent with the GERP ++RS (5.44) and PhyloP (7.97) scores for conservation, indicating a strong evolutionary and functional constraint on the mutant position.

6.1.3.5 Protein Modelling: Secondary Structure Analysis and Domain Search

Secondary structure analysis of wild type and mutant EYA1 sequences predicted multiple changes in the mutant protein due to the variant: this included loss of a helix structure at 107AA108, shortening of multiple helices (320DLER324 for instance), and gain or extension of helices and strands in the mutant (Supplementary material Figure S3A and B). These secondary structural changes were apparent in the 3D structures (Figure 10A). The 3D structure of wild type EYA1 indicates three hydrogen bonds formed between Aspartate-429 (D429) and Arginine-432 (R432) and Lysine (K433) (Figure 10B). In the mutant 3D protein, however, the mutant Glycine residue G429 forms a single hydrogen bond with K433 (Figure 10C). Based on the physicochemical properties, the mutant Glycine-429 residue reduces the net charge of the mutant protein by –1, and consequently renders the protein more unstable by increasing its instability index.

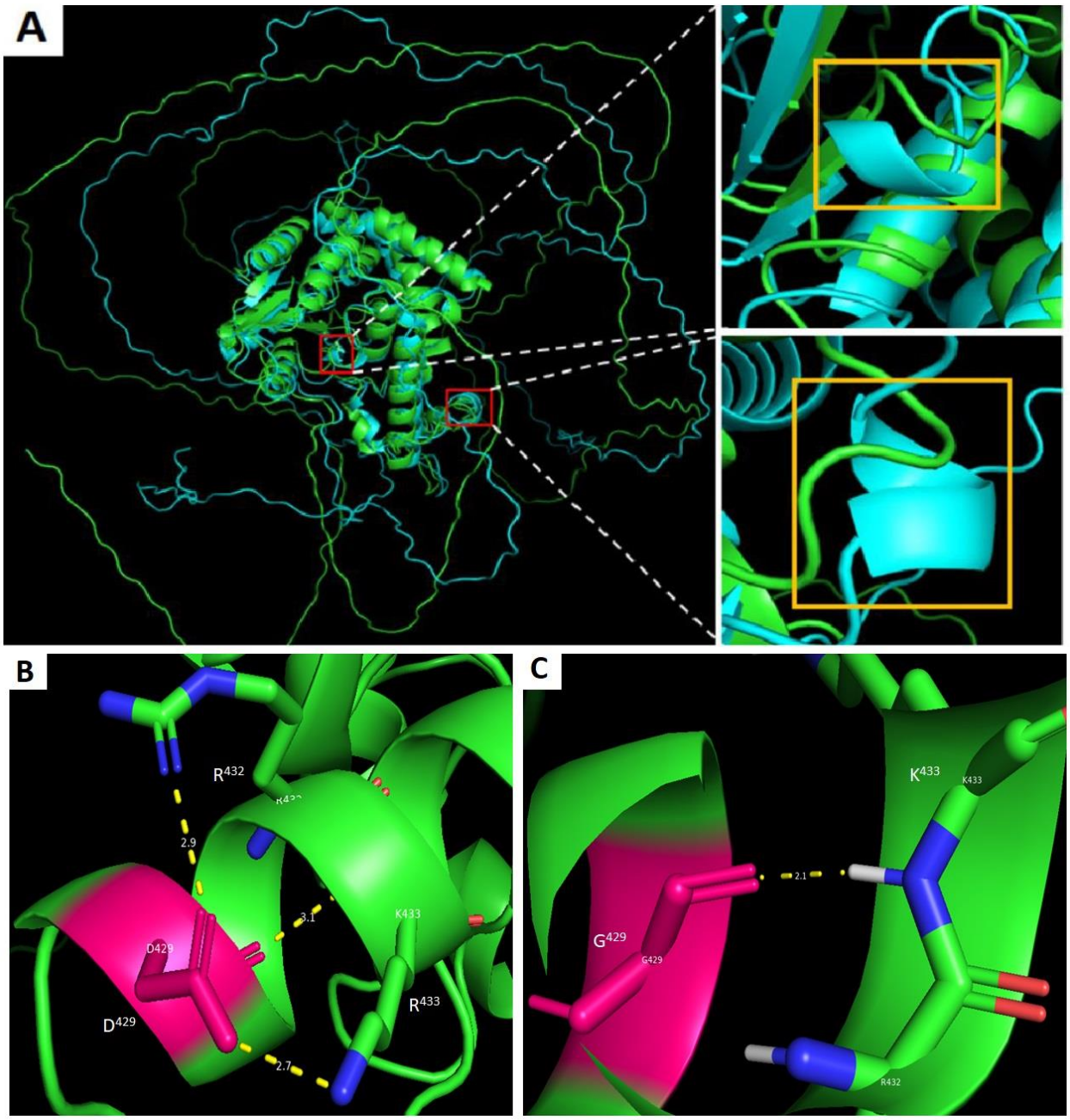


Figure 10: Three-dimensional (3D) secondary analysis and hydrogen bond analysis of EYA1 protein. (NP_000494.2). (a) superimposed three-dimensional structures of EYA1 wild type (green) and mutant (light blue) showing some regions with secondary structural changes (loss of helices in wild type or gain of helices in mutant), (b) wild type and (c) mutant protein structures. The wild type aspartate (D)- 429 residue forms three hydrogen bonds with arginine (R)- 432 and lysine (K)- 433. The mutant glycine (G)- 429 residue forms a single hydrogen bond with lysine (K)- 433.

6.1.4 Discussion

To date, three genes (*EYA1*, *SIX1*, and *SIX5*) have been implicated in BO syndrome. The *EYA1* gene is the most common gene associated with BO, and accounts for ~40% of all cases. *EYA1* encodes for a protein that plays a role in regulating the activity of other genes. The EYA1 protein interacts with several other proteins, including a group known as SIX proteins, to turn on (activate) and turn off (inactivate) genes that are important for normal development. Eyes absent (EYA) is a transcriptional coactivator, and an aspartyl-based protein tyrosine phosphatase that interacts with a broad variety of signaling pathways to regulate the development and homeostasis of organs and tissues such as eye, muscle, kidney, and ear. The variant (p.Asp429Gly) herein analysed is predicted to alter the net charge of the EYA1, impart changes to the secondary structure of the protein, alter hydrogen bond formation, and reduce the stability of the protein. Moreover, the mutant Glycine429 residue which resides in the aspartyl-based protein tyrosine phosphatase active site of the protein is smaller than the wildtype Aspartate residue. Although Aspartate-429 is not directly involved in the EYA1 phosphatase activity, it might be important in EYA1 interaction with its cofactors. Indeed, among its molecular functions, based on Gene Ontology, is metal ion binding, in which Mg2+ is identified as a cofactor. Therefore, the combined effects of net charge change, alteration of active site pocket, secondary structural changes, and instability are expected to affect the overall function of EYA1. The resulting genetic changes affect the development of organs and tissues before birth, which leads to the characteristic features of BOR/BO syndrome [177]. Despite the high number of cases reported elsewhere, only three studies have been reported in Africa,

of which two were genetically confirmed [190-192]. Moreover, these two genetically confirmed cases are not of black ancestry, a Tunisian and an Afrikaner. While the Afrikaner patient carried an *EYA1* variant, the BOR in the Tunisian patients was caused by a variant in *SIX1*. This is probably due to the lack of genetic testing facilities in many sub-Saharan African countries or the attribution of most HI to environmental causes that do not necessitate further genetic investigation. The BOR/BO syndrome is a heterogenous condition with high phenotype variability among individuals even in the same family as seen in our study. The phenotype found here is similar to what Namba et al. reported [193] and different from other studies such as the one from Clarke et al. [190], confirming the heterogeneity of the disease. The mixed HI was the most common type of HI seen in this study, corroborating what was reported in patients with BOR/BO syndrome in the literature [194]. In addition to this, the absence of renal morphological and functional abnormalities sustains the diagnosis of BO syndrome as reported in other studies [193]. This condition is predominantly inherited in an AD manner (~90%) as observed in the family reported here. Hundreds of *EYA1* variants were reported in numerous families around the world. Interestingly, our report represents the third, worldwide, and the first in Africa of this variant (p.Asp429Gly). This variant was previously reported in a family from the United States with multiple affected individuals and in a sporadic case from Japan [193, 195]. It would be interesting to further investigate the ancestry of the American family as most African Americans are of West African descent.

6.1.5 Conclusion

We identified a monoallelic variant in the *EYA1* gene in a Malian family with BO syndrome. It is the first time the identified variant has been reported in Africa, and the third time worldwide. With the decreasing cost of exome sequencing, genetic and genomic studies of the African population could identify more HI associated variants or genes which will improve our understanding of the pathophysiology of this condition.

Conflict of interest: The authors declare that they have no conflict of interest.

Acknowledgement: The study was funded by the National Institute of Health (NIH)/ National Institute of Neurological Disorders and Stroke NINDS grant U01HG007044 to GL; Wellcome Trust, grant 107755Z/15/Z AW (co-applicants); NIH/NHGRI grant U01-HG-009716 to AW; the African Academy of Science/Wellcome Trust, grant number H3A/18/001 to AW; and NIH/National Institute of Deafness and other Communication Disorders (NIDCD) grants R01 DC01165, R01 DC003594, and R01 DC016593 to SML. We thank Dr. Kenneth H. Fischbeck for his valuable supports.

Supporting information: the following additional supporting information were submitted online version of the article at the publisher's website (Table S2, S3, S4 and Figure S3)

6.2 Whole Exome Sequencing Reveals Known and Candidate Gene in NSHI in Mali

Abstract

Background: Hereditary hearing impairment (HHI), is a highly genetically heterogeneous disorder, caused by more than 120 genes to date. HHI is divided into syndromic HI (SHI) and non-syndromic (NSHI). Genetics of HI have extensively been studied in human, mostly in Caucasian population. However, rare studies that have focused on HI in sub-Saharan African populations showed that there is no significant contribution of the common HI genes in SSA. These genes and others associated with HI are largely unknown in Mali.

Methods: All patients were examined by a multidisciplinary team including ENT specialists, medical geneticists and neurologists. Informed consent was obtained from all participants and the study protocol was approved by the Ethics Committee of faculty of Medicine and Dentistry, Bamako, Mali and the University of Cape Town, South Africa. Only families that segregated at least two affected individuals with HI were included. Carefully clinical assessment was done and HI was confirmed by Pure Tone audiometry. DNA was extracted from peripheral blood. Whole exome sequencing (WES) was performed in selected probands then confirmed by Sanger sequencing. The variants were checked in other family members for segregation analysis.

Results: A total of twenty individuals (9 males and 11 females) belonging to eight families were found to be affected. The HI was congenital, profound and sensorineural in all patients. The consanguinity was seen in six families and the inheritance pattern was autosomal recessive in all. WES identified a novel compound heterozygous variants in four families located in OTOGL

[c.209-9C>G and c.5685C>A; p.(D1895E)], CDH23 [c.646C>G; p.(L216V), PJVK [c.461T>G; p.(V154G)], TMC1 (c.2003+2T>C) genes. A candidate gene UBFD1 [c.58G>A; p.(E20K)] was identified in one family. Additionally, known variants in three families including two in MYO15A [c.6331A>T, p.(N2111Y) and c.8158G>A, p.(D2720N)] and one in MYO7A [c.3945C>A; p.(C1315X)] were identified. These variants segregate with the phenotype and were found to be damaging with in silico tools and protein modelling.

Conclusion: This study reports the first genetically confirmed cases of NSHI in Mali and highlights the genetic heterogeneity of the African population and suggests novel gene discovery. The data will contribute to the understanding of the pathobiology of HI, globally.

Key-words: non-syndromic hearing impairment, novel variant, candidate gene, Mali, Africa

6.2.1 Introduction:

Hearing Impairment is the most common sensory disturbance which is caused by genetic and non-genetic factors. Several studies have highlighted that environmental factors remain the leading aetiology of HI in sub-Saharan African populations [1, 5]. In developed countries, more than 50% of congenital HI are due to genetic causes and is broadly divided into SHI and NSHI. SHI associates other organ symptoms and abnormalities, and NSHI whereas HI is isolated. NSHI is the most prevalent subtype of hereditary HI (HHI) and represents at least 70% of cases, while 30% are SHI. Most recessive NSHI are pre-lingual severe-to-profound, and dominant NSHI are post-lingual and progressive [196]. Genetics of HI are substantially studied mostly in the population with Caucasian ancestry where at least one half of NSHI are caused by mutations in the gap junction protein, beta-2 (GJB2, OMIM: 220290) and gap junction protein, beta-6 (GJB6, OMIM: 604418) genes [19, 197-200]. To date, more than one hundred mutations have been identified in the *GJB2* gene, of which the c.35delG variant is found with a highest frequency in the Caucasian population while c.235delC is highest in Southeast Asians [36]. Recent studies have shown the sparsity of the genetic epidemiology of HI from across diverse populations [36, 201, 202]. Moreover, a cohort study found *GJB2*, *SLC26A4* (OMIM: 605646), and mtDNA 12S rRNA as the most common genes while another study reported *MYO15A* (OMIM: 602666), *SLC26A4*, *USH2A* (OMIM: 608400), *MYO7A* (OMIM: 276903), *MYO6* (OMIM: 600970) and *TRIOBP* (OMIM: 609761) genes in NSHI [36, 203]. Until recently, genetic studies of HI were scant in Africa and the rare studies conducted reported that genetic causes of HI in populations with African ancestry seem to be different from those of non-

Africans populations. In addition, several other studies reported the poor contribution of the most prevalent HI genes in SSA [10, 19, 204, 205]. However, an exception is noted in Ghana and more recently in Senegal where mutations in *GJB2* gene have been found with a higher prevalence [23, 206, 207]. Like in other African countries, the genetics of HI has not been studied in Mali likely due to the lack of dedicated genetic investigation facilities and the heavy burden of the infectious diseases [23, 64]. The common and rare genes causing HI have not been investigated in Mali where the consanguinity and fertility rate are high in the general population [40]. In this pioneering study, we investigated HI cases with strong evidence of having genetic causes using WES approach.

6.2.2 Methods

6.2.2.1 Ethical compliance: This study was conducted in full accordance with the Helsinki declaration. Institutional ethical approval was obtained from the Faculty of Medicine and Dentistry of Bamako, Mali (N°2020/129/CE/FMOS/FAPH) and the University of Town, Cape Town, South Africa (HREC REF 691/2020). Written informed consent was obtained from all participants and assent for minors under 18 years before their enrolment.

6.2.2.2 Clinical assessment: All participants included in this study were carefully evaluated by medical geneticists, neurologists and ENT specialists. Detailed past medical history and clinical assessment were done through a questionnaire (available upon request). Only families that segregated at least two affected individuals with NSHI were included. We performed Pure Tone Audiometry (PTA) for air and bone conduction. HI was classified according

to the Bureau International of Audio-phonology classification (BIAP): Normal hearing (<20db), slight HI (21-40db), moderate HI (41-70db), severe (71-90db), profound (91-120db) and cophosis (>120db).

6.2.2.3 Molecular analysis

DNA was extracted from 10 ml of peripheral blood following a safety procedure by trained nurses and/or medical residents. DNA extraction was done using QIAGEN Blood DNA Kits following the manufacturer's instructions. WES was performed at Omega Bioservices (Norcross, GA, USA). DNA concentration was quantified using the Visible spectrophotometer NanoDrop™ 2000, Thermo Scientific (Wilmington, DE, USA) in the laboratory of neurogenetics, at the Faculty of Medicine and Dentistry, Mali.

6.2.2.4 Whole exome sequencing

DNA samples of probands and selected relatives from the eight families were sent to Omega Bioservices (Norcross, GA, USA) for exome sequencing (ES). Library preparation was performed with an Illumina Nextera Rapid Capture Exome Kit® (Illumina, San Diego, CA, USA) following the manufacturer's instructions, and the resulting libraries were hybridized with a 37 Mb probe pool to enrich exome sequences. Sequencing was performed on an Illumina HiSeq 2500 sequencer using the pair-end 150 bp run format as described in the section 6.1 of this document.

6.2.2.5 Annotation and Filtering Strategy

For the annotation and prioritization of variants, we used the previously published strategy [173]. Variants were prioritized regarding several considerations including the inheritance pattern, rare and known pathogenic

HI variants listed in ClinVar variants with a minor allele frequency (MAF) < 0.005 (for AR) and <0.0005 (for AD) in all populations of the genome aggregation database (gnomAD) and regardless their frequency. Also, a combination of tools including dbNSFP v3.0 to annotate and other prediction tools for the evaluation of both coding variants and non-coding including CADD score [208]. PhyloP, Genomic Evolutionary Rate Profiling (GERP), SiPhy, and phastCons scores were used to estimate the evolutionary conservation of the nucleotides and amino acid (aa) residues at which the variants occurred on the publicly available Varsome database (https://varsome.com/). Furthermore, the hereditary hearing loss homepage (HHL), online Mendelian inheritance in man (OMIM), human phenotype ontology (HPO), and ClinVar databases were used to determine the relationship between these variants and genes and HI. Finally, three criteria were considered for candidate variants: (1) they occurred in known HI genes (and genes expressed in the inner ear); (2) they had a predicted effect on protein function or pre-mRNA splicing (nonsense, missense, start-loss, frameshift, splicing, start-loss, etc.); and (3) they co-segregated with the HI phenotype within the family as previously described [178].

6.2.2.6 Sanger sequencing

Samples of all available family members were Sanger sequenced for segregation analysis. Primer pairs covering the variants of interest were designed on the free accessible Integrated DNA Technologies platform (https://www.idtdna.com) in our Human Genetics Laboratory, the University of Cape Town in South Africa. All the primer sequences and cycling conditions for these experiments are available upon request. PCR-amplified DNA products were Sanger sequenced using a BigDye^{TM} Terminator v3.1

Cycle Sequencing Kit and an ABI 3130XL Genetic Analyzer® (Applied Biosystems, Foster City, CA, USA) in the Division of Human Genetics, University of Cape Town, South Africa. Sequencing chromatograms were manually checked using FinchTV v1.4.0, and aligned in UGENE v34.0 to the gene references of *MYO15A* (ENST00000647165.2), *MYO7A* (ENST00000409709.9), *CDH23* (ENST00000224721.12), *PJVK* (ENST00000644580.2), *OTOGL* (ENST00000547103.7), *UBFD1* (ENST00000395878.8), and *TMC1* (ENST00000297784.10) respectively.

5.2.2.7 Evolutionary amino acid conservation

We aligned the sequences of the novel variants of genes found here with non-human similar proteins to provide more evidence of amino acid residue conservation. A PSI-BLAST search against the non-redundant protein database was performed. The multiple sequence alignment (MSA) was performed using MEGA X software [209]. Furthermore, the I-tasser homology webserver was used to access the tertiary structural features of both protein forms [210].

3D protein structure prediction for Functional Characterization of Novel Variants

Molecular dynamic (MD) simulations were conducted to assess the effect of novel variants on protein function. The amino acids sequence was obtained from UniProt PJVK/DFNB59, (https://www.uniprot.org/uniprot/Q0ZLH3), and *UBFD1* (https://www.uniprot.org/uniprot/O14562). The tertiary structure of the *PJVK* and *UBFD1* genes was generated using the I-tasser homology web server [210]. All MD simulations were conducted with the GROMACS package, version 5.6. [211-214] using Amber (AMBER99SB-ILDN) force

field [215]. The system was solvated in a dodecahedron box of tip4p water. The temperature and pressure were maintained at 300 K using the Parrinello-Donadio-Bussi V-rescale thermostat [216] and a pressure of 1 bar using the Berendsen barostat [217]. The short-range nonbonded interactions were modelled using Lennard Jones potentials. The long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) algorithm [218, 219]. The LINCS algorithm was used to constrain hydrogen bond lengths [33]. Then the velocities were assigned according to the Maxwell-Boltzman distribution at 300 K. The equilibration of the structure NVT (constant Number of particles, Volume, and Temperature) and NPT (constant Number of particles, Pressure, and Temperature) for 10 ns each. The MD production simulation was run for 50 ns for each structure.

6.2.2.8 Splice site prediction

The effect of splicing site mutation was provided by the CADD score [34]. Then, we used the freely available online tool (Fruitfly, http://www.fruitfly.org/seq_tools/splice.html) to evaluate the effect of 2000+2T>C on the *TMC1* gene.

6.2.3 Results

6.2.3.1 Clinical data

A total of twenty individuals (9 males and 11 females) belonging to eight families were enrolled. HI was congenital in most cases, and the mean age at diagnosis was 10.65 years ranging from 3 to 29 years. Affected individuals had a prelingual, bilateral and profound NSHI. Parental consanguinity was seen in 75% of families (n=6) and the inheritance pattern was consistent with autosomal recessive (Figure 11A-H).

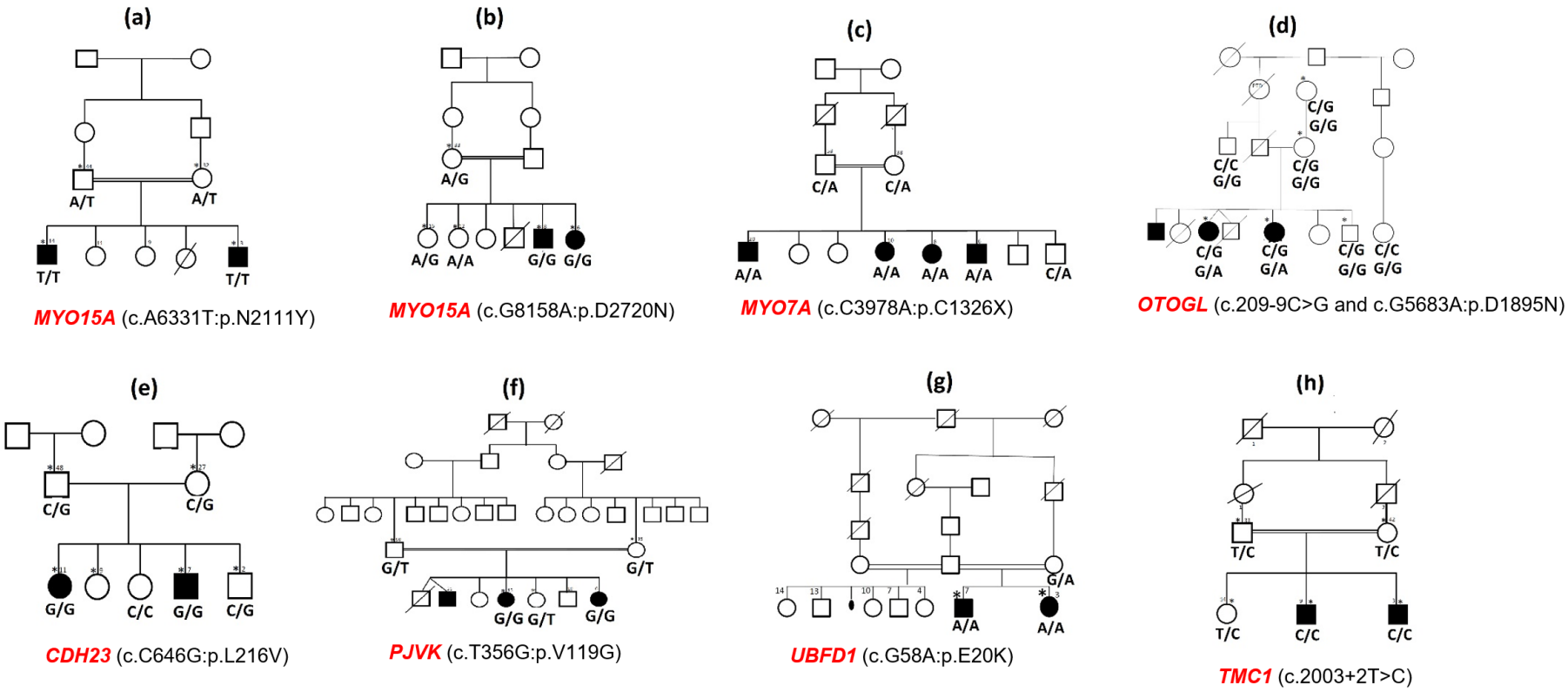


Figure 11: Pedigree and genotypic profile of patients with NSHI. Inheritance pattern consistent with autosomal recessive and the consanguinity reported in six families (a-c, f-h). Asterisk stands for those seen in clinic and capital letters for genotype profile and the variants.

6.2.3.2 Whole exome sequencing genes and variants identification

The average target region coverage was about 225×, with 96.30% of the target region being covered to a depth of 10 X or more. After applying various filtering criteria described in the methodology section. We retained pathogenic and likely pathogenic variants. WES performed in probands identified a known homozygous pathogenic missense variant in *MYO15* gene [c.A6331T; p.(N2111Y), c.G8158A; p.(D2720N), (OMIM: 602666)] in two unrelated families, and a pathogenic nonsense variant in *MYO7A* gene [c.C3978A; p.(C1326X); (OMIM: 276903)], and a novel compound heterozygous variant in *OTOGL* gene [c.209-9C>G and c.C5685A: p.(D1895E); (OMIM: 614925)], a novel missense variant in *CDH23* gene [c.C646G; p.(L216V); (OMIM: 605516)], *PJVK* gene [c.T461G; p.(V154G); (OMIM: 610219)], and homozygous splicing mutation in *TMC1* gene [c.2003+2T>C; (OMIM: 606706)] in one family respectively. Also, in another family ES revealed a missense pathogenic variant in *the UBFD1* gene [c.G58A: p.(E20K)]. However, *UBFD1* gene is not a known HI-associated gene. This could be a novel candidate gene causing HI. In addition, tissue-specific expression analysis showed that *UBFD1* is similarly expressed in the cochlear hair cells as *GJB2, GJB6*, *MYO15*, *OTOG, OTOGL* and *CDH23* genes according to RNAseq data available on gEAR portal (Figure 12). Several in silico prediction tools showed these variants as deleterious. This study reports the first genetically confirmed cases of NSHI in Mali.

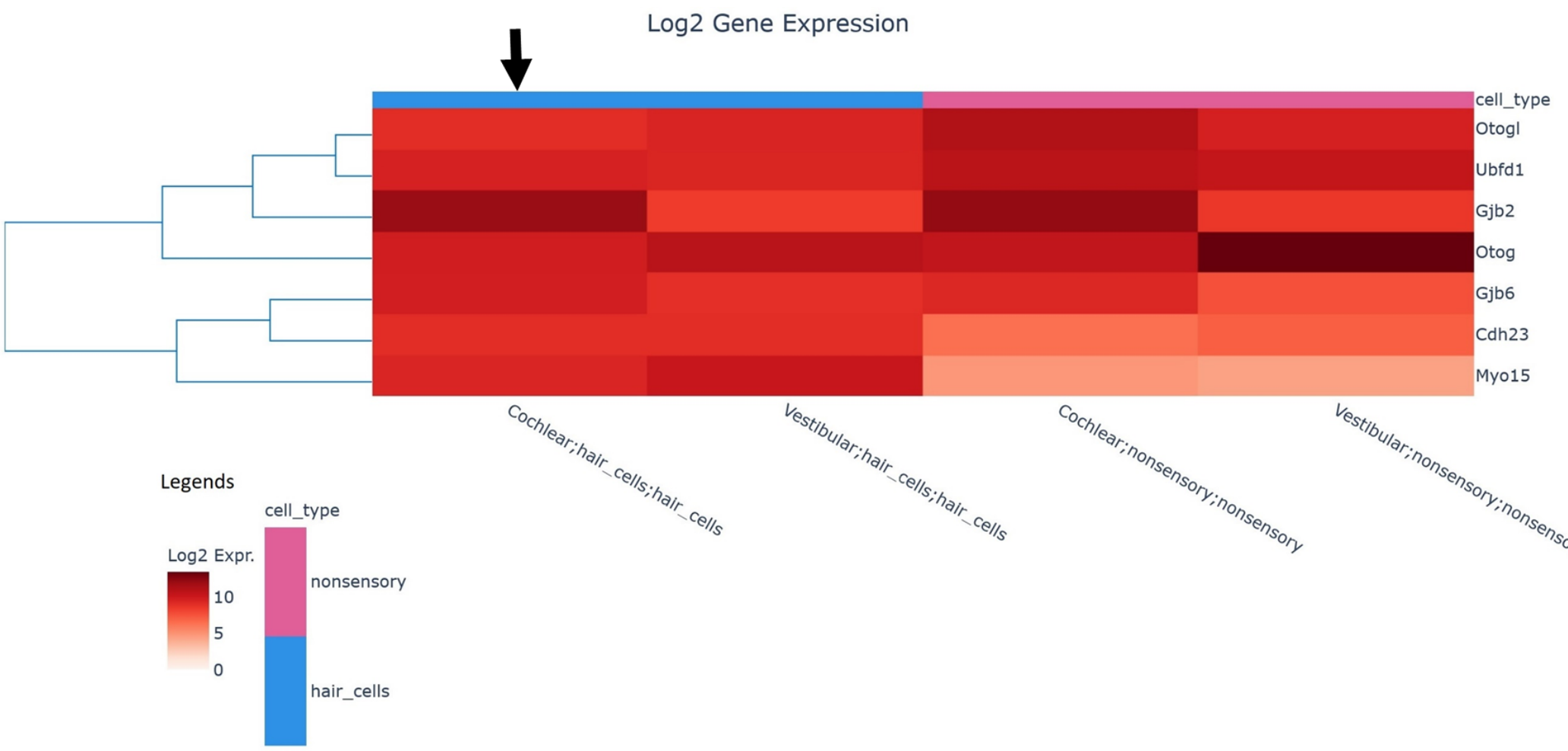


Figure 12: Gene expression analysis. Heatmaps showing the expression level of the known HI genes *GJB2*, *GJB6*, *MYO15*, *CDH23*, *OTOG*, *OTOGL* and *UBDF1* gene. Interestingly, *UBFD1* gene is similarly expressed in cochlear hair cells (white arrow). Map generated from gene expression analysis resources portal (https://umgear.org/).

6.2.3.3 Sanger sequencing confirmation of variants

Following the protocol described in the methods section, variants found using WES were confirmed by Sanger sequencing in all available family members (Figure 13A-H). These variants are predicted to be deleterious using several prediction tools (Table S4)

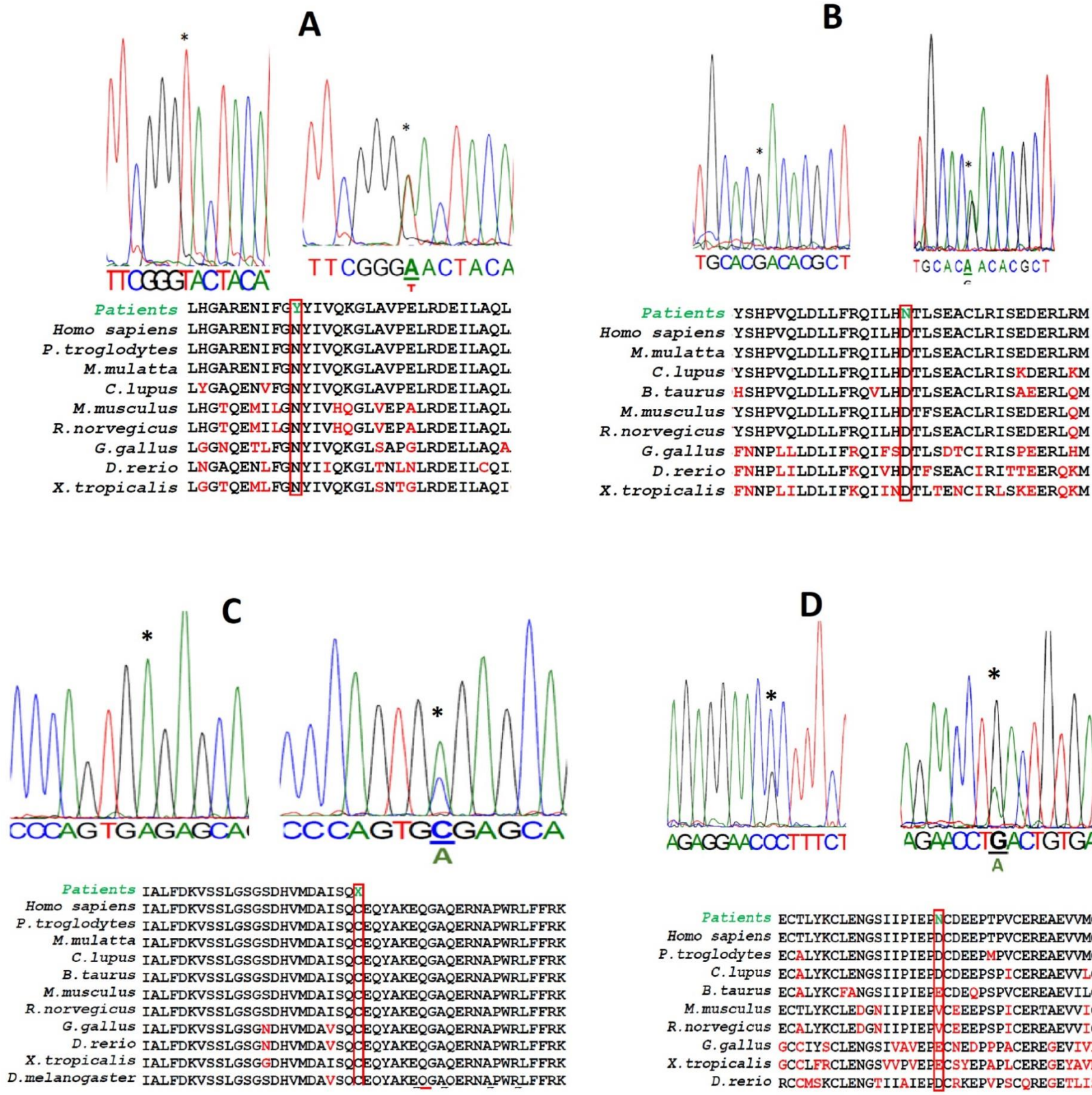


Figure 13: Genetic data of patients with NSHI. Chromatogram and the amino acid conservation across species. **(A)** *MYO15A* gene (c.A6331T:p.N2111Y), **(B)** *MYO15A* gene (c.G8158A:p.D2720N), **(C)** *MYO7A* gene (c.C3978A:p.C1326X), **(D)** *OTOGL* gene (c.209-9C>G and c.G5683A:p.D1895N). Red box indicates the amino acid of interests

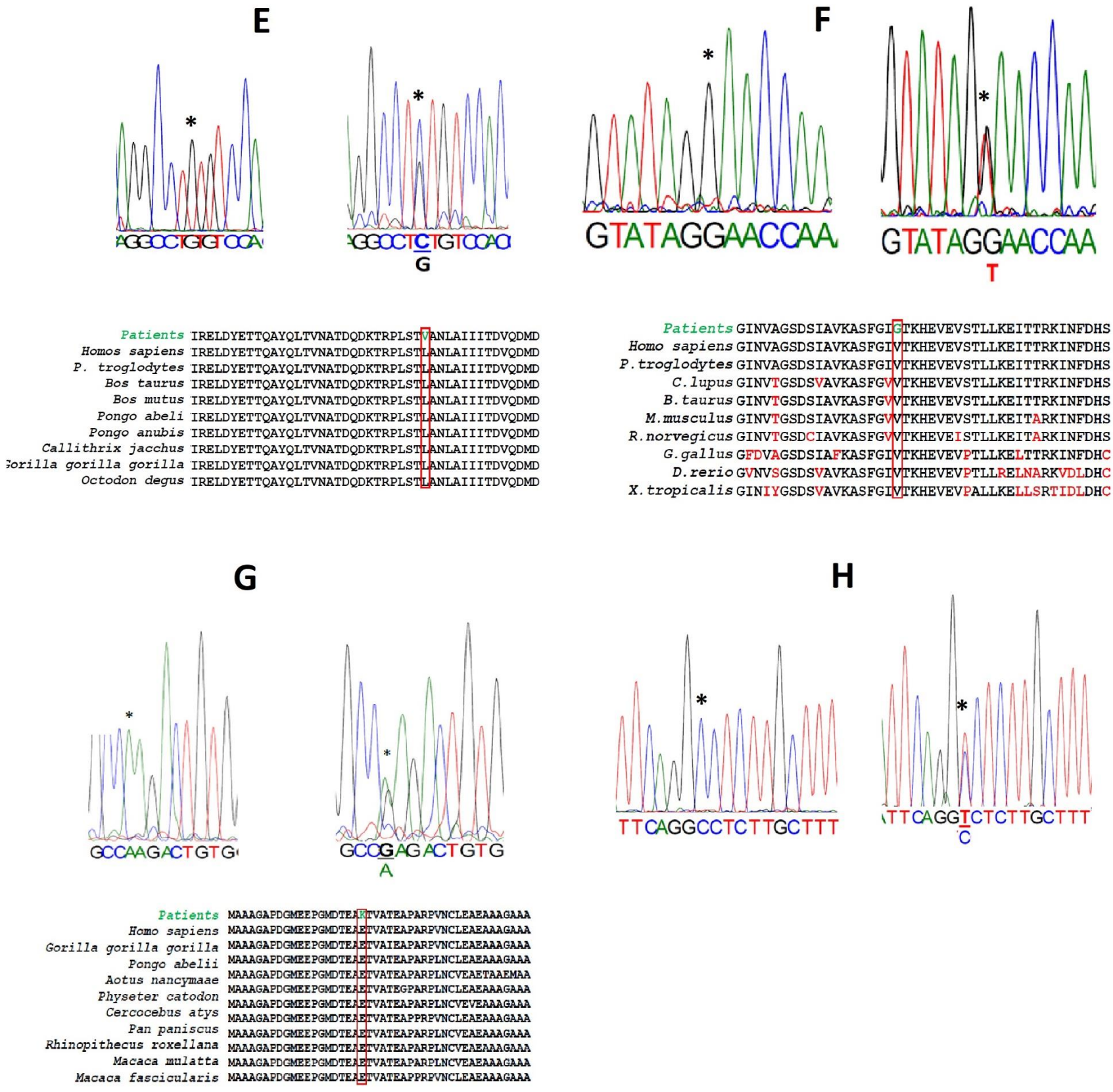


Figure 13 (Continued): Genetic data of patients' with NSHI. (E) *CDH23* gene (c.C646G:p.L216V), **(F)** *PJVK* gene (c.T356G:p.V119G), **(G)** *UBFD1* gene (c.G58A:p.E20K), **(H)** *TMC1* gene: c.2003+2T>C. Red box indicates the amino acid of interests.

6.2.3.4 Evolutionary conservation of amino acid

The NCBI PSI-BLAST search of the genes *MYO15A*, *MYO7A*, *CDH23*, *PJVK*, *OTOGL*, and *UBFD1* against the non-redundant protein database found all residues to be highly conserved across all non-human species retrieved in the top 500 BLAST hits (Figure 13A-G). As expected, there was substantial conservation across an extensive aa block (on which the variant resides). This was consistent with the GERP and PhyloP scores for conservation, indicating a strong evolutionary and functional constraint on the region of the protein.

6.2.3.5 Three-dimensional (3D) protein structure prediction for functional characterization of novel variants

Molecular modeling scheme of the PJVK/DFNB59, (c.T356G:p.V119G) and UBFD1 (c.G58A:p.E20K). A superposition of PJK/DFNB59 protein with mutated residue (Gly119) colored blue and the wild-type residue (val119) colored red (Figure 14A). The non-conservative substitution of uncharged non-polar Gly136 amino acid to uncharged polar amino acid SER136 render the protein structure flexible and impacts binding interactions, stability and conformation of the protein. Zoom of the superposition of PJVK/DFNB59 protein site comparing the wild type (colored red) and mutant (colored blue) (Figure 14B). Protein illustrating the flexibility of the structures. (D) UBFD1 protein shows the mutated residue (Lys20) colored blue and the wild-type residue (Glue20) colored red (Figure 14C). The substitution of the positively charged and hydrophilic amino acid His142 with uncharged polar and hydrophobic amino acid Trp142 could impact binding interactions and the stability of structure and conformation of the protein (Figure 14E, F) Zoom of

the superposition of UBFD1 protein site comparing the wild type (colored red) and mutant (colored blue) and illustrating the flexibility of the structures.

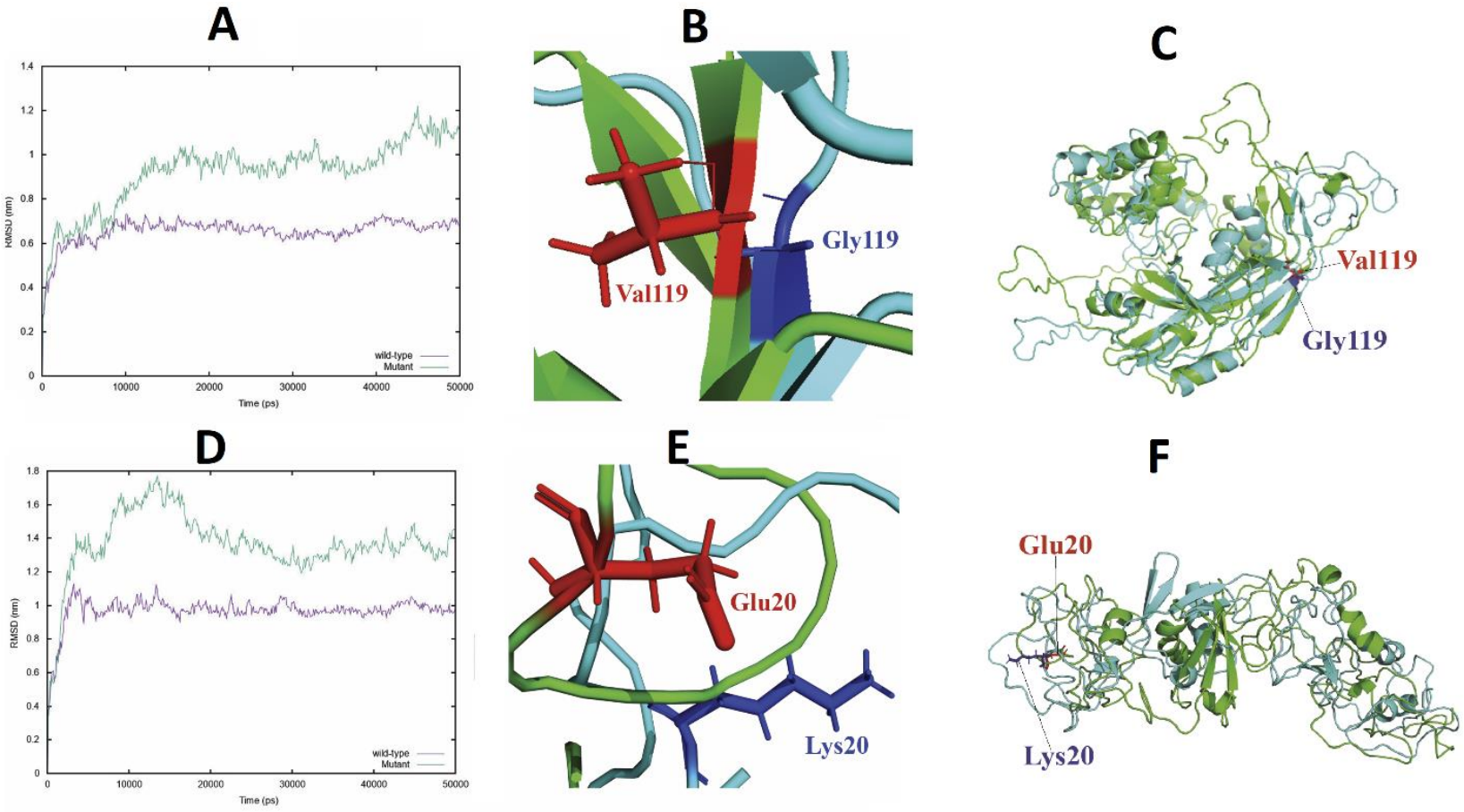


Figure 14: Molecular modelling of the PJVK and UBFD1 proteins. (A) Superposition of PJK/DFNB59 protein with mutated residue (Gly119) colored blue and the wild-type residue (val119) colored red. (B) Zoom of the superposition of PJVK/DFNB59 protein site comparing the wild type (colored red) and mutant (colored blue). (C) Protein and illustrating the flexibility of the structures. (D) UBFD1 protein shows the mutated residue (Lys20) colored blue and the wild-type residue (Glue20) colored red. (E) Zoom of the superposition of UBFD1 protein site comparing the wild type (colored red) and mutant (colored blue). and (F) illustrating the flexibility of the structures

6.2.4 Discussion

We report here the first genetically confirmed families segregating Non-Syndromic Hearing impairment (NSHI) in Mali. We used the WES approach to identify known and candidate novel genes associated with NSHI in the Malian population. WES technique has been widely used to identify variants in several genetic diseases including HI [220-223]. Since the first gene identification of the *GJB2* gene in 1997, several novel genes are being discovered and it is expected that at least 1% of all human genes are involved in the hearing process [8]. Although more than 120 genes are discovered, *GJB2* and *GJB6* still remain the most prevalent HI-associated genes in the population of Caucasian ancestry [15]. The genetic spectrum of deafness-causing genes varies greatly between populations [201]. Mali is a country where consanguinity and interethnic marriage is a common practice and the high rate of fertility offer a great opportunity to investigate these genes as demonstrated in other monogenic diseases [41-43]. By taking advantage of the WES, we successfully identified known and novel variants in the Malian population. The inheritance pattern was consistent with autosomal recessive in all families. This is not surprising since NSHI is inherited in an AR manner in almost 90% of cases [8].

The involvement of the myosin genes (*MYO3A*, *MYO6*, *MYO7A*, *MYO15A*, *MYH14* and *MYH9)* in HI is extensively documented to date [224]. The majority of *MYO15A* variants are associated with a congenital severe-to-profound HI except for *MYO15A* variants in exon 2, which cause a milder auditory phenotype, suggesting a genotype-phenotype correlation of *MYO15A* [225]. *MYO15A* codes for the unconventional myosin proteins and contains 66 exons with 3530 amino acids. Molecularly, it is expressed in the

cochlea hair cells [225]. *MYO15A* (OMIM: 602666) gene variants are the third or fourth most common causes of ARNSHL globally [225]. However, in other studies *MYO15A* gene was reported as the most common gene in NSHI in a cohort from multiple continents and in Egypt [36, 226]. In this study, the *MYO15A* variants p.Asp2720Ans and p.Asn2111Tyr) were found in two unrelated families associated with NSHI. The first variant was seen in a Pakistani family with NSHI and the second in an Indian family [227, 228]. Like previous reports, Chen Chi Wu et al. found the *MYO15A* gene as the fourth gene implicated in NSHI in Taiwanese children with HI, 77% of which had severe to profound NSHI [201, 226, 228]. Nevertheless, its frequency varies from one population to another and a low frequency of 0.89% has been reported in a Japanese cohort with 1120 NSHI [226, 229].

We identified a nonsense variant in the *MYO7A* gene (p.C1326X) in a family with two affected individuals. *MYO7A* gene encodes a protein classified as an unconventional myosin and was identified by Weil et al in 1995 [230]. *MYO7A* is a known HI-gene reported in 1997 in a Chinese family with NSHI and can cause both SHI and NSHI [15]. Several *MYO7A* variants have been reported to cause HI in humans. However, the variant found in our family was reported in one family only from India [36].

PJVK gene is one of the rare gene implicated in HI by altering auditory and neural signals in the inner ear [231]. A novel pathogenic missense variant (p.Val119Gly) in the *PJVK* gene was identified in a family with profound non progressive NSHI. Typically, the *PJVK* gene is associated with congenital NSHI presenting with progressive HI and auditory neuropathy and was seen in several family from different regions [231-233]. To date, about twenty variants are reported to be associated with HI and more than half of them

are missense or nonsense variants [234]. The variant we identified here was not previously reported and was absent from SNP database. The residue Val119 is highly conserved across different species and is located in a conserved domain of the protein. Bioinformatic tools predicted the variant to alter the protein's ability to interact with others and disrupt the protein stability.

CDH23 is a gene that instructs the making of Cadherin protein which includes a large family of intercellular adhesion proteins [235]. Cadherin protein plays a crucial role in making stick cell-cell together and helps to shape the structure of the inner ear [236]. The classical cadherins encode repeated extracellular motifs that provide cell-to-cell adhesion via homophilic, calcium-dependent interactions while the cytoplasmic domains interact with the cytoskeleton via β-catenin [237]. This protein is not only expressed in the inner ear but retinal where the short isoform is found [236]. *CDH23* is implicated in both Usher syndrome and NSHI [238]. The variant (p.L216V) identified in our family has not been reported previously and was absent from the public database. In addition, patients did not present any other symptoms suggestive of Usher syndrome such as speech disorder or early retinitis pigmentosa [238].

OTOGL (OMIMI: 614925) is a recessive HI gene that encode for the Otogelin-like protein containing 2353 amino acids [236]. A compound heterozygote variant has been identified in affected individuals of one family [c.209-9C>G and c.G5683A; p.Asp1895Asn]. The splicing variant was found in heterozygous state in gnomAD with the MAF of 0.0003802 and is predicted to be likely benign. The missense variant (p.Asp1895Asn) is also known variant reported in gnomAD with MAF of 0.006283. Although there

are reported as benign or likely benign, these two variants were found in only affected individuals, and parents and relatives are heterozygous for one of each variant. This implies that the variants can be deleterious when inherited together. Several type of mutations in *OTOGL* gene are reported including missense and compound heterozygous mutations [236].

TMC1 encoding for the transmembrane channel-like protein 1, is a component of mechano-transduction channels in hair cells of the inner ear [239]. Variants in the *TMC1* gene can cause both autosomal dominant and recessive HI in humans [240, 241]. The *TMC1* gene is reported to be the sixth most frequent ARNSHL gene [241]. Diverse variants are reported in *TMC1* gene including substitution, insertion, deletion, and splicing variants. The splice-site variant (c.2003+2T>C) seen in our family is novel and is absent from SNP databases and is predicted to disrupt the splicing site of the gene.

Ubiquitin Family Domain Containing 1 protein is encoded by *UBFD1* gene, located at 16p12.2 is an intracellular protein with 309 amino acids [242]. In mice model, *UBFD1* was shown to be ubiquitously expressed in several tissues, particularly in the ear and auditory system [243, 244]. Furthermore, gene expression analysis using RNAseq (retrieved from gEAR portal) of *GJB2*, *GJB6*, *MYO15*, *OTOG*, *CDH23*, *OTOGL*, *MYO7A* and *UBFD1* showed that *UBDF1* is similarly expressed in cochlear hair cell (https://umgear.org/) [245]. The pathogenic missense variant found in our family is predicted to be damaging and impact the binding interactions and the stability of structure and conformation of the protein. Functional studies are required to confirm its pathogenicity likely using mice models as this have been demonstrated for several novel candidates' genes.

6.2.5 Conclusion

We report a comprehensive genetic study of NSHI in Mali. It shows the efficiency of the WES in identifying genetic variants associated with HI. This is the first study reporting genetically confirmed cases with novel variants and a potential novel gene in Mali. As anticipated, *GJB2* and *GJB6* genes do not seem to contribute extensively to HI in the Malian population. More importantly, the candidate gene identified will be evaluated with functional studies likely animal models. Studying populations with different genetic backgrounds may be of global benefit to first uncover novel gene variants, and novel candidate genes i.e., *UBFD1* for a better understanding of the pathophysiology of HI. This study suggests the need for a large-scale study of the genetic aspects of HI in Mali.

Chapter 7: General discussion, conclusion and perspectives

7.1 General discussion

Charcot-Marie-Tooth disease profile

Charcot-Marie-Tooth disease, also called peroneal atrophy or hereditary motor and sensory neuropathy (HMSN) was first described independently by three physicians in 1886, Jean-Martin Charcot, Pierre Marie and Henri Haward Tooth [246, 247]. CMT is the most prevalent inherited peripheral neuropathy worldwide. Although there are limited data on a national level on CMT in most countries, a crude prevalence of 1/2500 was reported [26]. More than a century after the first description, Valentijn et al. showed that the *PMP22* gene (encoding for the Peripheral Myelin Protein 22) is located within the CMT1A (OMIM: 118220) duplication and concluded that overexpression of this gene is responsible for the CMT1A [248]. CMT1A is the most prevalent subtype of CMT worldwide, representing at least 60% of all genetically diagnosed cases [27]. Although, over one hundred genes have been identified to date, 90% of CMT are caused by mutations in *PMP22*, *MFN2* (encoding for Mutofusin 2 protein), *GJB1* (encoding for Gap Junction Beta 1 protein), and *MPZ* (encoding for Myelin Protein Zero) [27]. Until recently, genetic studies of CMT were less regarded in sub-Saharan African countries [249, 250]. An important part of the previous studies were performed in North Africa where a high prevalence of consanguinity was reported in the general population, increasing the recessive cases [79, 86]. Our study shows that the overall genetic profile of CMT in Africa seems to be different from that of populations with Caucasian ancestry. The common *PMP22* gene was reported in only four families in Africa and no case was

found in Mali. *LMNA* gene (OMIM: 150330) encoding Lamin A protein, *GDAP1* (OMIM: 606598) encoding for Ganglioside-induced differentiation-associated protein 1 and *SH3TC2* (OMIM: 608206) encoding the SH3 domain and tetratricopeptide repeat domain 2 genes were reported to be the most prevalent in Africa. This could be due to the high African genomic diversity or the limited data available from Africa [61, 249]. Like in most African countries, there are not enough data on CMT in Mali despite the high consanguinity and fertility rates, offering a great opportunity for genetic studies [40, 64]. Taking advantage of NGS, the genetic studies of CMT will likely facilitate novel gene discovery as this was shown in other hereditary neurological disorders [41-43]. Based on the genetic diversity seen in the populations with African ancestry, there is no surprise to see epidemiological gap in CMT compared to Eurasian populations. Therefore, large-scale genetic studies in Africa could uncover several genetic variants and contribute to the global effort to fight these untreatable diseases [251].

Candidate gene in CMTX1 and CMT2D cases in Mali

X-linked CMT type 1 is the commonest subtype after CMT1A and CMT2A (OMIM: 609260) [27] globally and is caused by mutations in the *GJB1* gene coding the Gap Junction Beta 1 protein [122]. CMTX1 spectrum is sparse, and it has been reported to be the most prevalent in some cohorts [123]. In this study, we used the CMT panel gene testing to successfully solve three unrelated families presenting CMTX1 phenotype. The variants found here were previously reported in other populations, but this is the first study reporting them in the African population [135, 252, 253]. No phenotype-genotype correlation was established in these families. It is generally admitted that most X-linked traits are less severe in the female when

compared to men, however, a female patient with a severe phenotype carrying the variant Phe235Cys was seen in this study, as previously reported [135]. However, similar to this study a case with a phenotype was reported in an American family.

CMT2D is a rare axonal type of CMT caused by the *GARS* (OMIM: 600287) gene, coding for the glycyl-tRNA synthetase 1 protein [254]. Interestingly, the variant we found has not previously been identified. This finding will contribute to changing the global epidemiology of these rare diseases. We took advantage of the increasing access to NGS to solve the case of this Malian family. As seen in this family, the disease onset is generally in the second decade of CMT2D patients and is inherited dominantly [254]. However, one particularity of this family is that the mother carried the variant but was found clinically unaffected, suggesting an incomplete penetrance as seen in other CMT2D families [254, 255].

Etiologies of Childhood HI in Mali

Hearing impairment is ranked among the most common disabling and common sensory disorders worldwide [1]. HI is classified according to several criteria such as age at onset, severity, mechanism underlying, and aetiology. According to WHO, developing countries are the most burdened area in the world compared to developed countries [1]. Only a very limited number of studies interested in the clinical description of childhood HI in Mali [163, 256]. In this study, we investigated the aetiologies of the childhood HI in the two schools for the deaf in the capital city (Bamako). The onset was mostly in the first decade and sensorineural type was the commonest mechanism like previous studies [10]. Following this study complemented with a literature review, we found that environmental factors are the leading

etiologies of childhood HI in Mali. The common factors identified included meningitis, chronic otitis, and birth complications among others. Similar findings were previously reported in many sub-Saharan African countries [5, 9, 22]. This could be due to the facilities that are not well equipped and the weakness in the healthcare systems in most of our countries.

Whole exome sequencing to identify causative variants in SHI and NSHI in Mali

In this last part of this project, we used the WES approach to seek causative variants in SHI and NSHI. The sequencing revolution in the past few decades has facilitated the identification of candidate variants and new genes in both monogenic and complex traits [222]. The extreme genetic heterogeneity of HI makes the traditional tool of investigation less efficient [257]. The completion of the Human Genome Project (HGP) in 2003 has been a critical step in the genomic revolution [258]. The benefits of the use of WES in gene identification have extensively been documented in several studies [37, 39, 226]. Despite that, WES is not routinely used in the clinical setting, known variants WES has successfully been used to identify new genes in several monogenic diseases including hereditary HI [15, 39, 222].

First, we applied WES in the case of a large family with BO syndrome and found a known but very rare variant in the *EYA1* gene (p.Gly429Asp). This variant was previously reported in Caucasian families making it the third worldwide [193]. We also used several tools to check its deleteriousness with several bioinformatic tools including protein modelling. This finding will contribute to changing the global genetic epidemiology of BO/BOR spectrum disorder [172, 259].

In this second part, we successfully identified known variants and novel candidate gene not previously reported. Of the eight families, known variants were found in Myosin genes (two in *MYO15* and one in *MYO7A*), a novel variant in four genes including *PJVK*, *CDH23*, *TMC1*, *OTOGL*. These are known HI genes, but the variants found here were not reported elsewhere. Moreover, the candidate gene identified (*UBFD1)* is not known in HI. Interestingly, similar findings were reported in a cohort of prescreened *GJB2*-negative families from Ghana with seven novel candidate HI genes identified accounting for 13.7% of their cohort [222].

7.2 Strengths and limitations

This study is a comprehensive study that investigated the genetics of HI and PN in Mali. Our study has several strengths. It provides the most complete clinical description and aetiological profile of HI in school-aged children in Mali. Despite being a common disorder, only a few studies have been conducted on HI in Mali. Also, a high consanguinity rate was seen and preventable diseases such as meningitis, chronic otitis and rubella were found with higher contribution to HI in Mali. In addition, the genetic profile of HI is sparse and is different from those of Caucasian populations with no cases of *GJB2* and *GJB6* found in our cohort. Furthermore, we successfully established the genetic basis of both NSHI and BO syndrome cases, the first genetically diagnosed in Mali. Despite being the most common inherited PN, this is the first study that focused on the clinical and genetic description of certain types of CMT (CMTX1 and CMT2D). These findings will lead to expanding the global epidemiology of these diseases and highlights the potential unique genetic discovery opportunity.

However, our study has some limitations. First, patients with HI included in this study were recruited in only the two schools for the deaf located in the capital city while there are eight administrative regions in Mali. This may not be representative of the whole country. Also, there is no national registry of HI or PN in Mali making the phenotype characterization more difficult. Second, there is an obvious epidemiological gap in HI between Mali and developed countries. The incidence, prevalence and frequency of HI are largely unknown in Mali. Third, *in silico* analysis with protein modeling were not done for all novel variants as some presented computational challenges (*CDH23* gene). Fourth, the novel candidate gene found will need to be validated through functional studies likely animal models to confirm their pathogenicity.

7.3 Conclusion

This study revealed that the most common CMT-genes do not contribute much in CMT cases in Africa. The common genes found in African patients are *LMNA*, *GDAP1* and *SH3TC2*. Subsequently, genetic testing confirmed CMT cases in four unrelated families with known variants in the *GJB1* gene in three families and a novel variant in the *GARS* gene. These CMT cases are the first cases in sub-Saharan African population. As previously reported, HI was found to be associated with CMT and recessive cases of CMT were found to be the most common type in Africa particularly in the northern African population. Our project highlighted the need of conducting large scale of genetic studies of HI and PN in population with African ancestry likely using next generation sequencing techniques. In addition, this study revealed a high contribution of the environmental factors in the occurrence of HI among school aged children in Mali with meningitis being the most

common factor identified. It showed that *GJB2* and *GJB6* genes do not contribute significantly to HI in the Malian population. In the context of multiplex HI, using NGS (gene panel and WES approaches), we successfully identified the causative variants in patients with hereditary HI. Interestingly, in this study we identified novel variants in four HI-genes including *TMC1*, *OTOGL*, *PJVK*, *CDH23* and known variants in *MYO15A* and *MYO7A* genes. More interestingly, a potential candidate gene was identified (*UBFD1)*, which was found to be highly expressed in cochlear hair cells in mouse.

Environmental factors such as meningitis, chronic otitis and birth incidents continue to play a major role in the occurrence of HI in Mali. Despite the limited data on the epidemiology of HI and PN in Mali, we were able to describe the aetiological profile of HI. This evidence argues for the next step forthe urgent actions which needs to be taken to vigorously fight preventable diseases to decrease the burden of HI in our communities. With the high consanguinity rate, the recessive cases of HI are anticipated to be prevalent, and this will offer an opportunity for gene identification. The global profile of CMT in Africa was determined and has been shown to be different from developed countries. More importantly, the well-known and significant CMT genes do not contribute much to CMT in Africa and specifically in Mali. In addition, several novel genetic findings were found in both CMT and HI which contribute to the global action to understand the pathobiology of PN and HI and prepare those patients for potential clinical trials.

7.4 Perspectives

Epidemiological features of HI and PN are highly elusive in Mali. There is need for the establishment of a national registry of HI and PN to the better

characterization and management of HI and PN. Therefore, a nationwide study should be carried out to better understand its global impact on the Malian population. Genetic counselling services are required to inform the population on the role of the consanguinity in recessive conditions to weaken their impacts on HI. The national immunization program against infectious diseases including meningitis and rubella should be reinforced with the collaboration of policymakers and stakeholders. In addition, the novel variants and putative genes found here will lead to the discovery of several other genes. As shown in this study, NGS and more specifically WES can be used as a gold standard in genetic studies. Although there is no efficient curative treatment of hereditary HI and PN, most of the clinical trials that are being developed are gene-specific therapy and such studies are highly important for precision medicine.

References:

1. WHO. *Deafness and hearing loss*. 2021 1 April 2021 [cited 2022 27 April 2021]; Available from: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss.
2. Morton, C.C. and W.E. Nance, *Newborn hearing screening--a silent revolution.* N Engl J Med, 2006. **354**(20): p. 2151-64.
3. Morton, N.E., *Genetic epidemiology of hearing impairment.* Ann N Y Acad Sci, 1991. **630**: p. 16-31.
4. Olusanya, B.O., K.J. Neumann, and J.E. Saunders, *The global burden of disabling hearing impairment: a call to action.* Bull World Health Organ, 2014. **92**(5): p. 367-73.
5. Wonkam, A., et al., *Aetiology of childhood hearing loss in Cameroon (sub-Saharan Africa).* Eur J Med Genet, 2013. **56**(1): p. 20-5.
6. Olusanya, B.O. and A.C. Davis, *Global investment to reduce the burden of hearing loss.* Lancet Glob Health, 2022. **10**(1): p. e4-e5.
7. Grosse, S.D., D.S. Ross, and S.C. Dollard, *Congenital cytomegalovirus (CMV) infection as a cause of permanent bilateral hearing loss: a quantitative assessment.* J Clin Virol, 2008. **41**(2): p. 57-62.
8. Helga V, T.S., D. Smith, *HEREDITARY HEARING LOSS AND ITS SYNDROMES*. 2013, United State of America: Oxford University Press. 749.
9. Adadey, S.M., et al., *Public Health Burden of Hearing Impairment and the Promise of Genomics and Environmental Research: A Case Study in Ghana, Africa.* Omics, 2017. **21**(11): p. 638-646.
10. Wonkam Tingang, E., et al., *Hearing Impairment Overview in Africa: the Case of Cameroon.* Genes (Basel), 2020. **11**(2).
11. Marazita, M.L., et al., *Genetic epidemiological studies of early-onset deafness in the U.S. school-age population.* Am J Med Genet, 1993. **46**(5): p. 486-91.
12. del Castillo, I., et al., *A deletion involving the connexin 30 gene in nonsyndromic hearing impairment.* N Engl J Med, 2002. **346**(4): p. 243-9.
13. Keats, B.J., *Genes and syndromic hearing loss.* J Commun Disord, 2002. **35**(4): p. 355-66.
14. Petit, C., J. Levilliers, and J.P. Hardelin, *Molecular genetics of hearing loss.* Annu Rev Genet, 2001. **35**: p. 589-646.
15. Van Camp G, S.R., *Hereditary Hearing Loss Homepage*. 2022.
16. Kelsell, D.P., et al., *Connexin 26 mutations in hereditary non-syndromic sensorineural deafness.* Nature, 1997. **387**(6628): p. 80-3.

17. Kenneson, A., K. Van Naarden Braun, and C. Boyle, *GJB2 (connexin 26) variants and nonsyndromic sensorineural hearing loss: a HuGE review.* Genet Med, 2002. **4**(4): p. 258-74.
18. Adadey, S.M., et al., *Hearing loss in Africa: current genetic profile.* Hum Genet, 2022. **141**(3-4): p. 505-517.
19. Adadey, S.M., et al., *Connexin Genes Variants Associated with Non-Syndromic Hearing Impairment: A Systematic Review of the Global Burden.* Life (Basel), 2020. **10**(11).
20. Kabahuma, R.I., et al., *Absence of GJB2 gene mutations, the GJB6 deletion (GJB6-D13S1830) and four common mitochondrial mutations in nonsyndromic genetic hearing loss in a South African population.* Int J Pediatr Otorhinolaryngol, 2011. **75**(5): p. 611-7.
21. Kabahuma, R.I., et al., *Spectrum of MYO7A Mutations in an Indigenous South African Population Further Elucidates the Nonsyndromic Autosomal Recessive Phenotype of DFNB2 to Include Both Homozygous and Compound Heterozygous Mutations.* Genes (Basel), 2021. **12**(2).
22. Manyisa, N., et al., *Hearing Impairment in South Africa and the Lessons Learned for Planetary Health Genomics: A Systematic Review.* Omics, 2022. **26**(1): p. 2-18.
23. Dia, Y., et al., *GJB2 Is a Major Cause of Non-Syndromic Hearing Impairment in Senegal.* Biology (Basel), 2022. **11**(5).
24. Giuliani, N., et al., *The audiologic profile of patients with Charcot-Marie Tooth neuropathy can be characterised by both cochlear and neural deficits.* Int J Audiol, 2019. **58**(12): p. 902-912.
25. Eggermann, K., et al., *Hereditary Neuropathies.* Dtsch Arztebl Int, 2018. **115**(6): p. 91-97.
26. Skre, H., *Genetic and clinical aspects of Charcot-Marie-Tooth's disease.* Clin Genet, 1974. **6**(2): p. 98-118.
27. Pipis, M., et al., *Next-generation sequencing in Charcot-Marie-Tooth disease: opportunities and challenges.* Nat Rev Neurol, 2019. **15**(11): p. 644-656.
28. Fridman, V., 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): p. 873-8.
29. Murphy, S.M., et al., *Charcot-Marie-Tooth disease: frequency of genetic subtypes and guidelines for genetic testing.* J Neurol Neurosurg Psychiatry, 2012. **83**(7): p. 706-10.

30. Saporta, A.S., et al., *Charcot-Marie-Tooth disease subtypes and genetic testing strategies.* Ann Neurol, 2011. **69**(1): p. 22-33.
31. Klein, C.J., *Charcot-Marie-Tooth Disease and Other Hereditary Neuropathies.* Continuum (Minneap Minn), 2020. **26**(5): p. 1224-1256.
32. Van Camp G, S.R., *Hereditary Hearing Loss Homepage.* *https://hereditaryhearingloss.org*. retriewed May, 2020.
33. Alzahrani, M., P. Tabet, and I. Saliba, *Pediatric hearing loss: common causes, diagnosis and therapeutic approach.* Minerva Pediatr, 2015. **67**(1): p. 75-90.
34. Kılıç, S., et al., *Comprehensive medical evaluation of pediatric bilateral sensorineural hearing loss.* Laryngoscope Investig Otolaryngol, 2021. **6**(5): p. 1196-1207.
35. Lebeko, K., et al., *Genetics of hearing loss in Africans: use of next generation sequencing is the best way forward.* Pan Afr Med J, 2015. **20**: p. 383.
36. Yan, D., et al., *Spectrum of DNA variants for non-syndromic deafness in a large cohort from multiple continents.* Hum Genet, 2016. **135**(8): p. 953-61.
37. Wonkam, A., et al., *Whole exome sequencing reveals a biallelic frameshift mutation in GRXCR2 in hearing impairment in Cameroon.* Mol Genet Genomic Med, 2021: p. e1609.
38. Wonkam, A., et al., *Whole exome sequencing reveals a biallelic frameshift mutation in GRXCR2 in hearing impairment in Cameroon.* Mol Genet Genomic Med, 2021. **9**(3): p. e1609.
39. Wonkam, A., et al., *Whole exome sequencing reveals pathogenic variants in MYO3A, MYO15A and COL9A3 and differential frequencies in ancestral alleles in hearing impairment genes among individuals from Cameroon.* Hum Mol Genet, 2021. **29**(23): p. 3729-3743.
40. Sangaré, M., et al., *Genetics of low spinal muscular atrophy carrier frequency in sub-Saharan Africa.* Ann Neurol, 2014. **75**(4): p. 525-32.
41. Landouré, G., et al., *A novel variant in the spatacsin gene causing SPG11 in a Malian family.* J Neurol Sci, 2020. **411**: p. 116675.
42. Landouré, G., et al., *Novel mutation in the ATM gene in a Malian family with ataxia telangiectasia.* J Neurol, 2013. **260**(1): p. 324-6.
43. Traoré, M., et al., *Novel mutation in the NHLRC1 gene in a Malian family with a severe phenotype of Lafora disease.* Neurogenetics, 2009. **10**(4): p. 319-23.
44. Callaghan, B.C., R.S. Price, and E.L. Feldman, *Distal Symmetric Polyneuropathy: A Review.* Jama, 2015. **314**(20): p. 2172-81.

45. Rossor, A.M., M.E. Shy, and M.M. Reilly, *Are we prepared for clinical trials in Charcot-Marie-Tooth disease?* Brain Res, 2020. **1729**: p. 146625.
46. DiVincenzo, C., et al., *The allelic spectrum of Charcot-Marie-Tooth disease in over 17,000 individuals with neuropathy.* Mol Genet Genomic Med, 2014. **2**(6): p. 522-9.
47. Lupski, J.R., et al., *DNA duplication associated with Charcot-Marie-Tooth disease type 1A.* Cell, 1991. **66**(2): p. 219-32.
48. Raeymaekers, P., et al., *Duplication in chromosome 17p11.2 in Charcot-Marie-Tooth neuropathy type 1a (CMT 1a). The HMSN Collaborative Research Group.* Neuromuscul Disord, 1991. **1**(2): p. 93-7.
49. Yalcouyé, A., et al., *A novel mutation in the GARS gene in a Malian family with Charcot-Marie-Tooth disease.* Mol Genet Genomic Med, 2019. **7**(7): p. e00782.
50. Aarli, J.A., A.G. Diop, and H. Lochmüller, *Neurology in sub-Saharan Africa: a challenge for World Federation of Neurology.* Neurology, 2007. **69**(17): p. 1715-8.
51. Cortese, A., et al., *Targeted next-generation sequencing panels in the diagnosis of Charcot-Marie-Tooth disease.* Neurology, 2020. **94**(1): p. e51-e61.
52. Chumakov, I., et al., *Polytherapy with a combination of three repurposed drugs (PXT3003) down-regulates Pmp22 over-expression and improves myelination, axonal and functional parameters in models of CMT1A neuropathy.* Orphanet J Rare Dis, 2014. **9**: p. 201.
53. Morena, J., A. Gupta, and J.C. Hoyle, *Charcot-Marie-Tooth: From Molecules to Therapy.* Int J Mol Sci, 2019. **20**(14).
54. Ridler, C., *Peripheral neuropathies: Antisense therapy for Charcot-Marie-Tooth disease?* Nat Rev Neurol, 2018. **14**(2): p. 64.
55. Cooper, A., et al., *APOL1 renal risk variants have contrasting resistance and susceptibility associations with African trypanosomiasis.* Elife, 2017. **6**.
56. Landouré, G., et al., *Hereditary spastic paraplegia type 43 (SPG43) is caused by mutation in C19orf12.* Hum Mutat, 2013. **34**(10): p. 1357-60.
57. Othmane, K.B., et al., *Identification of a new locus for autosomal recessive Charcot-Marie-Tooth disease with focally folded myelin on chromosome 11p15.* Genomics, 1999. **62**(3): p. 344-9.
58. Tingang Wonkam, E., et al., *GJB2 and GJB6 Mutations in Hereditary Recessive Non-Syndromic Hearing Impairment in Cameroon.* Genes (Basel), 2019. **10**(11).

59. Wonkam, A., *Investigating the missing heritability of fetal haemoglobin level in Africa.* Br J Haematol, 2020. **191**(5): p. 668-670.
60. Lebeko, K., et al., *Targeted genomic enrichment and massively parallel sequencing identifies novel nonsyndromic hearing impairment pathogenic variants in Cameroonian families.* Clin Genet, 2016. **90**(3): p. 288-90.
61. Wonkam, A., *Sequence three million genomes across Africa.* Nature, 2021. **590**(7845): p. 209-211.
62. McGuire, A.L., et al., *The road ahead in genetics and genomics.* Nat Rev Genet, 2020. **21**(10): p. 581-596.
63. Birouk, N., et al., *Phenotypical features of a Moroccan family with autosomal recessive Charcot-Marie-Tooth disease associated with the S194X mutation in the GDAP1 gene.* Arch Neurol, 2003. **60**(4): p. 598-604.
64. Landouré, G., et al., *Genetics and genomic medicine in Mali: challenges and future perspectives.* Mol Genet Genomic Med, 2016. **4**(2): p. 126-34.
65. Sohani, Z.N., et al., *Empirical evaluation of the Q-Genie tool: a protocol for assessment of effectiveness.* BMJ Open, 2016. **6**(6): p. e010403.
66. Hoy, D., et al., *Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement.* J Clin Epidemiol, 2012. **65**(9): p. 934-9.
67. Kandil, M.R., et al., *A community-based epidemiological study of peripheral neuropathies in Assiut, Egypt.* Neurol Res, 2012. **34**(10): p. 960-6.
68. Aiyesimoju, A.B., et al., *Hereditary neurodegenerative disorders in Nigerian Africans.* Neurology, 1984. **34**(3): p. 361-2.
69. LeGuern, E., et al., *Homozygosity mapping of an autosomal recessive form of demyelinating Charcot-Marie-Tooth disease to chromosome 5q23-q33.* Hum Mol Genet, 1996. **5**(10): p. 1685-8.
70. Kessali, M., et al., *A clinical, electrophysiologic, neuropathologic, and genetic study of two large Algerian families with an autosomal recessive demyelinating form of Charcot-Marie-Tooth disease.* Neurology, 1997. **48**(4): p. 867-73.
71. Meggouh, F., et al., *The first de novo mutation of the connexin 32 gene associated with X linked Charcot-Marie-Tooth disease.* J Med Genet, 1998. **35**(3): p. 251-2.
72. Bouhouche, A., et al., *A locus for an axonal form of autosomal recessive Charcot-Marie-Tooth disease maps to chromosome 1q21.2-q21.3.* Am J Hum Genet, 1999. **65**(3): p. 722-7.

73. Barhoumi, C., et al., *Linkage of a new locus for autosomal recessive axonal form of Charcot-Marie-Tooth disease to chromosome 8q21.3.* Neuromuscul Disord, 2001. **11**(1): p. 27-34.
74. Baxter, R.V., et al., *Ganglioside-induced differentiation-associated protein-1 is mutant in Charcot-Marie-Tooth disease type 4A/8q21.* Nat Genet, 2002. **30**(1): p. 21-2.
75. De Sandre-Giovannoli, A., et al., *Homozygous defects in LMNA, encoding lamin A/C nuclear-envelope proteins, cause autosomal recessive axonal neuropathy in human (Charcot-Marie-Tooth disorder type 2) and mouse.* Am J Hum Genet, 2002. **70**(3): p. 726-36.
76. Kakar, R., et al., *Clinical and genetic analysis of CMT1B in a Nigerian family.* Muscle Nerve, 2003. **27**(5): p. 628-30.
77. Azzedine, H., et al., *Mutations in MTMR13, a new pseudophosphatase homologue of MTMR2 and Sbf1, in two families with an autosomal recessive demyelinating form of Charcot-Marie-Tooth disease associated with early-onset glaucoma.* Am J Hum Genet, 2003. **72**(5): p. 1141-53.
78. Chaouch, M., et al., *The phenotypic manifestations of autosomal recessive axonal Charcot-Marie-Tooth due to a mutation in Lamin A/C gene.* Neuromuscul Disord, 2003. **13**(1): p. 60-7.
79. Tazir, M., et al., *Phenotypic variability in autosomal recessive axonal Charcot-Marie-Tooth disease due to the R298C mutation in lamin A/C.* Brain, 2004. **127**(Pt 1): p. 154-63.
80. Azzedine, H., et al., *Spine deformities in Charcot-Marie-Tooth 4C caused by SH3TC2 gene mutations.* Neurology, 2006. **67**(4): p. 602-6.
81. A Bösenberg, K.L., *Anaesthesia and Charcot-Marie-Tooth Disease.* Southern African Journal of Anaesthesia & Analgesia, 2006.
82. Onwuewe IO, E.O., Nwabueze AC, *Charcot-Marie-Tooth disease presenting in a young Nigerian Male: A case report.* Journal of College of Medicine, 2007. **12(1)**: p. 35-38.
83. Bouhouche, A., et al., *A novel GDAP1 mutation P78L responsible for CMT4A disease in three Moroccan families.* Can J Neurol Sci, 2007. **34**(4): p. 421-6.
84. Delague, V., et al., *Mutations in FGD4 encoding the Rho GDP/GTP exchange factor FRABIN cause autosomal recessive Charcot-Marie-Tooth type 4H.* Am J Hum Genet, 2007. **81**(1): p. 1-16.
85. Bouhouche, A., et al., *Autosomal recessive axonal Charcot-Marie-Tooth disease (ARCMT2): phenotype-genotype correlations in 13 Moroccan families.* Brain, 2007. **130**(Pt 4): p. 1062-75.

86. Hamadouche, T., et al., *Founder effect and estimation of the age of the c.892C>T (p.Arg298Cys) mutation in LMNA associated to Charcot-Marie-Tooth subtype CMT2B1 in families from North Western Africa.* Ann Hum Genet, 2008. **72**(Pt 5): p. 590-7.
87. Nouioua, S., et al., *Novel mutations in the PRX and the MTMR2 genes are responsible for unusual Charcot-Marie-Tooth disease phenotypes.* Neuromuscul Disord, 2011. **21**(8): p. 543-50.
88. Baudot, C., et al., *Two novel missense mutations in FGD4/FRABIN cause Charcot-Marie-Tooth type 4H (CMT4H).* J Peripher Nerv Syst, 2012. **17**(2): p. 141-6.
89. Boubaker, C., et al., *A novel mutation in FGD4/FRABIN causes Charcot Marie Tooth disease type 4H in patients from a consanguineous Tunisian family.* Ann Hum Genet, 2013. **77**(4): p. 336-43.
90. Mathis, S., et al., *Peripheral myelin protein 22 gene duplication with atypical presentations: a new example of the wide spectrum of Charcot-Marie-Tooth 1A disease.* Neuromuscul Disord, 2014. **24**(6): p. 524-8.
91. S Manyeruke, E.A., C Tinley, , *Macular oedema in Charcot-Marie-Tooth disease.* SA Ophthalmology Journal, 2020.
92. Bird, T.D., J. Ott, and E.R. Giblett, *Evidence for linkage of Charcot-Marie-Tooth neuropathy to the Duffy locus on chromosome 1.* Am J Hum Genet, 1982. **34**(3): p. 388-94.
93. Radhakrishnan, K., M.A. el-Mangoush, and S.E. Gerryo, *Descriptive epidemiology of selected neuromuscular disorders in Benghazi, Libya.* Acta Neurol Scand, 1987. **75**(2): p. 95-100.
94. Osuntokun, B.O., et al., *Neurological disorders in Nigerian Africans: a community-based study.* Acta Neurol Scand, 1987. **75**(1): p. 13-21.
95. Fridman, V., 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): p. 873-8.
96. Hoebeke, C., et al., *Retrospective study of 75 children with peripheral inherited neuropathy: Genotype-phenotype correlations.* Arch Pediatr, 2018. **25**(8): p. 452-458.
97. Tazir, M., et al., *Hereditary motor and sensory neuropathies or Charcot-Marie-Tooth diseases: an update.* J Neurol Sci, 2014. **347**(1-2): p. 14-22.
98. Harding, A.E. and P.K. Thomas, *Autosomal recessive forms of hereditary motor and sensory neuropathy.* J Neurol Neurosurg Psychiatry, 1980. **43**(8): p. 669-78.

99. Nagappa, M., S. Sharma, and A.B. Taly, *Charcot Marie Tooth*, in *StatPearls*. 2021, StatPearls Publishing

100. Landouré, G., et al., *Mutations in TRPV4 cause Charcot-Marie-Tooth disease type 2C.* Nat Genet, 2010. **42**(2): p. 170-4.
101. Hardie, R., et al., *Diaphragmatic weakness in hereditary motor and sensory neuropathy.* J Neurol Neurosurg Psychiatry, 1990. **53**(4): p. 348-50.
102. Romdhane, L., et al., *Consanguinity and Inbreeding in Health and Disease in North African Populations.* Annu Rev Genomics Hum Genet, 2019. **20**: p. 155-179.
103. Stavrou, M., et al., *Genetic mechanisms of peripheral nerve disease.* Neurosci Lett, 2021. **742**: p. 135357.
104. Rossor, A.M., et al., *Clinical implications of genetic advances in Charcot-Marie-Tooth disease.* Nat Rev Neurol, 2013. **9**(10): p. 562-71.
105. Gonzaga-Jauregui, C., et al., *Exome Sequence Analysis Suggests that Genetic Burden Contributes to Phenotypic Variability and Complex Neuropathy.* Cell Rep, 2015. **12**(7): p. 1169-83.
106. Landrum, M.J., et al., *ClinVar: improving access to variant interpretations and supporting evidence.* Nucleic Acids Res, 2018. **46**(D1): p. D1062-d1067.
107. Lek, M., et al., *Analysis of protein-coding genetic variation in 60,706 humans.* Nature, 2016. **536**(7616): p. 285-91.
108. Yalvac, M.E., et al., *AAV1.NT-3 gene therapy increases muscle fiber diameter through activation of mTOR pathway and metabolic remodeling in a CMT mouse model.* Gene Ther, 2018. **25**(2): p. 129-138.
109. Sahenk, Z. and B. Ozes, *Gene therapy to promote regeneration in Charcot-Marie-Tooth disease.* Brain Res, 2020. **1727**: p. 146533.
110. Pareyson, D., *Charcot-marie-tooth disease and related neuropathies: molecular basis for distinction and diagnosis.* Muscle Nerve, 1999. **22**(11): p. 1498-509.
111. Tsai, P.C., et al., *Clinical and biophysical characterization of 19 GJB1 mutations.* Ann Clin Transl Neurol, 2016. **3**(11): p. 854-865.
112. Vance, J.M., *Hereditary motor and sensory neuropathies.* J Med Genet, 1991. **28**(1): p. 1-5.
113. Murakami, T. and Y. Sunada, *Schwann Cell and the Pathogenesis of Charcot-Marie-Tooth Disease.* Adv Exp Med Biol, 2019. **1190**: p. 301-321.
114. Bienfait, H.M., et al., *Phenotype of Charcot-Marie-Tooth disease Type 2.* Neurology, 2007. **68**(20): p. 1658-67.

115. Lu, Y.Y., et al., *Clinical and Genetic Features of Chinese X-linked Charcot-Marie-Tooth Type 1 Disease.* Chin Med J (Engl), 2017. **130**(9): p. 1049-1054.
116. Kleopa, K.A., C.K. Abrams, and S.S. Scherer, *How do mutations in GJB1 cause X-linked Charcot-Marie-Tooth disease?* Brain Res, 2012. **1487**: p. 198-205.
117. Richards, 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): p. 405-24.
118. Totland, M.Z., et al., *Regulation of gap junction intercellular communication by connexin ubiquitination: physiological and pathophysiological implications.* Cell Mol Life Sci, 2020. **77**(4): p. 573-591.
119. Rash, J.E., et al., *Cell-specific expression of connexins and evidence of restricted gap junctional coupling between glial cells and between neurons.* J Neurosci, 2001. **21**(6): p. 1983-2000.
120. Willecke, K., et al., *Structural and functional diversity of connexin genes in the mouse and human genome.* Biol Chem, 2002. **383**(5): p. 725-37.
121. Nelis, E., N. Haites, and C. Van Broeckhoven, *Mutations in the peripheral myelin genes and associated genes in inherited peripheral neuropathies.* Hum Mutat, 1999. **13**(1): p. 11-28.
122. Bergoffen, J., et al., *Connexin mutations in X-linked Charcot-Marie-Tooth disease.* Science, 1993. **262**(5142): p. 2039-42.
123. Shahrizaila, N., et al., *X-linked Charcot-Marie-Tooth disease predominates in a cohort of multiethnic Malaysian patients.* Muscle Nerve, 2014. **49**(2): p. 198-201.
124. Abrams, C.K., *Diseases of connexins expressed in myelinating glia.* Neurosci Lett, 2019. **695**: p. 91-99.
125. Yiu, E.M., et al., *A retrospective review of X-linked Charcot-Marie-Tooth disease in childhood.* Neurology, 2011. **76**(5): p. 461-6.
126. Bahr, M., et al., *Central visual, acoustic, and motor pathway involvement in a Charcot-Marie-Tooth family with an Asn205Ser mutation in the connexin 32 gene.* J Neurol Neurosurg Psychiatry, 1999. **66**(2): p. 202-6.
127. Lee, H.J., et al., *Two novel mutations of GARS in Korean families with distal hereditary motor neuropathy type V.* J Peripher Nerv Syst, 2012. **17**(4): p. 418-21.
128. Panosyan, F.B., et al., *Cross-sectional analysis of a large cohort with X-linked Charcot-Marie-Tooth disease (CMTX1).* Neurology, 2017. **89**(9): p. 927-935.

129. Abrams, C.K., et al., *Pathogenesis of X-linked Charcot-Marie-Tooth disease: differential effects of two mutations in connexin 32.* J Neurosci, 2003. **23**(33): p. 10548-58.
130. Arthur-Farraj, P.J., et al., *Hand weakness in Charcot-Marie-Tooth disease 1X.* Neuromuscul Disord, 2012. **22**(7): p. 622-6.
131. Bissar-Tadmouri, N., et al., *Mutational analysis and genotype/phenotype correlation in Turkish Charcot-Marie-Tooth Type 1 and HNPP patients.* Clin Genet, 2000. **58**(5): p. 396-402.
132. Bone, L.J., et al., *Connexin32 and X-linked Charcot-Marie-Tooth disease.* Neurobiol Dis, 1997. **4**(3-4): p. 221-30.
133. Dubourg, O., et al., *Clinical, electrophysiological and molecular genetic characteristics of 93 patients with X-linked Charcot-Marie-Tooth disease.* Brain, 2001. **124**(Pt 10): p. 1958-67.
134. Senderek, J., et al., *X-linked dominant Charcot-Marie-Tooth disease: nerve biopsies allow morphological evaluation and detection of connexin32 mutations (Arg15Trp, Arg22Gln).* Acta Neuropathol, 1998. **95**(5): p. 443-9.
135. Shy, M.E., et al., *CMT1X phenotypes represent loss of GJB1 gene function.* Neurology, 2007. **68**(11): p. 849-55.
136. Scherer, S.S., et al., *Connexin32-null mice develop demyelinating peripheral neuropathy.* Glia, 1998. **24**(1): p. 8-20.
137. Liang, G.S., et al., *Severe neuropathy with leaky connexin32 hemichannels.* Ann Neurol, 2005. **57**(5): p. 749-54.
138. Wicklein, E.M., et al., *Missense mutation (R15W) of the connexin32 gene in a family with X chromosomal Charcot-Marie-Tooth neuropathy with only female family members affected.* J Neurol Neurosurg Psychiatry, 1997. **63**(3): p. 379-81.
139. Antonellis, A., et al., *Glycyl tRNA synthetase mutations in Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy type V.* Am J Hum Genet, 2003. **72**(5): p. 1293-9.
140. Sivakumar, K., et al., *Phenotypic spectrum of disorders associated with glycyl-tRNA synthetase mutations.* Brain, 2005. **128**(Pt 10): p. 2304-14.
141. Liao, Y.C., et al., *Two Novel De Novo GARS Mutations Cause Early-Onset Axonal Charcot-Marie-Tooth Disease.* PLoS One, 2015. **10**(8): p. e0133423.
142. Dubourg, O., et al., *The G526R glycyl-tRNA synthetase gene mutation in distal hereditary motor neuropathy type V.* Neurology, 2006. **66**(11): p. 1721-6.

143. Sleigh, J.N., et al., *Trk receptor signaling and sensory neuron fate are perturbed in human neuropathy caused by Gars mutations.* Proc Natl Acad Sci U S A, 2017. **114**(16): p. E3324-e3333.
144. Rohkamm, B., et al., *Further evidence for genetic heterogeneity of distal HMN type V, CMT2 with predominant hand involvement and Silver syndrome.* J Neurol Sci, 2007. **263**(1-2): p. 100-6.
145. Ibba, M. and D. Soll, *Aminoacyl-tRNA synthesis.* Annu Rev Biochem, 2000. **69**: p. 617-50.
146. Motley, W.W., et al., *Charcot-Marie-Tooth-linked mutant GARS is toxic to peripheral neurons independent of wild-type GARS levels.* PLoS Genet, 2011. **7**(12): p. e1002399.
147. Griffin, L.B., et al., *Impaired function is a common feature of neuropathy-associated glycyl-tRNA synthetase mutations.* Hum Mutat, 2014. **35**(11): p. 1363-71.
148. WHO, *The global burden of disease: 2004 update. Geneva: World Health Organization.* 2008.
149. Mehra, S., R.D. Eavey, and D.G. Keamy, Jr., *The epidemiology of hearing impairment in the United States: newborns, children, and adolescents.* Otolaryngol Head Neck Surg, 2009. **140**(4): p. 461-72.
150. Olusanya, B.O., R.J. Ruben, and A. Parving, *Reducing the burden of communication disorders in the developing world: an opportunity for the millennium development project.* Jama, 2006. **296**(4): p. 441-4.
151. Grindle, C.R., *Pediatric hearing loss.* Pediatr Rev, 2014. **35**(11): p. 456-63; quiz 464.
152. Moctar, E.C., et al., *Etiology and associated GJB2 mutations in Mauritanian children with non-syndromic hearing loss.* Eur Arch Otorhinolaryngol, 2016. **273**(11): p. 3693-3698.
153. A. AG Mohamed, S.S., S.K. Timbo, F. Konipo-Togola. , *Surdité de l'enfant en Afrique noire : cas de l'ecole des jeunes sourds de Bamako (Mali). Médecine d'Afrique Noire : 1996, 43 (11)*

154. Bayazit, Y.A. and M. Yilmaz, *An overview of hereditary hearing loss.* ORL J Otorhinolaryngol Relat Spec, 2006. **68**(2): p. 57-63.
155. Ravecca, F., et al., *Cx26 gene mutations in idiopathic progressive hearing loss.* J Otolaryngol, 2005. **34**(2): p. 126-34.
156. Sansovic, I., et al., *GJB2 mutations in patients with nonsyndromic hearing loss from Croatia.* Genet Test Mol Biomarkers, 2009. **13**(5): p. 693-9.

157. Hilgert, N., R.J. Smith, and G. Van Camp, *Forty-six genes causing nonsyndromic hearing impairment: which ones should be analyzed in DNA diagnostics?* Mutat Res, 2009. **681**(2-3): p. 189-96.
158. Kassim Diarra, B.G., Youssouf Sidibé, N'faly Konaté, Ibrahim Fofana, Drissa Kaloga Bagayoko, Fatogoma Issa Koné, Siaka Soumaoro, Doumbia Kadidiatou Singare, Samba Karim Timbo, Mohamed Amadou Keita, *Audiometric Profile of Deafness at the University Hospital Center Gabriel Toure of Bamako.* International Journal of Otolaryngology and Head & Neck Surgery, 2020, 9, 38-45, 2020.
159. Traore, H., et al., *[A case report on Waardenburg syndrome with cleft lip].* Mali Med, 2011. **26**(3): p. 53-5.
160. Imperato, P.J. and G.H. Imperato, *Clinical manifestations of Waardenburg syndrome in a male adolescent in Mali, West Africa.* J Community Health, 2015. **40**(1): p. 103-9.
161. Traore, M., et al., *Cauda equina syndrome and profound hearing loss after spinal anesthesia with isobaric bupivacaine.* Anesth Analg, 2006. **102**(6): p. 1863-4.
162. Sako, A.B., *[Particularities of hearing disorders in inhabitants of Mali].* Vestn Otorinolaringol, 1990(1): p. 18-20.
163. HB, S., *Child deafness in Mali.* Otolaryngology Online Journal, 2015. **5 issue 1 2015**.
164. Sacko HB, C.M., Traoré T, Telly n, Coulibaly S, Sanogo H, Fane S, Bagayoko HD, Bouaré M, *Audiometric findings of textile workers in Mali.* Otolaryngology Online Journal, 2016. **Volume 6, Issue 2**.
165. WHO, *World Health Organization: Mali Statistics.* 2013.
166. Meilleur, K.G., et al., *Genetic testing and counseling for hereditary neurological diseases in Mali.* J Community Genet, 2011. **2**(1): p. 33-42.
167. Shojaei, E., Z. Jafari, and M. Gholami, *Effect of Early Intervention on Language Development in Hearing-Impaired Children.* Iran J Otorhinolaryngol, 2016. **28**(84): p. 13-21.
168. Karanja, B.W., et al., *Risk Factors for Hearing Loss in Children following Bacterial Meningitis in a Tertiary Referral Hospital.* Int J Otolaryngol, 2013. **2013**: p. 354725.
169. Schrijver, I., *Hereditary Non-Syndromic Sensorineural Hearing Loss.* Journal of Molecular Diagnostics, Vol. 6, No. 4, November 2004, 2004.
170. Friedman, T.B. and A.J. Griffith, *Human nonsyndromic sensorineural deafness.* Annu Rev Genomics Hum Genet, 2003. **4**: p. 341-402.

171. Vona, B., et al., *Non-syndromic hearing loss gene identification: A brief history and glimpse into the future.* Mol Cell Probes, 2015. **29**(5): p. 260-70.
172. Fraser, F.C., et al., *Genetic aspects of the BOR syndrome--branchial fistulas, ear pits, hearing loss, and renal anomalies.* Am J Med Genet, 1978. **2**(3): p. 241-52.
173. Wonkam-Tingang, E., et al., *Bi-Allelic Novel Variants in CLIC5 Identified in a Cameroonian Multiplex Family with Non-Syndromic Hearing Impairment.* Genes (Basel), 2020. **11**(11).
174. Wonkam-Tingang, E., et al., *A novel variant in DXML2 gene is associated with autosomal dominant non-syndromic hearing impairment (DFNA71) in a Cameroonian family.* Exp Biol Med (Maywood), 2021: p. 1535370221999746.
175. Chen, P., et al., *EYA1 mutations leads to Branchio-Oto syndrome in two Chinese Han deaf families.* Int J Pediatr Otorhinolaryngol, 2019. **123**: p. 141-145.
176. Song, M.H., et al., *Mutational analysis of EYA1, SIX1 and SIX5 genes and strategies for management of hearing loss in patients with BOR/BO syndrome.* PLoS One, 2013. **8**(6): p. e67236.
177. Chang, E.H., et al., *Branchio-oto-renal syndrome: the mutation spectrum in EYA1 and its phenotypic consequences.* Hum Mutat, 2004. **23**(6): p. 582-9.
178. McKenna, A., et al., *The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data.* Genome Res, 2010. **20**(9): p. 1297-303.
179. Chang, C.C., et al., *Second-generation PLINK: rising to the challenge of larger and richer datasets.* Gigascience, 2015. **4**: p. 7.
180. Manichaikul, A., et al., *Robust relationship inference in genome-wide association studies.* Bioinformatics, 2010. **26**(22): p. 2867-73.
181. Choi, Y., et al., *Predicting the functional effect of amino acid substitutions and indels.* PLoS One, 2012. **7**(10): p. e46688.
182. Pollard, K.S., et al., *Detection of nonneutral substitution rates on mammalian phylogenies.* Genome Res, 2010. **20**(1): p. 110-21.
183. Sievers, F., et al., *Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.* Mol Syst Biol, 2011. **7**: p. 539.
184. Waterhouse, A.M., et al., *Jalview Version 2--a multiple sequence alignment editor and analysis workbench.* Bioinformatics, 2009. **25**(9): p. 1189-91.
185. Buchan, D.W.A. and D.T. Jones, *The PSIPRED Protein Analysis Workbench: 20 years on.* Nucleic Acids Res, 2019. **47**(W1): p. W402-w407.

186. Waterhouse, A., et al., *SWISS-MODEL: homology modelling of protein structures and complexes.* Nucleic Acids Res, 2018. **46**(W1): p. W296-w303.
187. Jones, P., et al., *InterProScan 5: genome-scale protein function classification.* Bioinformatics, 2014. **30**(9): p. 1236-40.
188. Heo, L., H. Park, and C. Seok, *GalaxyRefine: Protein structure refinement driven by side-chain repacking.* Nucleic Acids Res, 2013. **41**(Web Server issue): p. W384-8.
189. Schrödinger, *PyMOL | pymol.org.* 2019.
190. Clarke, J.C., et al., *A novel nonsense mutation in the EYA1 gene associated with branchio-oto-renal/branchiootic syndrome in an Afrikaner kindred.* Clin Genet, 2006. **70**(1): p. 63-7.
191. Mosrati, M.A., et al., *A novel dominant mutation in SIX1, affecting a highly conserved residue, result in only auditory defects in humans.* Eur J Med Genet, 2011. **54**(5): p. e484-8.
192. Nasir, S.B., et al., *Branchiootorenal syndrome: A case report.* Niger Postgrad Med J, 2018. **25**(1): p. 60-62.
193. Namba, A., et al., *Genetic features of hearing loss associated with ear anomalies: PDS and EYA1 mutation analysis.* J Hum Genet, 2001. **46**(9): p. 518-21.
194. Chen, A., et al., *Phenotypic manifestations of branchio-oto-renal syndrome.* Am J Med Genet, 1995. **58**(4): p. 365-70.
195. Orten, D.J., et al., *Branchio-oto-renal syndrome (BOR): novel mutations in the EYA1 gene, and a review of the mutational genetics of BOR.* Hum Mutat, 2008. **29**(4): p. 537-44.
196. Angeli, S., X. Lin, and X.Z. Liu, *Genetics of hearing and deafness.* Anat Rec (Hoboken), 2012. **295**(11): p. 1812-29.
197. Beheshtian, M., et al., *Heterogeneity of Hereditary Hearing Loss in Iran: a Comprehensive Review.* Arch Iran Med, 2016. **19**(10): p. 720-728.
198. Plevova, P., et al., *Genetic Aetiology of Nonsyndromic Hearing Loss in Moravia-Silesia.* Medicina (Kaunas), 2018. **54**(2).
199. Sloan-Heggen, C.M., et al., *Comprehensive genetic testing in the clinical evaluation of 1119 patients with hearing loss.* Hum Genet, 2016. **135**(4): p. 441-450.
200. Xiang, Y.B., et al., *Mutation analysis of common deafness-causing genes among 506 patients with nonsyndromic hearing loss from Wenzhou city, China.* Int J Pediatr Otorhinolaryngol, 2019. **122**: p. 185-190.

201. Chakchouk, I., et al., *Disparities in discovery of pathogenic variants for autosomal recessive non-syndromic hearing impairment by ancestry.* Eur J Hum Genet, 2019. **27**(9): p. 1456-1465.
202. Wu, C.C., et al., *Genetic Epidemiology and Clinical Features of Hereditary Hearing Impairment in the Taiwanese Population.* Genes (Basel), 2019. **10**(10).
203. Zhou, Y., et al., *Mutation analysis of common deafness genes among 1,201 patients with non-syndromic hearing loss in Shanxi Province.* Mol Genet Genomic Med, 2019. **7**(3): p. e537.
204. Lasisi, A.O., et al., *Common genes for non-syndromic deafness are uncommon in sub-Saharan Africa: a report from Nigeria.* Int J Pediatr Otorhinolaryngol, 2014. **78**(11): p. 1870-3.
205. Wonkam, A., et al., *No evidence for clinical utility in investigating the connexin genes GJB2, GJB6 and GJA1 in non-syndromic hearing loss in black Africans.* S Afr Med J, 2015. **105**(1): p. 23-6.
206. Brobby, G.W., B. Müller-Myhsok, and R.D. Horstmann, *Connexin 26 R143W mutation associated with recessive nonsyndromic sensorineural deafness in Africa.* N Engl J Med, 1998. **338**(8): p. 548-50.
207. Rudman, J.R., et al., *The genetic basis of deafness in populations of African descent.* J Genet Genomics, 2017. **44**(6): p. 285-294.
208. Rentzsch, P., et al., *CADD: predicting the deleteriousness of variants throughout the human genome.* Nucleic Acids Res, 2019. **47**(D1): p. D886-d894.
209. Kumar, S., et al., *MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms.* Mol Biol Evol, 2018. **35**(6): p. 1547-1549.
210. Zhang, Y., *I-TASSER server for protein 3D structure prediction.* BMC Bioinformatics, 2008. **9**: p. 40.
211. Van Der Spoel, D., et al., *GROMACS: fast, flexible, and free.* J Comput Chem, 2005. **26**(16): p. 1701-18.
212. Pronk, S., et al., *GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit.* Bioinformatics, 2013. **29**(7): p. 845-54.
213. Lindahl E, H.B., Van Der Spoel D *GROMACS 3.0: A package for molecular simulation and trajectory analysis.* J. Mol. Mod

2001. **7**: p. 306-317.
214. Berendsen, H.J.C., van der Spoel, D., & van Drunen, R, *GROMACS - A MESSAGE-PASSING PARALLEL MOLECULAR-DYNAMICS IMPLEMENTATION*

Computer Physics Communications, 1995. **91** https://doi.org/10.1016/0010-4655(95)00042-e(1-3): p. 43-56.
215. Lindorff-Larsen, K., et al., *Improved side-chain torsion potentials for the Amber ff99SB protein force field.* Proteins, 2010. **78**(8): p. 1950-8.
216. Bussi, G., D. Donadio, and M. Parrinello, *Canonical sampling through velocity rescaling.* J Chem Phys, 2007. **126**(1): p. 014101.
217. Berendsen Hjc, P.J., Gunsteren Wfv, Dinola a, Haak Jr *Molecular dynamics with coupling to an external bath molecular dynamics with coupling to an external bath.* J. Chem. Physics

1984. **81**: p. 3684.
218. Darden T, Y.D., Pedersen L. , Particle mesh Ewald: an N -log (N) method for Ewald sums in large systems, J. Chem. Physics 1993. **98**: p. 10089–10092.
219. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., and Lee, H *A smooth particle mesh Ewald method.* J. Chem. Physics 1995 **103**: p. 8577-8593.
220. Chen, T., et al., *Cochlear supporting cells require GAS2 for cytoskeletal architecture and hearing.* Dev Cell, 2021. **56**(10): p. 1526-1540.e7.
221. Vona, B., et al., *A biallelic variant in CLRN2 causes non-syndromic hearing loss in humans.* Hum Genet, 2021. **140**(6): p. 915-931.
222. Wonkam, A., et al., *Exome sequencing of families from Ghana reveals known and candidate hearing impairment genes.* Commun Biol, 2022. **5**(1): p. 369.
223. Cortese, A., et al., *Biallelic mutations in SORD cause a common and potentially treatable hereditary neuropathy with implications for diabetes.* Nat Genet, 2020. **52**(5): p. 473-481.
224. Friedman, T.B., I.A. Belyantseva, and G.I. Frolenkov, *Myosins and Hearing.* Adv Exp Med Biol, 2020. **1239**: p. 317-330.
225. Rehman, A.U., et al., *Mutational Spectrum of MYO15A and the Molecular Mechanisms of DFNB3 Human Deafness.* Hum Mutat, 2016. **37**(10): p. 991-1003.
226. Budde, B.S., et al., *Comprehensive molecular analysis of 61 Egyptian families with hereditary nonsyndromic hearing loss.* Clin Genet, 2020. **98**(1): p. 32-42.
227. Naz, S., et al., *Genetic causes of moderate to severe hearing loss point to modifiers.* Clin Genet, 2017. **91**(4): p. 589-598.
228. Wang, L., et al., *Targeted Next-Generation Sequencing Identified Compound Heterozygous Mutations in MYO15A as the Probable Cause of*

Nonsyndromic Deafness in a Chinese Han Family. Neural Plast, 2020. **2020**: p. 6350479.
229. Nishio, S.Y. and S. Usami, *Deafness gene variations in a 1120 nonsyndromic hearing loss cohort: molecular epidemiology and deafness mutation spectrum of patients in Japan.* Ann Otol Rhinol Laryngol, 2015. **124 Suppl 1**: p. 49s-60s.
230. Weil, D., et al., *Defective myosin VIIA gene responsible for Usher syndrome type 1B.* Nature, 1995. **374**(6517): p. 60-1.
231. Delmaghani, S., et al., *Mutations in the gene encoding pejvakin, a newly identified protein of the afferent auditory pathway, cause DFNB59 auditory neuropathy.* Nat Genet, 2006. **38**(7): p. 770-8.
232. Ebermann, I., et al., *Truncating mutation of the DFNB59 gene causes cochlear hearing impairment and central vestibular dysfunction.* Hum Mutat, 2007. **28**(6): p. 571-7.
233. Collin, R.W., et al., *Involvement of DFNB59 mutations in autosomal recessive nonsyndromic hearing impairment.* Hum Mutat, 2007. **28**(7): p. 718-23.
234. Stenson, P.D., et al., *The Human Gene Mutation Database (HGMD(®)): optimizing its use in a clinical diagnostic or research setting.* Hum Genet, 2020. **139**(10): p. 1197-1207.
235. Astuto, L.M., et al., *CDH23 mutation and phenotype heterogeneity: a profile of 107 diverse families with Usher syndrome and nonsyndromic deafness.* Am J Hum Genet, 2002. **71**(2): p. 262-75.
236. Di Palma, F., et al., *Mutations in Cdh23, encoding a new type of cadherin, cause stereocilia disorganization in waltzer, the mouse model for Usher syndrome type 1D.* Nat Genet, 2001. **27**(1): p. 103-7.
237. Shapiro, L. and W.I. Weis, *Structure and biochemistry of cadherins and catenins.* Cold Spring Harb Perspect Biol, 2009. **1**(3): p. a003053.
238. Ahmed, Z.M., et al., *The molecular genetics of Usher syndrome.* Clin Genet, 2003. **63**(6): p. 431-44.
239. Pan, B., et al., *TMC1 and TMC2 are components of the mechanotransduction channel in hair cells of the mammalian inner ear.* Neuron, 2013. **79**(3): p. 504-15.
240. Gao, X., et al., *Targeted gene capture and massively parallel sequencing identify TMC1 as the causative gene in a six-generation Chinese family with autosomal dominant hearing loss.* Am J Med Genet A, 2015. **167a**(10): p. 2357-65.

241. Wang, H., et al., *Identification of four TMC1 variations in different Chinese families with hereditary hearing loss.* Mol Genet Genomic Med, 2018. **6**(4): p. 504-13.
242. Loftus, B.J., et al., *Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.* Genomics, 1999. **60**(3): p. 295-308.
243. Diez-Roux, G., et al., *A high-resolution anatomical atlas of the transcriptome in the mouse embryo.* PLoS Biol, 2011. **9**(1): p. e1000582.
244. Ko, M.S., et al., *Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome.* Hum Mol Genet, 1998. **7**(12): p. 1967-78.
245. Requena, T., A. Gallego-Martinez, and J.A. Lopez-Escamez, *Bioinformatic Integration of Molecular Networks and Major Pathways Involved in Mice Cochlear and Vestibular Supporting Cells.* Front Mol Neurosci, 2018. **11**: p. 108.
246. HH, T., *The Peroneal Type of Progressive Muscular Atrophy*. 1886., London: H.K. Lewis

247. Charcot JM, M.P., *Sur une form particulie're d¢atrophie musculaire progressive, souvant familiale, debutant par les pieds et les jambes, et atteignant plus tard les mains.* Rev Med (Paris) 1886. **6**: p. 97-138. .
248. Valentijn, L.J., et al., *The peripheral myelin gene PMP-22/GAS-3 is duplicated in Charcot-Marie-Tooth disease type 1A.* Nat Genet, 1992. **1**(3): p. 166-70.
249. Ju, D., et al., *Importance of Including Non-European Populations in Large Human Genetic Studies to Enhance Precision Medicine.* Annu Rev Biomed Data Sci, 2022.
250. Wonkam, A., et al., *Five Priorities of African Genomics Research: The Next Frontier.* Annu Rev Genomics Hum Genet, 2022.
251. Echaniz-Laguna, A., et al., *[Treating hereditary neuropathies : a dream come true?].* Rev Med Suisse, 2022. **18**(779): p. 813-816.
252. Takashima, H., C.F. Boerkoel, and J.R. Lupski, *Screening for mutations in a genetically heterogeneous disorder: DHPLC versus DNA sequence for mutation detection in multiple genes causing Charcot-Marie-Tooth neuropathy.* Genet Med, 2001. **3**(5): p. 335-42.

253. Hsu, Y.H., et al., *Mutation spectrum of Charcot-Marie-Tooth disease among the Han Chinese in Taiwan.* Ann Clin Transl Neurol, 2019. **6**(6): p. 1090-1101.
254. Ionasescu, V., et al., *Autosomal dominant Charcot-Marie-Tooth axonal neuropathy mapped on chromosome 7p (CMT2D).* Hum Mol Genet, 1996. **5**(9): p. 1373-5.
255. Sambuughin, N., et al., *Autosomal dominant distal spinal muscular atrophy type V (dSMA-V) and Charcot-Marie-Tooth disease type 2D (CMT2D) segregate within a single large kindred and map to a refined region on chromosome 7p15.* J Neurol Sci, 1998. **161**(1): p. 23-8.
256. A. AG Mohamed, S.S., S.K. Timbo, F. Konipo-Togola. , *Surdité de l'enfant en Afrique noire : cas de l'ecole des jeunes sourds de Bamako (Mali). Médecine d'Afrique Noire : 1996, 43 (11).*
257. Ross, J.P., P.A. Dion, and G.A. Rouleau, *Exome sequencing in genetic disease: recent advances and considerations.* F1000Res, 2020. **9**.
258. Green, E.D., J.D. Watson, and F.S. Collins, *Human Genome Project: Twenty-five years of big biology.* Nature, 2015. **526**(7571): p. 29-31.
259. Fraser, F.C., et al., *Autosomal dominant duplication of the renal collecting system, hearing loss, and external ear anomalies: a new syndrome?* Am J Med Genet, 1983. **14**(3): p. 473-8.

Appendices

Appendix 1: Supplementary materials

GJB1 variants in Charcot-Marie-Tooth disease X-linked type 1 in Mali

Abdoulaye Yalcouyé, Seybou H Diallo, Lassana Cissé, Mamadou Karembé, Salimata Diallo, Thomas Coulibaly, Salimata Diarra, Dramane Coulibaly, Mohamed Keita, Cheick O Guinto, Kenneth Fischbeck, Ambroise Wonkam, Guida Landouré

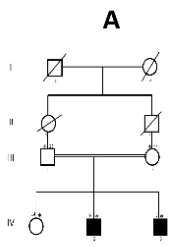
Table S1: Description of the variants with pathogenicity prediction

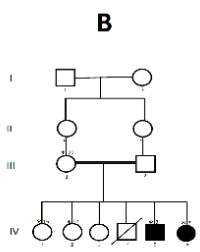
Prediction tools	**Variants**		
	c.271G>A, p.Val91Met	c.43C>T, p.Arg15Trp	c.704T>G, p.Phe235Cys
BayesDel addAF	Damaging	Damaging	Damaging
BayesDel noAF	Damaging	Damaging	Damaging
DANN	0.9989	0.9991	0.6584
DEOGEN2	Damaging	Damaging	Damaging
FATHMM	Damaging	Damaging	Damaging
FATHMM-MKL	Damaging	Damaging	Damaging
LIST-S2	0.9854	0.9725	0.5999
LRT	Deleterious	Deleterious	Neutral
MVP	0.9972	0.9988	0.8923
MetaRNN	Damaging	Damaging	Tolerated
MetaSVM	Damaging	Damaging	Damaging
MutPred	Pathogenic	Pathogenic	Neutral
MutationAssessor	3.555	2.36	0.345
MutationTaster	Disease causing	Disease causing	Disease causing
PROVEAN	Damaging	Damaging	Tolerated
PrimateAI	Damaging	Damaging	Tolerated
REVEL	Pathogenic	Pathogenic	0.381
SIFT	0.001	0	0.175
SIFT4G	0.004	0.001	0.165
CADD	25.7	28.6	17.12
PolyPhen2	Probably damaging	Probably damaging	Possibly damaging
ACMG criterias	PS4, PM1, PM2, PM5, PP1, PP2, PP3, PP5	PS3, PM1, PM2, PP1, PP3, PP5	PP2, BS1, BS2, BP6, BP4

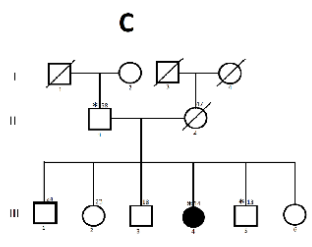
Figure S1: Supplementary material: Pedigrees of selected families segregating HI in Mali

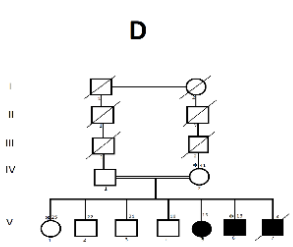
Etiologies of Childhood Hearing Impairment in Schools for the Deaf in Mali

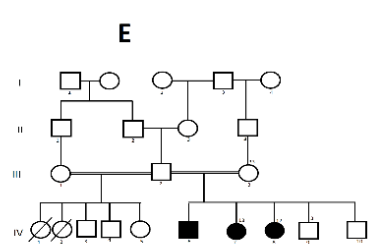
Abdoulaye Yalcouyé, Oumou Traoré, Abdoulaye Taméga, Alassane B Maïga, Fousseyni Kané, Oluwafemi G Oluwole, Cheick Oumar Guinto, Mohamed Kéita, Samba Karim Timbo, Carmen De Kock, Guida Landouré, Ambroise Wonkam

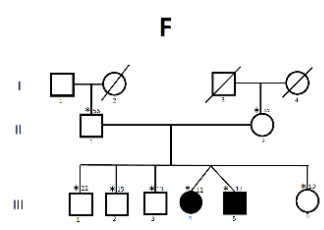

A

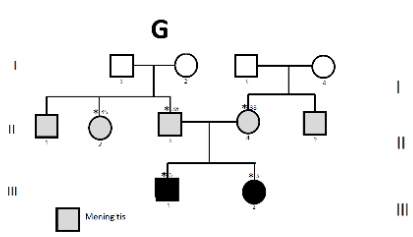

B

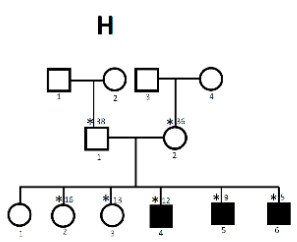

C

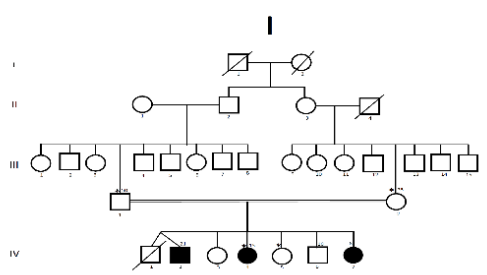

D

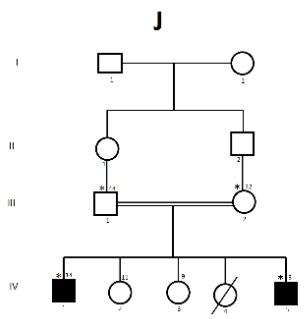

E

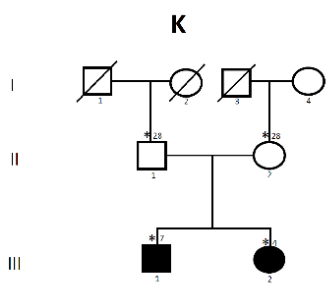

F

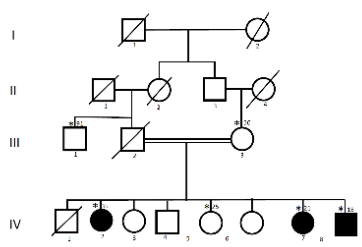

G
Meningitis

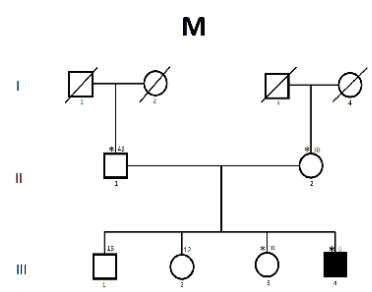

H

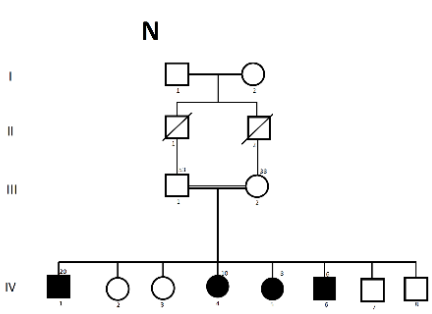

I

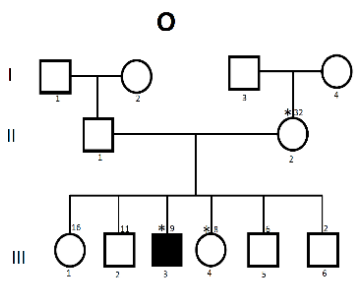

J


K


L


M


N


O

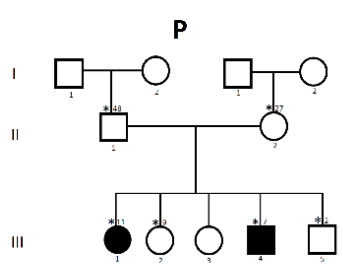

P

Table S2: Primer information of *EYA1*

A monoallelic variant in EYA1 is associated with Branchio- Otic syndrome in a Malian family

Abdoulaye Yalcouyé, Traoré Oumou, Salimata Diarra, Isabelle Schrauwen, Kevin Esoh, Magda Kamila Kadlubowska, Thashi Bharadwaj, Samuel Mawuli Adadey, Mohamed Kéita, Cheick O. Guinto, Suzanne M. Leal, Guida Landouré, Ambroise Wonkam

Gene	Primer	Forward	Reverse	Product size(bp)	Annealing Temperature
EYA1	Exon1	5'-AGAGGACGCGTGTGTGTTGA-3'	5'-AAAGCGCTCCTGTTAGCTTGG-3'	827	62
	Exon2	5'-CACGATCATTTTGTCATTCG-3'	5'-TTTCCAACAGAGGCTGTTACT-3'	538	57
	Exon3	5'-TGTGCAAGTGTGTTTCAAAGG-3'	5'-GCATTTGGTGAAACGAAACC-3'	446	58
	Exon4	5'-TTGTTTCCTAAGGGGAAAGAAC-3'	5'-TGTGACTCTGAGCAAAGCTATTTT-3'	449	53
	Exon5	5'-TTGCGAAGATTGTGATGACG-3'	5'-GGAACATGTGGGCACAGAC-3'	403	58
	Exon6	5'-GCTGGAATTTGTGATGTGGTT-3'	5'-ATTGGCCAAAGATTGGGTCT-3'	329	58
	Exon7	5-'TGAGATAAGATTGGGGAAGCA-3'	5'-CAGGTGTCCTGCCTCTAAGC-3'	443	58
	Exon8	5'-CCAAATACCAATTCTGCCTTTT-3'	5'-TGAAAACCAAACAACTCACCA-3'	345	58
	Exon9	5'-GTACTTGTATGCCCCCGTGT-3'	5'-CCAAAATTGTGCAACCACTG-3'	417	58
	Exon10	5'-ACCAGCGCAAGTAAAAGACG-3'	5'-GCATCTGATACCTTAACCACTGC-3'	390	58
	Exon11	5'-TCATTCATCTTCCGTTTCAAGA-3'	5'-CACTGGGGTCTGAATAAGCA-3'	434	58
	Exon12	5'-CATCAACATTTGGGGCTCTT-3'	5'-AGGCAAAACACATTGCCATA-3'	444	58
	Exon13	5'-GACTGCCACCTACTGATTGACA-3'	5'-GGAAAGCCATCTGTTCCAAA-3'	341	58
	Exon14	5'-AAGGTGAGCACCCTTGAATG-3'	5'-GGCCAGTGAGATGAAACTG-3'	406	58
	Exon15-16	5'-CAAAGCCGAAGAAATATGTTG-3'	5'-TCCTGAAGGAAAAGAGCTGA-3'	450	58

Table S3: Description of the variant with pathogenicity prediction tools

	c.1286A>G, (p.Asp429Glu)
Predicted effect	Missense
GERP	5.44
PhyloP	7.969
PhastCons	1
SiPhy	15.666
SIFT	0.001
Polyphen2 HDIV	0.899
Polyphen2 HVAR	0.826
MutationAssessor High	2.975
LRT	0
M-CAP13	0.259301
REVEL	0.939
MutPred	0.924
PROVEAN	Damaging
MetaSVM	Damaging
MetaLR	Damaging
MutationTaster	1
Eigen	0.836
Eigen-PC	0.803
FATHMM-MKL	Damaging
CADD	26.8
DANN	0.999
ACMG classification	Likely pathogenic

RefSeq transcript used NM_172059

Table S4: Online resources

ANNOVAR	https://annovar.openbioinformatics.org/
Bureau international d'audiophonologie (BIAP)	https://www.biap.org/en/recommandations/recommendations/tc-02-classification
ClinVar	https://www.ncbi.nlm.nih.gov/clinvar/
dbNSFP (including dbscSNV)	https://sites.google.com/site/jpopgen/dbNSFP
dbSNP	https://www.ncbi.nlm.nih.gov/snp/
DRAGEN germline pipeline	https://emea.illumina.com/products/by-type/informatics-products/basespace-sequence-hub/apps/edico-genome-inc-dragen-germline-pipeline.html
Ensembl	https://www.ensembl.org/index.html
Gene ontology (GO)	http://geneontology.org/
Genome aggregation database (gnomAD)	https://gnomad.broadinstitute.org/
Genome analysis toolkit (GATK)	https://gatk.broadinstitute.org/hc/en-us

Hereditary hearing loss homepage (HHL)	https://hereditaryhearingloss.org/
Human phenotype ontology (HPO)	https://hpo.jax.org/app/
Human splice finder (HSF)	https://hsf.genomnis.com/home
InterProScan	http://www.ebi.ac.uk/InterProScan/
MODELLER	http://www.salilab.org/modeller
NCBI-BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi
Online Mendelian inheritance in man (OMIM)	https://omim.org/
PDB	https://www.wwpdb.org/
PSIPRED	http://bioinf.cs.ucl.ac.uk/psipred/
PYMOL	http://www.pymol.org/
RefSeq	https://www.ncbi.nlm.nih.gov/refseq/
Swiss-Model	https://swissmodel.expasy.org/
Uniprot	https://www.uniprot.org/uniprot/Q9NZA1
UK10K	https://www.uk10k.org/

World Health Organisation	https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss

Figure S3: Secondary structure analysis of the EYA1 wildtype and mutant protein

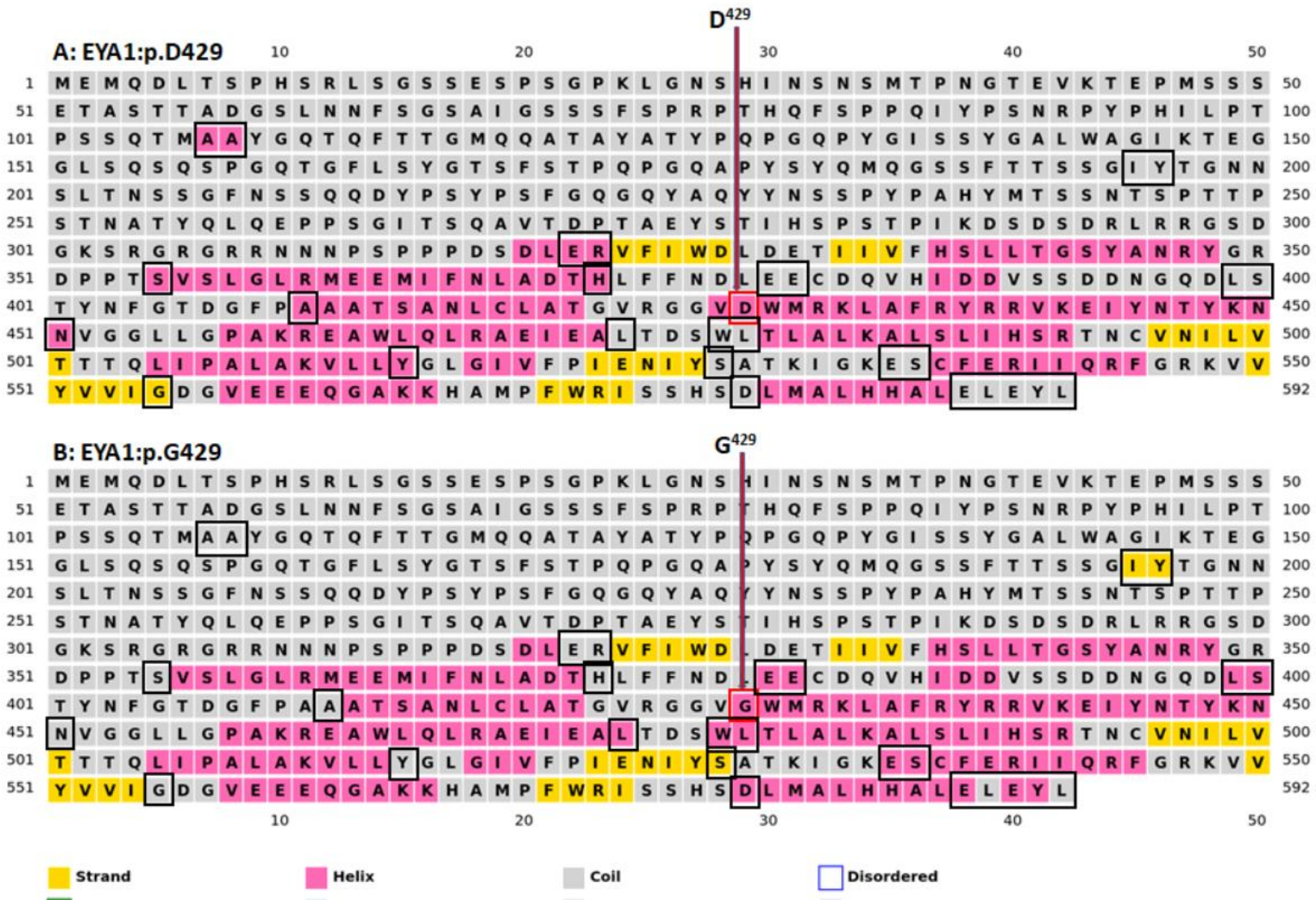


Table S5: Predictions scores of the variants

Prediction tools	*MYO15A* (p.N2111Y)	*MYO15A* (p.D2720N)	*MYO7A* (p.C1326X)	*PJVK* (p.V119G)	*CDH23* (p.L216V)	*OTOGL* (c.209-9C>G)	*OTOGL* (p.D1895N)	*TMC1* c.2003+2T>C)	*UBFD1* (p.E20K)
SIFT_score	0.001	0.016	.	0.001	0.005	.	0.395	.	0.153
SIFT_pred	D	D	.	D	D	.	T	.	T
Polyphen2_HDIV_score	1	0.999	.	0.987	0.993	.	.	.	0.134
Polyphen2_HDIV_pred	D	D	.	D	D	.	.	.	B
Polyphen2_HVAR_score	0.967	0.763	.	0.864	0.971	.	.	.	0.006
Polyphen2_HVAR_pred	D	P	.	P	D	.	.	.	B
LRT_score	.	.	0	0.01	0.003	.	.	.	0.119
LRT_pred	.	.	D	U	N	.	.	.	N
MutationTaster_score	1	1	1	1	0.914	.	1	1	0.977
MutationTaster_pred	A	D	A	D	D	.	N	D	D
MutationAssessor_score	3.215	2.525	.	2.3	2.74	.	0.945	.	0.205
MutationAssessor_pred	M	M	.	M	M	.	L	.	N
FATHMM_score	-2.93	-3.52	.	1.85	4.14	.	2.52	.	.
FATHMM_pred	D	D	.	T	T	.	T	.	.
PROVEAN_score	-6.32	-4.42	.	-4.44	-1.33	.	-0.92	.	-0.54
PROVEAN_pred	D	D	.	D	N	.	N	.	N
VEST3_score	0.597	0.249	.	0.942	0.64	.	0.034	.	0.272
MetaSVM_score	0.934	0.967	.	-1.063	-1.152	.	-1.018	.	-0.983
MetaSVM_pred	D	D	.	T	T	.	T	.	T
MetaLR_score	0.867	0.891	.	0.116	0.038	.	0.03	.	0.083
MetaLR_pred	D	D	.	T	T	.	T	.	T
M.CAP_score	0.552	0.53	.	0.039	0.037	.	0.01	.	0.253
M.CAP_pred	D	D	.	D	D	.	T	.	D
REVEL_score	0.87	0.78	.	0.466	0.18	.	0.075	.	0.184
MutPred_score	0.876	0.879	.	0.774	0.622	.	.	.	.
CADD_raw	4.023	5.197	10.247	6.02	3.584	.	0.059	4.968	2.471
CADD_phred	23.6	25.5	36	27.9	23.2	.	3.173	25.1	19.28
DANN_score	0.99	0.997	0.986	0.997	0.99	.	0.632	0.996	0.996
fathmm.MKL_coding_score	0.971	0.983	0.403	0.997	0.937	.	0.094	0.99	0.191
fathmm.MKL_coding_pred	D	D	N	D	D	.	N	D	N

Notes: A: Disease causing, Medium, L: low, D: Damaging, T: Tolerated, N: Neutral, U: Unknown

Appendix 2: Published manuscripts excluded with significant contribution of the candidate

1. Hotchkiss, J., Manyisa, N., Adadey, S. M., Oluwole, O. G., Wonkam, E., Mnika, K., Yalcouye, A., Nembaware, V., Haendel, M., Vasilevsky, N., Mulder, N. J., Jupp, S., Wonkam, A., & Mazandu, G. K. (2019). The Hearing Impairment Ontology: A Tool for Unifying Hearing Impairment Knowledge to Enhance Collaborative Research. Genes, 10(12). https://doi.org/10.3390/genes10120960. (Status: published)

Contribution to authorship

G.K.M., V.N., and A.W. are leading the H.I.O. implementation.

J.H., N.M., E.W., V.N., N.V., G.K.M., and A.W. curated the current

H.I.O. content. N.M., J.H., E.W., S.M.A., O.G.O., K.M., A.Y (**Abdoulaye Yalcouyé**)., and A.W. provided the H.I.O. content.

G.K.M., N.J.M., M.H., N.V., S.J. contributed to the ontology structure development and implementation.

G.K.M., J.H., N.M., and V.N. wrote the manuscript.

S.M.A., O.G.O., E.W., K.M., A.Y **(Abdoulaye Yalcouyé)**., M.H., N.V., N.J.M., S.J., and A.W. revised the manuscript.

Article

The Hearing Impairment Ontology: A Tool for Unifying Hearing Impairment Knowledge to Enhance Collaborative Research

Jade Hotchkiss [1,†], Noluthando Manyisa [1,†], Samuel Mawuli Adadey [1,2], Oluwafemi Gabriel Oluwole [1], Edmond Wonkam [1], Khuthala Mnika [1], Abdoulaye Yalcouye [1,3], Victoria Nembaware [1], Melissa Haendel [4], Nicole Vasilevsky [4], Nicola J. Mulder [5], Simon Jupp [6], Ambroise Wonkam [1,*] and Gaston K. Mazandu [1,5,7,*]

1 Division of Human Genetics, Department of Pathology, University of Cape Town, Health Sciences Campus, Anzio Rd, Observatory 7925, South Africa; giant.plankton@gmail.com (J.H.); mnynol006@myuct.ac.za (N.M.); smadadey@st.ug.edu.gh (S.M.A.); oluwafemi.oluwole@uct.ac.za (O.G.O.); wonkamedmond@yahoo.fr (E.W.); mnikakhuthala@gmail.com (K.M.); YLCABD001@myuct.ac.za (A.Y.); vnembaware@gmail.com (V.N.)
2 Department of Neurology, Point G Teaching Hospital, University of Sciences, Techniques and Technology, Bamako, Mali
3 West African Centre for Cell Biology of Infectious Pathogens, College of Basic and Applied Sciences, University of Ghana, Accra, Ghana
4 School of Medicine, Oregon Health & Science University, Portland, OR 97217, USA; haendel@ohsu.edu (M.H.); vasilevs@ohsu.edu (N.V.)
5 Computational Biology Division, Department of Integrative Biomedical Sciences, University of Cape Town, Health Sciences Campus. Anzio Rd, Observatory 7925, South Africa; nicola.mulder@uct.ac.za
6 European Bioinformatics Institute (EMBL-EBI), Cambridge CB10 1SD, UK; simon.jupp@gmail.com
7 African Institute for Mathematical Sciences, 5-7 Melrose Road, Muizenberg 7945, Cape Town, South Africa
* Correspondence: ambroise.wonkam@uct.ac.za (A.W.); gaston.mazandu@uct.ac.za (G.K.M.)
† First two authors are joint first author.

Received: 18 September 2019; Accepted: 11 November 2019; Published: 21 November 2019

Abstract: Hearing impairment (HI) is a common sensory disorder that is defined as the partial or complete inability to detect sound in one or both ears. This diverse pathology is associated with a myriad of phenotypic expressions and can be non-syndromic or syndromic. HI can be caused by various genetic, environmental, and/or unknown factors. Some ontologies capture some HI forms, phenotypes, and syndromes, but there is no comprehensive knowledge portal which includes aspects specific to the HI disease state. This hampers inter-study comparability, integration, and interoperability within and across disciplines. This work describes the HI Ontology (HIO) that was developed based on the Sickle Cell Disease Ontology (SCDO) model. This is a collaboratively developed resource built around the 'Hearing Impairment' concept by a group of experts in different aspects of HI and ontologies. HIO is the first comprehensive, standardized, hierarchical, and logical representation of existing HI knowledge. HIO allows researchers and clinicians alike to readily access standardized HI-related knowledge in a single location and promotes collaborations and HI information sharing, including epidemiological, socio-environmental, biomedical, genetic, and phenotypic information. Furthermore, this ontology illustrates the adaptability of the SCDO framework for use in developing a disease-specific ontology.

Keywords: hearing impairment; hearing loss; ontology; data harmonization; meta-analysis

1. Introduction

Hearing impairment (HI), the partial or total inability to hear, is a communication barrier and language development impediment. It can thus have a huge effect on one's quality of life [1]. HI has the highest rate for age-standardized disability of life in the world [2,3]. According to the most recent World Health Organization (WHO) estimates [4], over six percent of the world's population, representing approximately 460 million individuals, are currently living with a disabling HI, of which 93% are adults and mostly males (242 million males vs 190 million females). The financial burden associated with HI, which includes costs for healthcare, education, social support, and loss of productivity [5], is estimated to be 750 billion US dollars annually [6]. Even though 60% of HI cases can be prevented [5], the number of cases is expected to significantly increase to over 900 million in 2050 [4] with huge negative economic implications, unless action is taken. As a matter of urgency, there is a need to strengthen collaborative HI research efforts aimed at curbing the projected increased burden of HI globally. The Global Hearing Loss project (https://thespindle.org/project/global-hearing-loss-database/) has highlighted that HI research data is commonly unstructured, stored in natural language format, and hardly shared. The general lack of standardization of research data on rare or neglected diseases across studies [7] hampers presentation, sharing, integration, and interoperability of important information, such as prevalence, socio-environmental, biomedical, and phenotypic information. The need for harmonized HI datasets motivated the World-Wide Hearing group to develop a standard platform, the Global Hearing Loss Database (GHLD). The GHLD is based on WHO protocols with a web portal and a smartphone application to ease HI data collection and sharing processes. However, given the complexity of HI etiologies and phenotypes, analyses of these datasets and inter-study comparability would require a standard knowledge representation of the HI knowledge domain [8]. A standard knowledge representation of the HI concepts or terms would include concise descriptions to ensure a common understanding of the domain and to enable automated reasoning and inferencing. Moreso, with the constant evolution of biomedical knowledge [7], a human- and machine-readable upgradeable system is needed for standardized and well-defined HI knowledge representation to enhance collaborative research in the field.

Recent advances in artificial intelligence have fostered the use of ontology models to represent knowledge and information-based systems [9] in a human- and machine-readable format to help process, reuse, and re-apply knowledge [10,11]. An ontology is useful in establishing a common and controlled vocabulary system, describing key concepts, properties, and hierarchical relationships between concepts [12], with precise definitions for clear and unambiguous communication. In the biomedical research context, several human disease-related ontologies have been introduced, including the Human Disease Ontology (DO), which consistently defines various concepts encountered in disease domains [13], the Mondo Disease Ontology, which provides a merged and comprehensive cross-species disease ontology (https://monarch-initiative.github.io/mondo/), and the Human Phenotype Ontology (HPO), providing controlled vocabularies of abnormal phenotypes encountered in human disease [14,15]. However, as previously argued in support of developing the Sickle Cell Disease Ontology (SCDO) [5], none of the existing ontologies comprehensively captures related concepts specific to HI due to the complex nature of HI etiologies and phenotypes.

We present the Hearing Impairment Ontology (HIO), built upon the SCDO framework. The HIO was compiled by a working group, which includes HI and ontology experts, who defined, in detail, essential aspects of the HI knowledge domain (e.g., phenotypes, genetics, therapeutics, diagnostics, etc.) and how these aspects are related. Similar to how the SCDO was built around the central concept, 'Hemoglobinopathy' [7], the HIO is built around the 'Hearing Impairment' concept. However, because HI can be associated with a myriad of phenotypes and/or syndromes, caused by various genetic, environmental, and/or unknown factors, the SCDO framework was adapted to account for the additional complexities of HI. To date, this developing HIO represents the most comprehensive standardized HI domain knowledge portal, which will allow for the application of ontology-driven mining approaches for the identification of pertinent research questions.

2. Materials and Methods

For this first version of the HIO, the working group consisted of experts from an existing Hearing Impairment Genetics Studies in Africa (HI-GENES Africa) project at the University of Cape Town, Division of Human Genetics in the Department of Pathology, and expert ontology developers to provide technical guidance. The experts from the HI-GENES Africa project included PhD students and Postdocs, with a varied range of expertise which included clinicians, biomedical scientists, geneticists, and bioinformatics experts. SCDO curators and developers led the design, following the published ontology development reporting guidelines [16] and best practice.

2.1. SCDO Model-Based HIO Development

After attempting to "fit" different forms of HI and their associated causes into the same upper classes used in the SCDO (except replacing 'Hemoglobinopathy' with 'Hearing Impairment'), it was apparent that the HIO model needed to be carefully adjusted to account for the marked differences in the causes and pathophysiology between these diseases. A schema for the HIO was drawn up to formalize how the HIO would be modelled (what main classes would be needed and what relationships would be described between these classes). It is worth noting that these two diseases have complex phenotypic expressions, influenced by several genetic and environmental factors. Since SCDO was built around the central concept 'Hemoglobinopathy' to include more factors influencing its phenotypes [7], likewise, HIO is built around the central concept 'Hearing Impairment' to ensure that all aspects influencing the disease outcome and phenotypic manifestations are captured. This is achieved by relating different HI concepts specified in the ontology to various factors, including genetic and environmental factors, that contribute to the disease outcome. The overview of different steps in the modeling, from populating the ontology, checking different concepts and relations to the release of the HIO by curators, domain and ontology experts via internal and external reviews, is described in Figure 1.

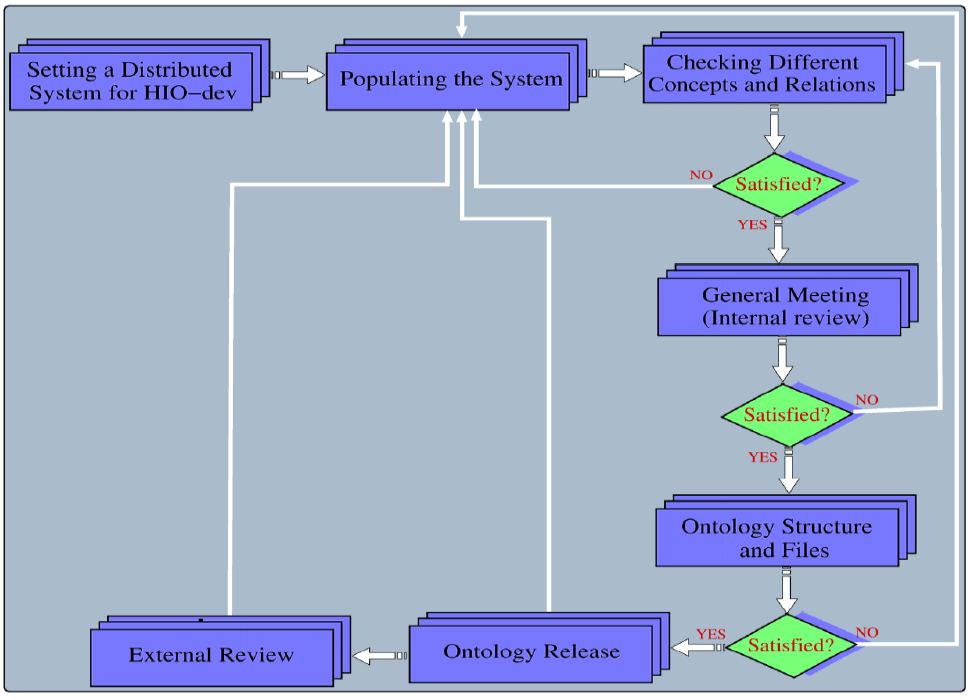


Figure 1. Flow chart of the dynamic and iterative ontology development process. It starts by setting up an online collaborative ontology development tool, WebProtege, which provides a highly distributed

ontology content management system, enabling domain experts, ontology curators, and developers to share and update information, and easily visualise the ontology classes and structure. A general discussion meeting (or internal review) is called to share a common understanding of existing Hearing Impairment (HI) knowledge currently included in the ontology and resolve any disagreement about a given concept.

2.2. Building Different HIO Objects

Annotation properties (both required and optional) were re-used from the SCDO. The additional annotation property, 'deprecated synonym', was included by the working group in order to indicate when a term has a synonym that is no longer acceptable. Terms to be included in the ontology were added and annotated in a shared online spreadsheet by the working group. The relationships between classes were also captured in shared online spreadsheets (each sheet dedicated to the relationships made by a certain object property). To keep track of terms reused from existing ontologies, the 'existence in other ontologies' annotation property was used to assign an 'existence status' to each term. The frequency of existence statuses was subsequently used also to evaluate ontology terms unique to the HIO and its contribution to updating HI terms in other ontologies.

2.3. Distributed Model-Based HIO Design

Coordinating an ontology development with groups of contributors from heterogeneous specialized backgrounds to derive a unified domain conceptualization is challenging. To ease the process, the online collaborative ontology development tool, WebProtege [17], was used to draft the skeleton structure (labels of terms only) of the HIO in Ontology Web Language (OWL) format. This tool provides a highly distributed ontology content management system, enabling domain experts, ontology curators, and developers to share and update information, and easily visualise the ontology classes and structure.

2.4. HIO File Refinement and Evolution

For quality control assurance of different concepts in the ontology and considering the dynamic evolution of the ontology structure, an iterative and collaborative process was used for refining different definitions and properties, as well as the topological structure. These concepts were validated by experts before being included into the hio-edit.owl file in WebProtege. Each term had to be checked by at least two members of the working group, including at least one Hl expert. Once terms were validated during the general online discussion meeting (or internal review), curators added their annotations from the spreadsheet into the WebProtege project (recording in the spreadsheet which terms had been added). Thereafter, ROBOT, an open source tool for automating ontology development workflows and tasks [18], was used to compile the complete ontology release files, which are Web Ontology Language (OWL) and Open Biomedical Ontology (OBO) formats.

3. Results

3.1. HIO General Description

Other than the 'HIO' and 'deprecated terms' classes, the HIO currently consists of 10 upper classes (see Figure 2). It contains 495 terms that are topologically connected by 543 links (is_a relationships) and expected to be associated with each other by 45 object properties (excluding the owl:topObjectProperty). Figure 2 shows how the SCDO upper classes were adapted for the HIO. Figure 3 shows the number of concepts in the HIO's upper level classes (Graph A) and the number of relationships currently asserted between concepts via object properties (Graph B) in this first release version of the ontology.

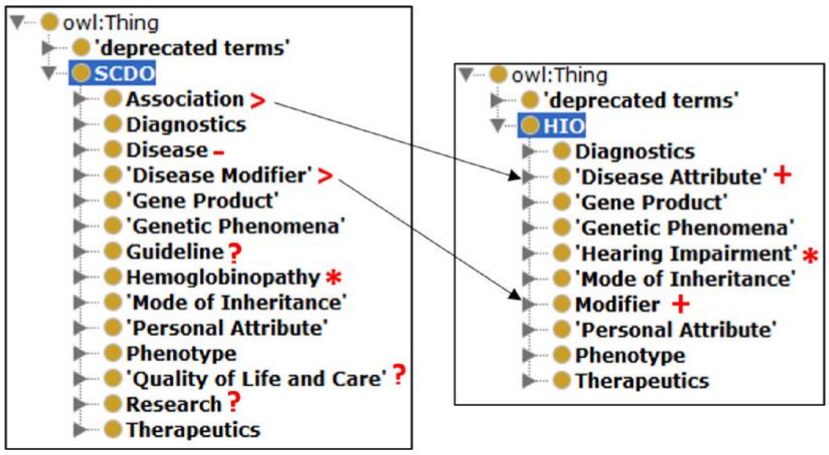


Figure 2. The upper classes of the Sickle Cell Disease Ontology (SCDO) and HI Ontology (HIO). (*) indicates the ontologies' central classes. (+) indicates classes in HIO but not in the SCDO. (-) indicates a class in SCDO but not in HIO. (>) indicates classes in SCDO that were incorporated in other HIO classes. (?) indicates SCDO classes that still need to be reviewed and adapted as necessary for inclusion into the HIO.

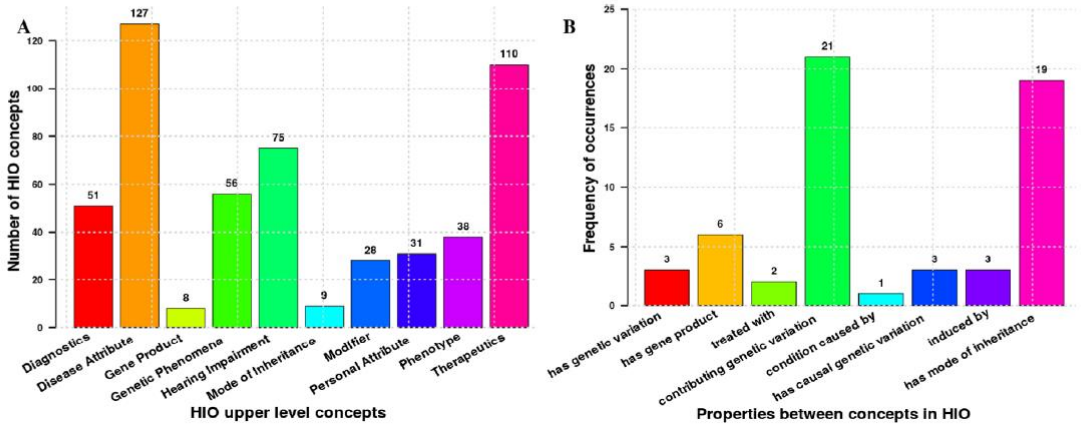


Figure 3. Summary statistics of current concepts and properties in the current HIO. Numbers at the top of bars represent the number of different HIO sub-classes topologically linked to upper level classes (**A**) and the occurrence frequency of a given property or association in the ontology (**B**). Note that 'contributing genetic variation' is used as the short hand label for the 'gene carrying contributing genetic variation' property and 'has causal genetic variation' for the 'condition has causal or contributing genetic variation' property.

The following upper classes from the SCDO are included in this first draft of the HIO: Diagnostics, Gene Product, Genetic Phenomena, Mode of Inheritance, Personal Attribute, Phenotype and Therapeutics. A new HIO identifier is assigned to each reused concept from other ontologies and attached to cross references to the source ontology. The SCDO's central 'Hemoglobinopathy' class has been replaced by a new central 'Hearing Impairment' class, which contains four main subclasses: 'Hearing Impairment by Cause', 'Hearing Impairment by Ear Affected', 'Hearing Impairment by Onset', and 'Hearing Impairment by Physiopathology Mechanism' (see Figure 4), which are comprehensively populated with the current HI domain knowledge, capturing various aspects associated with HI.

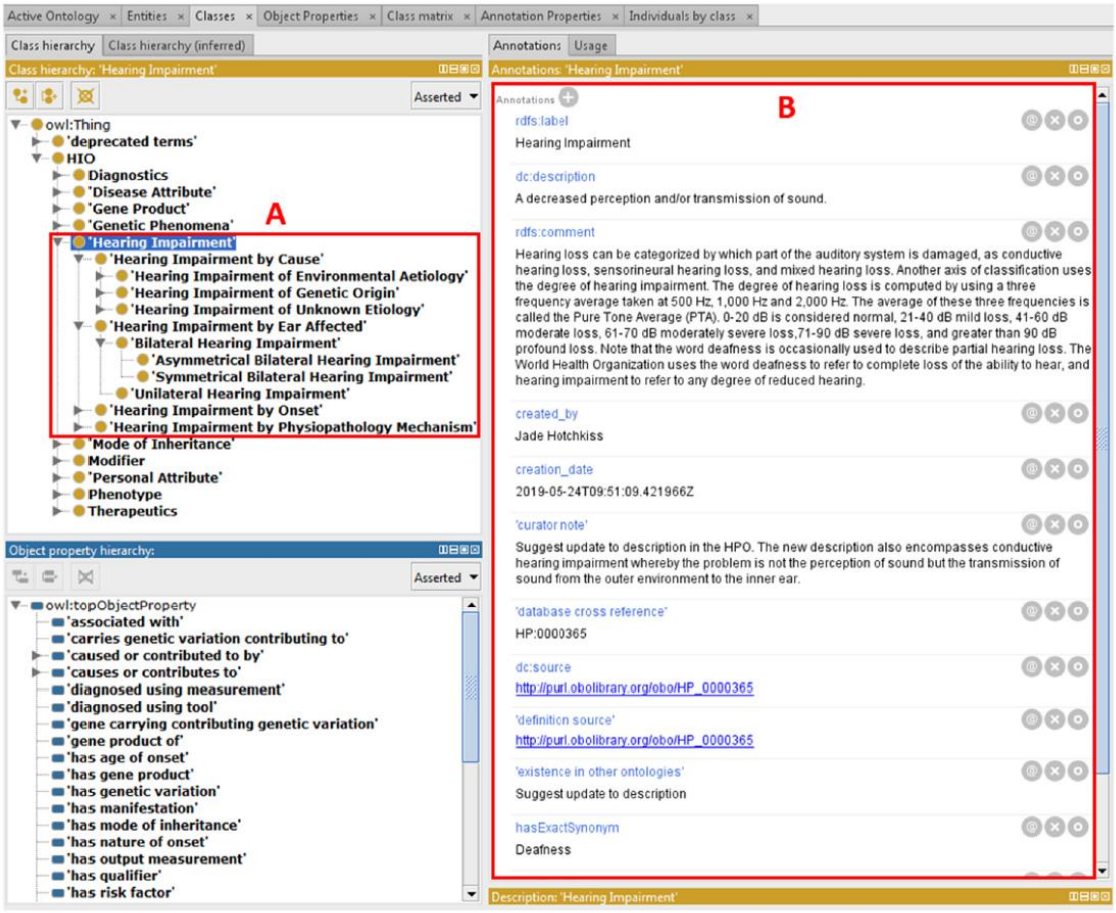


Figure 4. The 'Hearing Impairment' class within the HIO. (**A**) General categorization of hearing impairments in the 'Hearing Impairment' upper class and (**B**) annotations of the 'Hearing Impairment' class.

Two new (compared to the SCDO) upper classes have been included, namely: 'Disease Attribute' and 'Modifier' (see Figure 2). The 'Disease Attribute' class (see Figure 5) incorporates content similar to the SCDO's 'Association' class but notably also includes a 'Disease Cause' sub-class, which was found necessary due to the varied and often complex causes of hearing impairments [19–22]. The 'Disease Cause' class contains the term 'Unknown Etiology' and the two subclasses 'Environmental Disease Cause' and 'Intrinsic Disease Cause' [23–25], which are populated comprehensively with factors that cause or contribute in some way to HI.

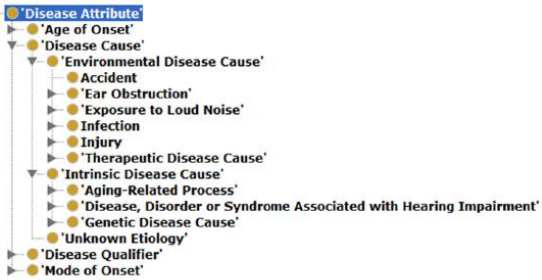


Figure 5. The 'Disease Attribute' class structure within the HIO. This hierarchy is intended to contain all possible features specific to or leading to HI.

The 'Modifier' class includes the SCDO's 'Disease Modifier' upper class as a sub-class, along with a new 'Disease Cause Modifier' class. This additional type of modifier was included because causes of hearing impairment sometimes have modifying factors that determine whether or not the disease

is in fact caused/present, e.g., ototoxicity induced by drugs (one cause of hearing impairment) has numerous modifying factors (e.g., dose, duration of therapy, concurrent renal failure, infusion rate, lifetime dose, coadministration with other drugs having ototoxic potential, genetic susceptibility) [26].

The HIO's central 'Hearing Impairment' class links, either directly or indirectly, to sub-classes of all other upper classes through numerous associations, as can be seen in Figure 6. For simplicity, associations made with terms in the 'Modifier' upper class are not shown (see these in Figure 7) and only associations with the 'Disease Cause' sub-class of the 'Disease Attribute' class are displayed. As shown by the large yellow c-shape in Figure 6, the 'Intrinsic Disease Cause' class encompasses portions of most of the other upper classes, namely 'Genetic Phenomena', 'Phenotype', 'Gene Product', 'Personal Attribute' and 'Therapeutics'.

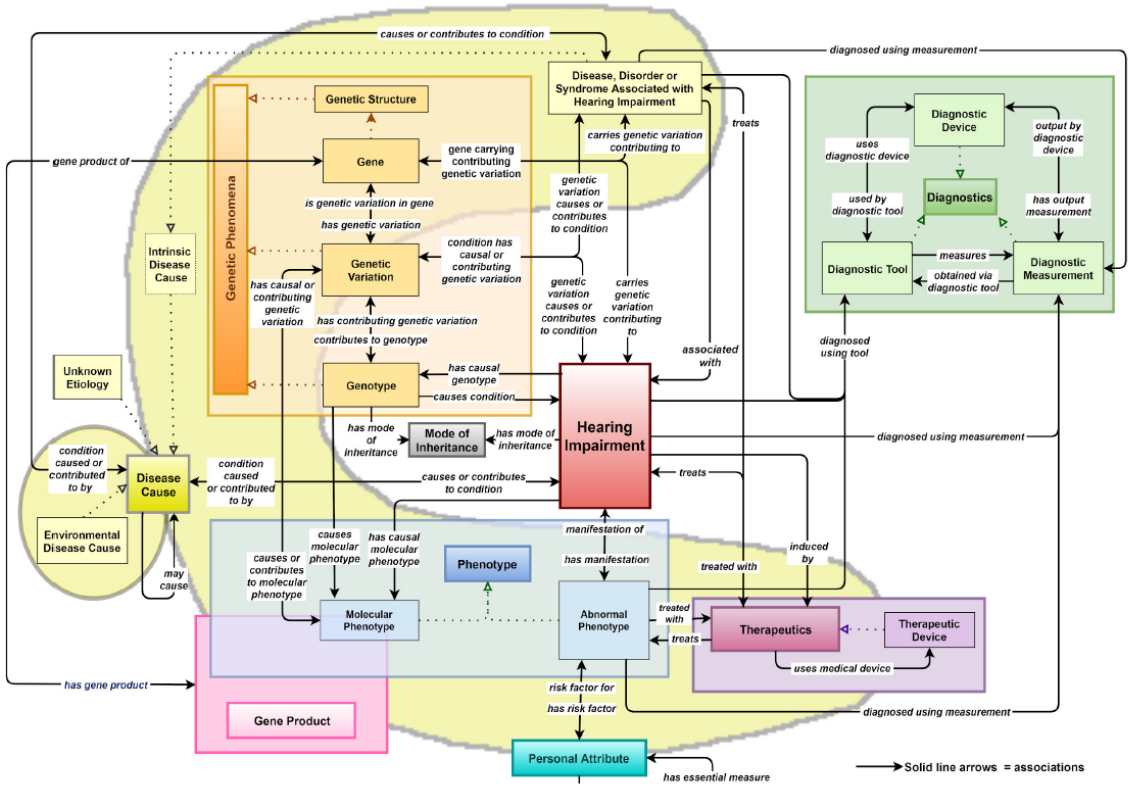


Figure 6. Associations made in the HIO between upper level (close to the root of the ontology) classes (excluding 'Modifier' class and only including 'Disease Cause' sub-class (yellow shapes) of the 'Disease Attribute' class). The 'Hearing Impairment' class is the central class.

The remaining associations made in the ontology, i.e., with the other disease attributes not in the 'Disease Cause' class (namely: 'Age of Onset', 'Mode of Onset', and 'Disease Qualifier' (which includes qualifiers such as 'Rare and 'Acquired')) and with the 'Modifier' class, are shown in Figure 7.

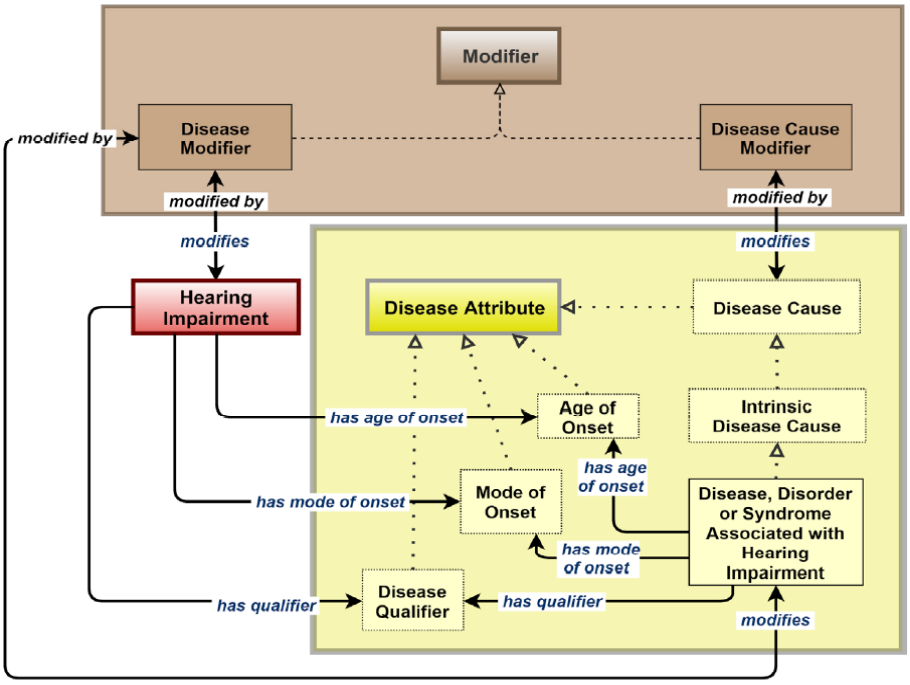


Figure 7. Associations made in the HIO to and from the 'Modifier' and 'Disease Attribute' classes, excluding those already shown in Figure 5 for the 'Disease Cause' class.

3.2. Assessing the Relevance of HIO

Of the 493 terms in the ontology (excluding 'deprecated terms' and upper 'owl:Thing'), at the time of writing this, 399 terms have descriptions and are considered as having all minimum required annotations (i.e., *rdfs:label*; *dc:description*; *definition source*, if relevant; *database cross reference*, if applicable; *dc:creator*, if the term was defined by an HIO curator using a source other than available ontologies; *existence in other ontologies*, used to record the existence status of the term prior to inclusion in the HIO; and *hasExactSynonyms*, to indicate synonyms, where relevant). Analysis of the existence statuses ascribed to these terms using the 'existence in other ontologies' annotation property (see Table 1) shows that the HIO, even in its first draft, is making contributions towards including and standardizing HI terms that were previously not included in other ontologies. Some of these terms are provided in Table 2 for illlustration. Where applicable, terms unique to the HIO will be recommended for inclusion into other related ontologies, such as the HPO, Orphanet, DO, MESH, and NCIT.

3.3. HIO Release and License

The HIO is released every two months with possible special releases when there are significant incidental changes. It is freely available under the Creative Commons Attribution 4.0 Unported License (CC:https://creativecommons.org/licenses/by/4.0/legalcode) and further copyrighted to maintain the quality and integrity of the term vocabulary, meaning that any modification to the HIO can only be done by HIO developers and curators.

3.4. Different HIO Access Platforms

In order to foster the dissemination of and easy access to this novel ontology the latest OWL file produced has been uploaded to the NCBO BioPortal at https://bioportal.bioontology.org/ontologies/HIO. This also facilitates the searching and viewing of different HIO concepts. In addition, the OWL and OBO files can be accessible via the GitHub repository at https://github.com/hiodev/hi-ontology.

Table 1. Summary of terms' existence statuses prior to inclusion in the HIO.

Existence Status	Explanation of Status	No. Terms	% Terms
Sufficient	Exists in other ontology and has appropriate description	284	71.2
Suggest update to description	Used term from existing ontology but will suggest they update their description to ours	27	6.8
Suggest update to label	Used term from existing ontology but will suggest they update their label to ours	0	0
Suggest update to label and description	Used term from existing ontology but will suggest they update their label and description to ours	0	0
Few but definitions not available	Term exists in a few ontologies but has not been given a description in any	3	0.8
Few but definitions not freely available	Term exists in a few ontologies but the description is not freely available	8	2.0
Few but definitions not specific enough	Term exists in a few ontologies but the definitions are not specific enough for the HIO's needs	9	2.3
Not relevant to context of hearing impairment	Term exists in other ontologies but the definitions are not relevant to the HI field	3	0.8
Negligible	No description or outdated ontology	2	0.5
None	Not in any existing ontology	63	15.8

Table 2. Some of terms that are unique to HIO.

Term Label	Term ID	Term Description
Symmetrical Bilateral Hearing Impairment	HIO:0000365	When the severity and configuration of hearing impairment is approximately the same in both ears.
Asymmetrical Bilateral Hearing Impairment	HIO:0000366	When each ear has a different severity and configuration of hearing impairment.
Postlingual Hearing Impairment	HIO:0000475	Hearing impairment which develops after the acquisition of speech and language, usually after the age of six.
Prelingual Hearing Impairment	HIO:0000476	Hearing impairment which is either congenital or develops before the acquisition of speech and language, usually before the age of 6.
Temporal Bone Fracture with Otic Capsule Involvement	HIO:0000287	Traumatic injury to the temporal bone in which the continuity of the bone is broken and violation of the otic capsule is involved.
Temporal Bone Fracture without Otic Capsule Involvement	HIO:0000288	Traumatic injury to the temporal bone in which the continuity of the bone is broken and violation of the otic capsule is not involved.
Pseudo-Dominant Inheritance	HIO:0000228	When the inheritance of a recessive trait mimics a dominant pattern of inheritance.
Cisplatin-Induced Hearing Impairment	HIO:0000215	Hearing loss caused by cisplatin (a chemotherapeutic agent) ototoxicity.
Neomycin-Induced Hearing Impairment	HIO:0000285	Partial or complete loss of hearing following ingestion of neomycin.
Maternal Medical History	HIO:0000362	A record of a patient's biological mother's background regarding health and the occurrence of disease events of the mother.
Hearing Impairment based on Immaturity	HIO:0000514	Hearing impairment that occurs due to premature birth (birth at or before 37 weeks of gestational age).

4. Discussion

Making use of ontological reasoning approaches may play a significant role in solving scalability and interoperability issues associated with current large-scale biological high-throughput datasets. This implies that building and maintaining biomedical ontologies is essential, especially in this current data rich era with an extensive consideration of big data analytics. With the contribution of HI domain experts, we have designed the HIO, which enables knowledge acquisition and harmonization, verification and validation of data available in different databases. This ontology is set to be the most comprehensive standardized HI domain knowledge portal, which will allow for the application of ontology-driven mining approaches for the identification of pertinent research questions. The HIO will foster clear and unambiguous communication and also facilitate sharing of information within the field.

4.1. HIO Structure, Other Disease Ontologies and HI Online Datasets

As pointed out previously, the HIO reuses concepts from other ontologies, including HPO, DO, and especially SCDO (see Table 1: summarizing the number of HI specific concepts vs reused concepts). These concepts were adjusted, where applicable, to incorporate new concepts specific to HI and relevant in various areas, such as HI subtypes, phenotypic expressions, genetic phenomena and different modes of inheritance. It is worth noting that, although we did not foresee all the adaptations that would be required, the use of the SCDO as a template for the HIO has been a very useful exercise. Whereas the general structure of the SCDO is more readily transferable to other monogenic diseases, we believe the HIO can be used as a template for designing disease-specific ontologies for diseases with a broader range of causes. Finally, note that there exist several online resources storing HI datasets and containing or enabling the retrieval of HI information. Table 3 lists some of these resources.

Table 3. Some existing online hearing impairment resources.

Scheme	Description	Types	URL	Reference
HHL	Hereditary Hearing Loss Homepage	An up-to-date overview of the genetics of hereditary hearing impairment for researchers and clinicians working in the field.	https://hereditaryhearingloss.org/	-
SHIELD	The Shared Harvard Inner Ear Laboratory Database	An integrative gene expression database for inner ear research	https://shield.hms.harvard.edu	[27]
DVD	Deafness Variation Database	A comprehensive resource integrating available genetic, genomic, and clinical data together with expert curation to generate a single classification for each variant in 152 genes implicated in syndromic and non-syndromic deafness.	http://deafnessvariationdatabase.org/	[28]
LOVD	Leiden Open Variation Databse	Retinal and hearing impairment genetic variant database	https://databases.lovd.nl/shared/genes/OTOF	[29]
NIDCD	National Institute on Deafness and Other Communication Disorders	A resource providing knowledge about Hearing, Ear Infections, and Deafness Diseases and Conditions. It also provides NIDCD Temporal Bone Registry at https://www.tbregistry.org/, a resource for learning about the pathology and pathophysiology of otologic disorders, which serves as a resource for scientists to analyze data from a collection of more than 12,000 temporal bone specimens.	https://www.nidcd.nih.gov/health/hearing-ear-infections-deafness	-
gEAR	Gene Expression Analysis Resource	Visualization and analysis of multiomic data both in public and private domains.	https://umgear.org/	-
OMIM	Online Mendelian Inheritance in Man	An Online Catalog of Human Genes and Genetic Disorders	https://www.omim.org	[30]

4.2. HIO Potential Future Applications

Even though we have shown the relevance of this new ontology by looking at how many of the classes are HI specific by querying against NCBO BioPortal, it should be noted that an ontology should be applied in order to appropriately assess its impact and suitability. We plan to use this ontology in data representation, which includes data harmonization, interoperability, and integration. For this, the HIO will be an essential resource in designing an ontology-based case report forms, providing essential data elements and controlled terminology. Different datasets can then be mapped to these data elements, making these datasets interoperable, thus easing the data integration process and meta-analysis. In the context of HI research, this will orient data analysis and enable the use of machine learning approaches with sufficient statistical power [31] for predicting disease clinical outcomes [32] and optimal therapeutic interventions, based on the disease pathophysiology mechanisms and other clinical parameters in patient records. It is expected that this HIO will contribute to fostering the subsequent HI research translation into healthcare, inferring knowledge based on patient clinical records, and the development of ontology-powered artificial intelligence medical tools helping in therapeutic interventions, prognosis, and diagnosis, as well as predictive models for an improved understanding of disease processes.

4.3. Challenges and Future Direction

Although the HIO has been designed in a manner that takes into account the various complexities of HI, there is admittedly still much content to be included, both with regards to terms and associations. New discoveries are also regularly being made in this field towards technological advances for diagnostics and therapeutic. This suggests that the ontology should be dynamic, in continuous evolution, keeping the HI knowledge up to date as new knowledge is accessible. There will thus be a need for ongoing input and maintenance of the HIO. For this, there is already a dedicated curation team aiming at assuring the quality and accuracy of the information contained in this ontology and also keeping it updated as HI knowledge evolves.

Going forward, these remaining upper classes of the SCDO will also be evaluated, and where necessary, adapted, for inclusion into the HIO: 'Research', 'Guidelines', and 'Quality of Life and Care'. We also plan to use competency questions defined by HI experts to evaluate the scope and domain coverage of the HIO. Beyond the use of competency questions, this ontology will also be assessed on its HI concept inclusion power, i.e., in terms of percentage of HI clinical terminologies from a given database, such as the GHLD database, or HI associated clinical reports or selected literature found in the HIO. This is particularly useful as it will provide an indication on the HIO ability in HI text mining tasks. Finally, the next critical challenge is to introduce this HIO into the dynamic clinical setting. This necessitates the development of testable and actionable health informatics applications to ensure clinical system-wide adherence. As indicated previously, HIO addresses the issue of unifying research clinical data from diverse sources. This ontology already paves the way towards the integration of clinical data into electronic medical records, which should facilitate the development of effective health informatics tools to potentially assist in the public and clinical management of hearing impairment conditions.

5. Conclusions

We have developed the HIO, a common controlled HI vocabulary, which is expected to enhance collaborative research. This ontology is currently the most comprehensive and standardized human- and machine-readable resource that unambiguously defines HI concepts and terminology for researchers, patients, and clinicians in order to help process, reuse, and re-apply existing HI knowledge in biomedical research and health-care systems. In the context of big data analytics, this ontology may facilitate retrospective data harmonization and contribute to mapping HI datasets to functional knowledge to enable the subsequent HI research translation into clinical applications and policy guidelines.

The HIO will allow researchers, clinicians, and patients to readily access standardized HI-related knowledge in a single location and promote HI data integration, interoperability, and sharing, including epidemiological, socio-environmental, biomedical, genetic, and phenotypic datasets.

Author Contributions: G.K.M., V.N., and A.W. are leading the H.I.O. implementation. J.H., N.M., E.W., V.N., N.V., G.K.M., and A.W. curated the current H.I.O. content. N.M., J.H., E.W., S.M.A., O.G.O., K.M., A.Y., and A.W. provided the H.I.O. content. G.K.M., N.J.M., M.H., N.V., S.J. contributed to the ontology structure development and implementation. G.K.M., J.H., N.M., and V.N. wrote the manuscript. All authors revised and approved the manuscript.

Funding: The development and implementation of the HIO is funded by the National Institutes of Health (NIH) Common Fund grants U01-HG-009716 and 1U01HG007459-01 to AW and the African Academy of Science/Wellcome Trust Ref H3A/18/001 and 1U54HG009790-01 to AW. Some developers and ontologists are supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Number U24HL135600 to AW. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funders.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chadha, S.; Cieza, A.; Reyes, K. Public health approach to hearing across the life course: A call-for-papers. *Bull. World Health Organ.* **2018**, *96*, 592. [CrossRef] [PubMed]
2. Murray, E.; McCabe, P.; Heard, R.; Ballard, K.J. Differential diagnosis of children with suspected childhood apraxia of speech. *J. Speech Lang. Hear. Res.* **2015**, *58*, 43–60. [CrossRef] [PubMed]
3. Vos, T.; Allen, C.; Arora, M.; Barber, R.M.; Bhutta, Z.A.; Brown, A.; Carter, A.; Casey, D.C.; Charlson, F.J.; Chen, A.Z.; et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: A systematic analysis for the global burden of disease study 2015. *Lancet* **2016**, *388*, 1545–1602. [CrossRef]
4. *Global Estimates on Hearing Loss*; The World Health Organization: Geneva, Switzerland, 2018. Available online: http://www.who.int/pbd/deafness/estimates/en/ (accessed on 27 June 2019).
5. Graydon, K.; Waterworth, C.; Miller, H.; Gunasekera, H. Global burden of hearing impairment and ear disease. *J. Laryngol. Otol.* **2019**, *133*, 18–25. [CrossRef] [PubMed]
6. WHO. *Global Costs of Unaddressed Hearing Loss and Cost-Effectiveness of Intervention*; The World Health Organization: Geneva, Switzerland, 2017.
7. The SCDO Working Group. The Sickle Cell Disease Ontology: Enabling universal sickle cell-based knowledge representation. *Database* **2019**, in press.
8. Löhler, J.; Walther, L.E.; Hansen, F.; Kapp, P.; Meerpohl, J.; Wollenberg, B.; Schönweiler, R.; Schmucker, C. The prevalence of hearing loss and use of hearing aids among adults in Germany: A systematic review. *Eur. Arch. Otorhinolaryngol.* **2019**, *276*, 945–956. [CrossRef]
9. Martinez-Cruz, C.; Blanco, I.J.; Vila, M.A. Ontologies versus relational databases: Are they so different? A comparison. *Artif. Intell. Rev.* **2011**, *38*, 271–290. [CrossRef]
10. Gruber, T.R. Toward principles for the design of ontologies used for knowledge sharing? *Int. J. Hum-Comput. St.* **1995**, *43*, 907–928. [CrossRef]
11. Mazandu, G.K.; Mulder, N.J. A topology-based metric for measuring term similarity in the Gene Ontology. *Adv. Bioinform.* **2012**, *2012*, 17. [CrossRef]
12. Mazandu, G.K.; Chimusa, E.R.; Mulder, N.J. Gene ontology semantic similarity tools: Survey on features and challenges for biological knowledge discovery. *Brief. Bioinform.* **2018**, *18*, 886–901. [CrossRef]
13. Kibbe, W.A.; Arze, C.; Felix, V.; Mitraka, E.; Bolton, E.; Fu, G.; Mungall, C.J.; Binder, J.X.; Malone, J.; Vasant, D.; et al. Disease Ontology 2015 update: An expanded and updated database of human diseases for linking biomedical knowledge through disease data. *Nucleic Acids Res.* **2015**, *43*, D1071–D1078. [CrossRef] [PubMed]
14. Köhler, S.; Vasilevsky, N.A.; Engelstad, M.; Foster, E.; McMurry, J.; Aymé, S.; Baynam, G.; Bello, S.M.; Boerkoel, C.F.; Boycott, K.M.; et al. The Human Phenotype Ontology in 2017. *Nucleic Acids Res.* **2017**, *45*, D865–D876. [CrossRef] [PubMed]
15. Köhler, S.; Carmody, L.; Vasilevsky, N.; Jacobsen, J.O.B.; Danis, D.; Gourdine, J.P.; Gargano, M.; Harris, N.L.; Matentzoglu, N.; McMurry, J.A.; et al. Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. *Nucleic Acids Res.* **2019**, *47*, D1018–D1027. [CrossRef] [PubMed]

16. Matentzoglu, N.; Malone, J.; Mungall, C.; Stevens, R. MIRO: Guidelines for minimum information for the reporting of an ontology. *J. Biomed. Semant.* **2018**, *9*, 6. [CrossRef]
17. Horridge, M.; Gonçalves, R.S.; Nyulas, C.; Musen, M.A. WebProtégé: A Cloud-Based Ontology Editor. *arXiv* **2019**, arXiv:1902.08251.
18. Jackson, R.C.; Balhoff, J.P.; Douglass, E.; Harris, N.L.; Mungall, C.J.; Overton, J.A. ROBOT: A Tool for Automating Ontology Workflows. *BMC Bioinform.* **2019**, *20*, 407. [CrossRef]
19. Nance, W.E. The genetics of deafness. *Ment. Retard. Dev. Disabil. Res. Rev.* **2003**, *9*, 109–119. [CrossRef]
20. Adadey, S.M.; Awandare, G.; Amedofu, G.K.; Wonkam, A. Public Health Burden of Hearing Impairment and the Promise of Genomics and Environmental Research: A Case Study in Ghana, Africa. *OMICS* **2017**, *21*, 638–646. [CrossRef]
21. Fook, L.; Morgan, R. Hearing impairment in older people: A review. *Postgrad. Med. J.* **2000**, *76*, 537–541. [CrossRef]
22. Szyfter, W.; Wróbel, M.J.; Szyfter-Harris, J.; Greczka, G. Hearing impairment in Polish infants. *Epidemiology* **2013**, *24*, 333. [CrossRef]
23. Cunningham, L.L.; Tucci, D.L. Hearing Loss in Adults. *N. Engl. J. Med.* **2017**, *377*, 2465–2473. [CrossRef] [PubMed]
24. Rudman, J.R.; Kabahuma, R.I.; Bressler, S.E.; Feng, Y.; Blanton, S.H.; Yan, D.; Liu, X.Z. The genetic basis of deafness in populations of African descent. *J. Genet. Genom.* **2017**, *44*, 285–294. [CrossRef] [PubMed]
25. Angela, S.; Lin, X.; Liu, X.Z. Genetics of Hearing and Deafness. *Anat. Rec.* **2012**, *295*, 1812–1829. [CrossRef] [PubMed]
26. Lustig, L.R. Drug-Induced Ototoxicity. 2018. Available online: https://www.msdmanuals.com/professional/ear,-nose,-and-throat-disorders/inner-ear-disorders/drug-induced-ototoxicity (accessed on 28 June 2019).
27. Shen, J.; Scheffer, D.I.; Kwan, K.Y.; Corey, D.P. SHIELD: An integrative gene expression database for inner ear research. *Database* **2015**, *2015*, bav071. [CrossRef] [PubMed]
28. Azaiez, H.; Booth, K.T.; Ephraim, S.S.; Crone, B.; Black-Ziegelbein, E.A.; Marini, R.J.; Shearer, A.E.; Sloan-Heggen, C.M.; Kolbe, D.; Casavant, T.; et al. Genomic Landscape and Mutational Signatures of Deafness-Associated Genes. *Am. J. Hum. Genet.* **2018**, *103*, 484–497. [CrossRef]
29. Fokkema, I.F.; Taschner, P.E.; Schaafsma, G.C.; Celli, J.; Laros, J.F.; den Dunnen, J.T. LOVD v.2.0: The next generation in gene variant databases. *Hum. Mutat.* **2011**, *32*, 557–563. [CrossRef]
30. Amberger, J.S.; Bocchini, C.A.; Scott, A.F.; Hamosh, A. OMIM.org: Leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Res.* **2019**, *47*, D1038–D1043. [CrossRef]
31. Mazandu, G.K.; Kyomugisha, I.; Geza, E.; Seuneu, M.; Bah, B.; Chimusa, E.R. *Designing Data-Driven Learning Algorithms: A Necessity to Ensure Effective Post-Genomic Medicine and Biomedical Research, Artificial Intelligence—Applications in Medicine and Biology*; IntechOpen: London, UK, 2019. Available online: https://www.intechopen.com/books/artificial-intelligence-applications-in-medicine-and-biology/designing-data-driven-learning-algorithms-a-necessity-to-ensure-effective-post-genomic-medicine-and- (accessed on 16 September 2019).
32. Zhao, Y.; Li, J.; Zhang, M.; Lu, Y.; Xie, H.; Tian, Y.; Qiu, W. Machine Learning Models for the Hearing Impairment Prediction in Workers Exposed to Complex Industrial Noise: A Pilot Study. *Ear Hear.* **2019**, *40*, 690–699. [CrossRef]

2. Noluthando Manyisa, Samuel Mawuli Adadey, Edmond Wonkam Tingang, **Abdoulaye Yalcouye**, Ambroise Wonkam. Hearing Impairment in South Africa and the Lessons Learned for Planetary Health Genomics: A Systematic Review. OMICS. 2022 Jan;26(1): 2-18.doi: 10.1089/omi.2021.0181. PMID: 35041532.

Authors' Contributions

N.M. performed the literature search and drafted and revised the article; S.M.A, E.T.W., and **Abdoulaye Yalcouye**. double checked the article selection and performed the quality control analysis. A.W. conceived and supervised the entire project; all authors have read and agreed to the final version of the article, and made a significant intellectual contribution for authorship.

Information