Assessing non-Mendelian Inheritance in Inherited Axonopathies

Dana M. Bis-Brewer[1], Ziv Gan-Or[3,4,12], Patrick Sleiman[2], Inherited Neuropathy Consortium, Hakon Hakonarson[2], Sarah Fazal[1], Steve Courel[1], Vivian Cintra[1], Feifei Tao[1], Mehrdad A. Estiar[3,4], Mark Tarnopolsky[5], Kym M. Boycott[6], Grace Yoon[7,8], Oksana Suchowersky[9], Nicolas Dupré[10,11], Andrew Cheng[13], Thomas E. Lloyd[13], Guy Rouleau[3,4,12], Rebecca Schüle, Stephan Züchner[1]

1. Dr. John T. Macdonald Foundation Department of Human Genetics, John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL, USA.
2. Center for Applied Genomics, The Children's Hospital of Philadelphia; Division of Human Genetics, Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.
3. Department of Human Genetics, McGill University, Montréal, Quebec, Canada.
4. Montreal Neurological Institute and Hospital, McGill University, Montréal, Quebec, Canada.
5. Neuromuscular and Neurometabolics Division, Department of Pediatrics, McMaster University, Hamilton, Ontario, Canada.
6. Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, Ontario, Canada.
7. Division of Clinical and Metabolic Genetics, Department of Paediatrics, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada.
8. Division of Neurology, Department of Paediatrics, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada.
9. Department of Medicine, Medical Genetics and Pediatrics. University of Alberta, Edmonton AB, Canada.
10. Division of Neurosciences, CHU de Québec, Université Laval, Québec City, QC, Canada.
11. Department of Medicine, Faculty of Medicine, Université Laval, Québec City, QC, Canada.

12. Department of Neurology and Neurosurgery, McGill University, Montréal, Quebec, Canada.
13. Department of Neurology and Neuroscience, School of Medicine, Johns Hopkins University, Baltimore, MD, USA

Disclosure

Nothing to disclose.

Key words

Charcot-Marie-Tooth disease, Hereditary Spastic Paraplegia, inherited axonopathy, mutational burden, oligogenic inheritance, risk alleles

Corresponding author

Stephan Zuchner, MD, PhD
University of Miami Miller School of Medicine
Biomedical Research Building (BRB)
Room 616, LC: M-860
1501 NW 10th Avenue
Miami, FL 33136
Phone 305-243-2281

Abstract

Inherited axonopathies (IA) are rare, clinically and genetically heterogeneous diseases that lead to length-dependent degeneration of the long axons in central and peripheral nervous systems. Progressive axonal degeneration can lead to both Charcot-Marie-Tooth type 2 (CMT2) and hereditary spastic paraplegia (HSP) depending on whether the peripheral or central nerves are affected, respectively. Mendelian high-penetrance alleles in over one hundred different genes have been shown to cause IA; yet, about 50% of IA cases do not receive a genetic diagnosis. A more comprehensive spectrum of causative genes and alleles is warranted, including causative and risk alleles, as well as oligogenic multilocus inheritance. Through international collaboration, IA exome studies are beginning to be sufficiently powered to perform a pilot rare variant burden analysis. After extensive quality control, our cohort contained 343 CMT cases, 515 HSP cases, and 935 non-neurological controls. We assessed the cumulative mutational burden across disease genes, explored the evidence for multilocus inheritance, and performed an exome-wide rare variant burden analysis. We replicated the previously described mutational burden in a much larger cohort of CMT cases, and observed the same effect in HSP cases. We identified a preliminary risk allele for CMT in the *EXOC4* gene (*p*-value= 6.9x10-6, OR=2.1) and explored the possibility of multi-locus inheritance in IA.

Introduction

Inherited axonopathies (IA) are a group of disorders unified by a common pathological mechanism: length-dependent axonal degeneration. They are traditionally classified into two broad genetic disorders: Hereditary Spastic Paraplegia (HSP) and Charcot-Marie-Tooth disease (CMT) depending on upper or lower motor neuron involvement, respectively. Historically, CMT and HSP are treated as distinct disorders, but their increasingly apparent clinical and genetic overlap challenges this classification. CMT and HSP can be caused by mutations within the same gene (for example, *KIF1A, REEP1,* and *BSCL2*), yet the additional factors which determine peripheral or central nerve involvement affected in each IA patient remains unclear. Currently, more than 50% of cases do not receive a genetic diagnosis from whole-exome sequencing (WES).[1] The percentage of genetically undiagnosed IA cases suggests that rare mutational mechanisms or modes of inheritances may be overlooked in standard WES analysis.

The phenotypic variability and reduced penetrance observed within IA supports the possibility of multilocus inheritance or genetic modification. These two events are closely related and both result in phenotypic effects that are caused by more than a single Mendelian allele. A distinction between these processes lies in the sufficiency of the primary allele to cause disease.[2] If the presence of the primary allele alone manifests the phenotype, then the secondary allele is a genetic modification of the phenotype, such as the severity of progression or the age at onset.[2] However, if the presence of an allele in a second gene or multiple genes is required to cause disease, then inheritance is multilocus in nature.[2] Non-Mendelian modes of inheritance have been independently

demonstrated in both CMT and HSP[3,4], but cohort-level sequencing analyses are limited. In this pilot study, we gathered over 800 exomes from IA cases to determine whether multilocus inheritance warrants deeper investigation in classically Mendelian disease groups.

Materials and Methods

The cases were collected from the Inherited Neuropathy Consortium, the University of Miami (UM), Children's Hospital of Philadelphia (CHOP), the University Hospital Tübingen, and McGill University, and all participating individuals gave informed consent prior to initiating this study in agreement with the institutional review boards. Families included in the study are affected by IA (either CMT or HSP). CMT cases were diagnosed with CMT (type 1, 1A, 2, 4, or intermediate), distal hereditary motor neuropathy, hereditary sensory autonomic neuropathy, or hereditary sensory neuropathy; HSP cases were diagnosed with pure or complicated HSP. WES was performed at UM (CMT and HSP cases), McGill (HSP cases), and at CHOP (controls). Enrolled cases had previously negative testing for key IA genes; however, solved research cases were included in the cohort. All samples were sequenced on Illumina HiSeq2000 and joint-genotyped according to the GATK (v.3.3) germline WES best practices. After extensive quality control (including duplication percentage, sex and relatedness, depth and missingness metrics, ancestry), the cohort contained 343 CMT cases, 515 HSP cases, and 935 non-neurological controls.

To detect risk alleles, a gene-based rare variant association test was performed by the C-alpha test in the PLINK/SEQ suite. Following recommended protocol, tests with an i-statistic greater than 10^{-3} were removed, and Bonferroni correction was applied.[5] To compare the mutational burden across known disease genes (CMT:*n*=88, HSP:*n*=95), the number of rare variants (non-synonymous or loss-of-function at ExAC MAF $\leq$ 0.01 and $\leq$ 0.001) within disease genes was computed for each sample, and the average counts were compared between case and control using a Mann-Whitney-Wilcoxon test followed by 10,000 iterations of affection permutation for significance. To assess multilocus inheritance, the number of disease genes carrying at least one qualifying variant was determined per sample. Case and control carrier status was organized into 2x2 contingency tables and assessed by Fisher's exact test.

Results

Association of EXOC4 with CMT cases

Exome-wide association analysis was performed at 17,637 protein coding loci by the C-alpha test. The PLINK/SEQ suite computes an estimate of the minimal achievable *p*-value for a locus, the i-statistic. We followed the recommended protocol to filter out loci with an i-statistic greater than 10^{-3} before Bonferroni correction to remove non-contributing genes.[5] Based on the 2,145 remaining loci, the *p*-value threshold for an experiment-wide significance (alpha=0.05) was 2.3×10^{-5}. After filtering results by the PLINK/SEQ i-statistic and applying Bonferroni multiple-testing correction, three genes, *KDM5A* (*p*-value= 9.9×10^{-7}, OR=3.6), *EXOC4* (*p*-value= 6.9×10^{-6}, OR=2.6), and *CEP78* (*p*-value= 2.3×10^{-5}, OR=4.4), reached experiment-wide significance (Fig. 1A). *KDM5A* and *EXOC4* both

contained a single allele in cases that drove the association: NM_001042603.1(*KDM5A*):c.11T>G and NM_021807.3(*EXOC4*):c.1648G>A. Sanger sequencing confirmed the driver allele in *EXOC4* (Fig. 1B) and revealed a false call in *KDM5A*. We did not follow up with *CEP78* since the gene did not contain a single driving allele. At the *EXOC4* gene-level, heterozygous carriers were 2.6 (95% CI: 1.28-5.37) times more likely to be affected, while at the driver allele-level, heterozygous carriers were 9.07 times more likely to be affected (95% CI: 2.94-28.01) (Fig. 1C). The variant was a non-synonymous missense change (p.Gly550Arg) predicted to be disease causing by MutationTaster[15] with a gnomAD MAF of 0.00411.

Increased mutational burden across known disease genes in IA cases

IA cohorts were independently tested for a mutational burden (an excess of rare variants) across known disease genes. In our CMT and HSP cohorts, we identified a significant mutational burden (Mann-Whitney, nominal p-value ≤ 0.05) in each tested variant set (non-synonymous and loss-of-function variation at ExAC MAF ≤ 0.001 or ≤ 0.01; (Fig. 1D-E). As a further test of our observations, we repeated the mutational burden comparison with permutated case/control status over 10,000 iterations. We found that each tested variant set remained statistically significant (empirical p-value ≤ 0.05), thus supporting that the mutational burden found across disease genes is specific to each IA cohort.

Multilocus inheritance suggested in IA cases

Next, we sought to determine whether the observed mutational burden was more likely to follow a monogenic (single gene), digenic (two genes), or oligogenic (more than two

genes) inheritance. Unlike the mutational burden, the significance of each inheritance type was influenced by the minor allele frequency (Fig. 1F-G). HSP cases showed consistent evidence for oligogenic inheritance (≥ 3 genes) of non-synonymous (NS) variation and monogenic inheritance (1 gene) of loss-of-function (LoF) variation at both ExAC MAF ≤ 0.01 and ≤ 0.001 (Fisher's exact, *p*-value ≤ 0.05). HSP cases also displayed significant di/oligogenic inheritance (≥ 2 genes) of NS variation at the less common ExAC MAF ≤ 0.001 (*p*-value ≤ 0.05). Furthermore, di/oligogenic inheritance of both NS and LoF variation for HSP cases is suggested at ExAC MAF ≤ 0.01 (*p*-value = 0.0598 and 0.0572, respectively). Evidence for inheritance types in CMT was not as consistent as in HSP, possibly due to a lower CMT sample size. At ExAC MAF ≤ 0.01, CMT cases demonstrated monogenic inheritance for LoF variation and oligogenic inheritance for NS variation (*p*-value ≤ 0.05) with potential di/oligogenic inheritance for NS variation (p-value = 0.521). Lastly, at ExAC MAF ≤ 0.001, CMT cases only showed significant evidence for monogenic inheritance of NS variation (*p*-value ≤ 0.05) with potential evidence for oligogenic NS inheritance and monogenic LoF inheritance (*p*-value = 0.0641 and 0.0536, respectively). The counts of samples carrying variants is summarized in Fig. 1H-I.

Discussion

As the cost and availability of next-generation sequencing continues to drop, we are now reaching large enough sample sizes to apply statistical approaches to rare diseases. In this study, we sought to assess the mutational burden and multilocus involvement of rare variation in a cohort of inherited axonopathies as well as identify potential risk loci.

To identify genes that could potentially carry non-Mendelian risk alleles, we performed an unbiased exome-wide rare variant burden analysis with the C-alpha test. After filtering results and performing Sanger sequencing, *EXOC4* stood out as a candidate CMT gene. *EXOC4* is involved in vesicle transport and membrane tethering in polarized cells and is expressed in Schwann cells.[6] In a CMT4B1 mouse model, Exoc4 (Sec8) formed a complex with Mtmr2 and Dlg1 to coordinate homeostatic control of myelination.[6] Exoc4 is abundantly expressed at the *Drosophila* neuromuscular junction and required for *in vivo* regulation of synaptic microtubule formation.[7] Furthermore, Exoc4 is suggested to play a central role in oligodendrocyte membrane formation through the regulation of vesicular transport of myelin proteins.[8] Though *EXOC4* has biologically plausibility, this result should be interpreted with a degree of caution. Stronger genetic evidence for *EXOC4* is necessary, including replication of the association or identification of highly penetrant Mendelian variants. Unfortunately, a secondary large CMT exome cohort does not currently exist for follow-up replication analysis.

From the rare variant burden analysis, we were also able to re-identify several established monogenic CMT2 genes, including *MME*[9], *MORC*[16] and *MFN2*[10]. This is despite a general effort to exclude cases with *MFN2* and other common CMT genes from exome analysis. Similarly, known familial ALS genes showed strong associations in a gene-based rare variant burden analysis of sporadic ALS cases.[11] These results give us confidence about the utility of association studies in rare disease cohorts, and may indicate the presence of additional risk alleles contributing to the phenotype in these known CMT genes. We interpret these results as additional evidence supporting cohort-

level statistical approaches to identify Mendelian and non-Mendelian factors involved in classically monogenic disease.

Additionally, we observed a significant mutational burden across CMT and HSP disease genes in cases compared to non-neurological controls. The aggregation of rare, damaging alleles in disease-associated genes may contribute to risk, severity, and clinical heterogeneity. This inheritance model has been suggested in CMT based on exome sequencing from 37 individuals.[12] Gonzaga-Jauregui *et al* observed an average of 1.8 variants per case across 58 neuropathy genes compared to 1.3 variants per control. They followed up this observation with a second small cohort of 32 cases from Turkish descent, and observed a mutational burden of 2.1 vs 1.6 nonsynonymous rare variants per in cases vs controls. The mutational burden hypothesis was functionally evaluated *in vivo* in zebrafish experiments, which resulted in increased phenotypic severity when pairs of neuropathy genes were inactivated.[11] Our cohort is roughly 10 times larger than the previous cohorts, and is now the third independent CMT cohort to support the mutational burden hypothesis. Furthermore, rare non-synonymous variation was also significantly distributed across 2 or 3 disease genes in our cohort, indicating multilocus inheritance – which remains underexplored in rare diseases because of functional validation challenges. Additional variants in multiple disease genes can have either a combinatorial effect on the same biological pathways or a destabilizing effect on the entire disease module.

The primary goal of this study was to move beyond the 'one-disease-one-gene' model to assess an expanded genetic architecture in IA. An appreciation for the extent of allelic and locus heterogeneity, reduced penetrance, and variable expressivity within IA

has come from traditional family-based approaches. These insights across Mendelian diseases are driving the genetics community to delineate the more complicated and nuanced patterns of inheritance: 1. A gene-based variant burden test was successfully applied to a cohort of ALS cases and identified a new risk gene.[11] Using this approach, we observed an enrichment of qualifying variants (in a candidate gene and in known disease genes) that influence disease risk. We are extremely cautious about overstating any potential involvement of *EXOC4* in disease pathogenesis. However, we interpret these results as evidence supporting the hypothesis of 'risk alleles' in IA. 2. A mutational burden that can modulate phenotypic severity was observed in two small CMT cohorts, and the increased burden of protein-altering variants was functionally tested in a zebrafish model and demonstrated phenotypic modification.[11] We have replicated this finding in a larger cohort of CMT cases and discovered a similar result in HSP cases. Beyond CMT and HSP, a rare variant aggregation has also been show to influence susceptibility to Parkinson disease, and the age-of-onset of ALS.[13]

Concluding Remarks

Concepts such as risk alleles, mutational burden, and multilocus inheritance within rare Mendelian diseases lie at the intersection of rare and common diseases. Recent discoveries have shed light on the architecture of common disease, including increased risk for a common disease from heterozygous alleles in recessive Mendelian genes.[14] However, the impacts of multilocus inheritance on Mendelian disease, including phenotypic severity, oligogenic inheritance, blended phenotypes, and phenotypic expansion, requires further exploration.[13] Investigating these non-Mendelian concepts will

lead to a unified model of human disease and facilitate precision genetic therapies. Here, we continue pushing these boundaries in IA, suggest potential involvement of *EXOC4* in disease pathogenesis of CMT, and provide further evidence supporting a multilocus Mendelian model.

Acknowledgements

We thank the patients and their families for their participation. This work was done as a part of a Canadian collaboration to study HSP (CanHSP). The CanHSP cohort was funded by CIHR Emerging Team Grant, in collaboration with the Canadian Organization for Rare Disorders (CORD), grant number RN127580 – 260005.

References

1. Murphy SM, Laurá M, Fawcett K, et al. Charcot–Marie–Tooth disease: frequency of genetic subtypes and guidelines for genetic testing. *J Neurology Neurosurg Psychiatry*. 2012;83(7):706 710. doi:10.1136/jnnp-2012-302451

2. Eldomery MK, Coban-Akdemir Z, Harel T, et al. Lessons learned from additional research analyses of unsolved clinical exome cases. *Genome Med*. 2017;9(1):26. doi:10.1186/s13073-017-0412-6

3. Bis-Brewer DM, Zuchner S. Genetics modifiers and non-Mendelian aspects of CMT. *Brain Research.* 2020;1726:146459. doi:10.1016/j.brainres.2019.146459

4. Bis-Brewer DM, Zuchner S. Perspectives on the Genomics of HSP Beyond Mendelian Inheritance. *Front Neurol*. 2018;9:25 11. doi:10.3389/fneur.2018.00958

5. Kiezun A, Garimella K, Do R, et al. Exome sequencing and the genetic basis of complex traits. *Nat Genet*. 2012;44(6):623 630. doi:10.1038/ng.2303

6. Bolis A, Coviello S, Visigalli I, et al. Dlg1, Sec8, and Mtmr2 Regulate Membrane Homeostasis in Schwann Cell Myelination. *J Neurosci*. 2009;29(27):8858 8870. doi:10.1523/jneurosci.1423-09.2009

7. Liebl FL, Chen K, Karr J, Sheng Q, Featherstone DE. Increased synaptic microtubules and altered synapse development in Drosophila sec8 mutants. *Bmc Biol*. 2005;3(1):27. doi:10.1186/1741-7007-3-27

8. Anitei M. A role for Sec8 in oligodendrocyte morphological differentiation. *J Cell Sci*. 2006;119(5):807 818. doi:10.1242/jcs.02785

9. Auer-Grumbach M, Toegel S, Schabhüttl M, et al. Rare Variants in MME, Encoding Metalloprotease Neprilysin, Are Linked to Late-Onset Autosomal-Dominant Axonal Polyneuropathies. *Am J Hum Genetics*. 2016;99(3):607 623. doi:10.1016/j.ajhg.2016.07.008

10. Zuchner S, Mersiyanova I, Muglia M, et al. Mutations in the mitochondrial GTPase mitofusin 2 cause Charcot-Marie-Tooth neuropathy type 2A. *Nat Genet*. 2004;36(5):449 451. doi:10.1038/ng1341

11. Cirulli ET, Lasseigne BN, Petrovski S, et al. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science*. 2015;347(6229):1436 1441. doi:10.1126/science.aaa3650

12. Gonzaga-Jauregui C, Harel T, Gambin T, et al. Exome Sequence Analysis Suggests that Genetic Burden Contributes to Phenotypic Variability and Complex Neuropathy. *Cell Reports*. 2015;12(7):1169 1183. doi:10.1016/j.celrep.2015.07.023

13. Posey JE, O'Donnell-Luria AH, Chong JX, et al. Insights into genetics, human biology and disease gleaned from family based genomic studies. *Genet Med*. 2019;21(4):798-812. doi:10.1038/s41436-018-0408-7

14. Ji W, Foo J, O'Roak BJ, et al. Rare independent mutations in renal salt handling genes contribute to blood pressure variation. *Nat Genet*. 2008;40(5):592-599. doi:10.1038/ng.118