4. Müller T, Schlegel E, Zingler S, Thiede HM. Effects of one-day application of levodopa/carbidopa/entacapone versus levodopa/carbidopa/Opicapone in Parkinson's disease patients. Cell 2022;11(9):1511. https://doi.org/10.3390/Cells11091591
5. Isobe C, Murata T, Sato C, Terayama Y. Increase of total homocysteine concentration in cerebrospinal fluid in patients with Alzheimer's disease and Parkinson's disease. Life Sci 2005;77(15):1836–1843.
6. Müller T, Fowler B, Kuhn W. Levodopa intake increases plasma levels of S-adenosylmethionine in treated patients with Parkinson disease. Clin Neuropharmacol 2005;28(6):274–276.

Exploring *THAP11* Repeat Expansion beyond Chinese-Ancestry Cohorts: An Examination of 1000 Genomes and UK Biobank Data

Tan et al report novel CAG repeat expansion in *THAP11* associated with spinocerebellar ataxia (SCA) in two Chinese families. They observed 45–100 repeats, three CAA interruptions, and a long uninterrupted 3' tail in sequencing of ataxic individuals from a single family.[1]

We investigated presence and size of *THAP11* expansion in short-read next-generation sequencing of individuals with other ancestries in 1000 Genomes and the UK Biobank.[2] Repeat genotypes of up to 50 triplets can be accurately typed with short read sequencing and bioinformatic tools.[3]

Methods

We genotyped CAG expansion in *THAP11* (GRCh38 position chr16:67842863-67842950) with ExpansionHunter 5.0.0 and REViewer[4] in 138 individuals with whole-genome sequencing in the UK Biobank[2] and a history of hereditary ataxia or ataxia of unknown cause (ICD10 G110-G119; R270). We obtained *THAP11* ExpansionHunter genotypes from 2504 unrelated individuals in 1000 Genomes.[5]

Results

One European-ancestry ataxic individual in the UK Biobank has a 46/29 CAG *THAP11* genotype, exceeding the proposed pathogenic threshold,[1] and an uninterrupted 22-repeat CAG repeat expansion in *CACNA1A* that likely causes spinocerebellar ataxia type 6 (SCA6).[6] REViewer visualization of ExpansionHunter genotyping shows six CAA interruptions to the *THAP11* expansion (Fig. 1A) and no interruptions in the SCA6 expansion (Fig. 1B).

The individual had primary care diagnoses of hereditary ataxia at 40, Parkinson's disease at 53, and received repeated prescriptions for trihexyphenidyl, a drug used in management of movement disorders. More-precise age of onset is not available as medical records in the UK Biobank are partial and participants are not re-contactable.

THAP11 genotypes in the 1000 Genomes cohort were between 19 and 39 repeats. In the UK Biobank ataxia cohort, the range was 25 to 46 repeats. In both cohorts, the median was 29 repeats (IQR, 28–29) (Fig. 1C).

Discussion

We report the first finding of *THAP11* CAG expansion in an ataxic individual of European ancestry. Interpretation of this expansion is complicated by detection of an uninterrupted pathogenic full-penetrance length SCA6 expansion and diagnosis of both ataxia and Parkinson's disease. This individual has six CAA interruptions, with nine uninterrupted 3′ repeats. We also provide length distributions of the *THAP11* allele.

Tan et al suggest toxicity of CAA-interrupted repeats based on CAG-pure sequences of ataxic individuals. They report one family where ataxic individuals with *THAP11* expansion have three interruptions and 32 to 87 uninterrupted repeats in the 3′ end of the expansion, and unaffected family members have five to six interruptions and shorter tails. They also report *THAP11* expansion in an unrelated individual (patient II-1), with six interruptions and 10 uninterrupted 3′ repeats (Tan et al's supporting information Figure S3),[1] similar to our European-ancestry individual.

Studies suggest CAA interruptions to CAG expansions stabilize intergenerational variability in repeat length,[7] which may contribute to expansion instability in the family with fewer interruptions. Our findings show further work is necessary to elucidate the role of rare *THAP11* expansion and its composition in ataxia and demonstrates that *THAP11* expansion is detectable with bioinformatic approaches. It also further highlights the emerging complexity of expansion composition in tandem repeat-mediated disease.

Acknowledgments: We thank Professor Paul Lockhart, Dr. Justin Read and Kayli Davies of the Murdoch Children's Research Institute, Melbourne, Australia for comments and feedback. This research was conducted with data from UK Biobank (www.ukbiobank.ac.uk), a

Key Words: ataxia, population studies, biobanks, repeat expansions, bioinformatic screening

***Correspondence to:** Dr. Melanie Bahlo, Population Health and Immunity Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria 3052, Australia; E-mail: bahlo@wehi.edu.au

Relevant conflicts of interest/financial disclosures: The authors have no conflicts of interest to declare and no financial disclosures to report.

Received: 20 May 2023; **Revised:** 15 August 2023; **Accepted:** 29 August 2023

Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.29636

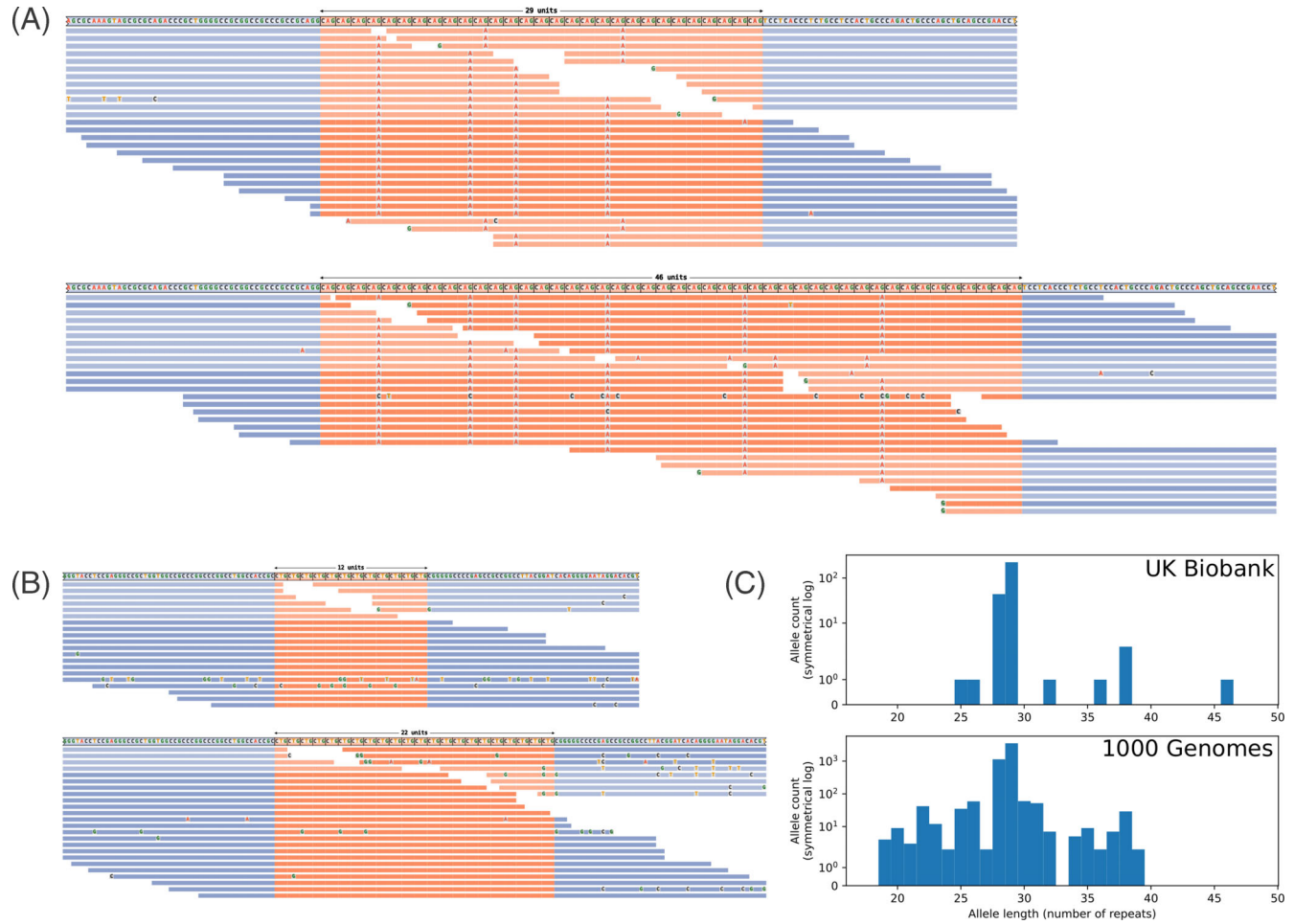

FIG.. 1. (**A**) REViewer plot of *THAP11* CAG expansion in a European ancestry participant in the UK Biobank. CAA interruptions are present in the expanded allele. (**B**) REViewer plot of expanded *CACNA1A* allele in the same individual. (C) Distribution of *THAP11* allele sizes in 139 individuals with ataxia in the UK Biobank and 2504 unrelated individuals in the 1000 Genomes cohort.

major biomedical database, under data use agreement 36610 (PI Bahlo). Open access publishing facilitated by The University of Melbourne, as part of the Wiley - The University of Melbourne agreement via the Council of Australian University Librarians.

Data Availability Statement

The data that support the findings of this study are available from the UK Biobank. Restrictions apply to the availability of UK Biobank data. Data from the Illumina Repeat Catalog are openly available at https://github.com/Illumina/RepeatCatalogs.

Liam G. Fearnley, PhD,[1,2,3] Haloom Rafehi, PhD,[1,2,3] Mark F. Bennett, PhD,[1,2,4] and Melanie Bahlo, PhD[1,2*]

[1]*Population Health and Immunity Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria, Australia,* [2]*Department of Medical Biology, The University of Melbourne, Parkville, Victoria, Australia,* [3]*Bruce Lefroy Centre for Genetic Health Research, Murdoch Children's Research Institute, Parkville, Victoria, Australia, and* [4]*Epilepsy Research Centre, Department of Medicine, University of Melbourne, Austin Health, Heidelberg, Victoria, Australia*

References

1. Tan D, Wei C, Chen Z, et al. CAG repeat expansion in THAP11 is associated with a novel spinocerebellar ataxia. Mov Disord 2023;38:1282–1293. https://doi.org/10.1002/mds.29412

2. Halldorsson BV, Eggertsson HP, Moore KHS, et al. The sequences of 150,119 genomes in the UK biobank. Nature 2022;607:732–740. https://doi.org/10.1038/s41586-022-04965-x

3. Ibañez K, Polke J, Hagelstrom RT, et al. Whole genome sequencing for the diagnosis of neurological repeat expansion disorders in the UK: a retrospective diagnostic accuracy and prospective clinical validation study. Lancet Neurol 2022;21(3):234–245. https://doi.org/10.1016/S1474-4422(21)00462-2

4. Dolzhenko E, Deshpande V, Schlesinger F, et al. ExpansionHunter: a sequence-graph-based tool to analyze variation in short tandem repeat regions. Bioinformatics 2019;35(22):4754–4756. https://doi.org/10.1093/bioinformatics/btz431

5. Qiu, Y., Deshpande, V., Avdeyev, P. et al. Illumina Repeat Catalogs https://github.com/Illumina/RepeatCatalogs.

6. Craig K, Takiyama Y, Soong BW, et al. Pathogenic expansions of the SCA6 locus are associated with a common CACNA1A haplotype across the globe: founder effect or predisposing chromosome? Eur J Hum Genet 2008;16(7):841–847. https://doi.org/10.1038/ejhg.2008.20

7. Pandey N, Mittal U, Srivastava AK, et al. SMARCA2 and THAP11: potential candidates for polyglutamine disorders as evidenced from polymorphism and protein-folding simulation studies. J Hum Genet 2004;49:596–602. https://doi.org/10.1007/s10038-004-0194-8

Reply to: "*THAP11* CAG Expansion Beyond Chinese-Ancestry Cohorts: An Examination of 1000 Genomes and UK Biobank"

Fearnley et al reported an ataxia case with a 46 of 29 CAG repeats in *THAP11*, comorbid with a 22-CAG repeat expansion in *CACNA1A* from European-ancestry ataxia cohort. This patient was diagnosed as hereditary ataxia at the age of 40 and Parkinson's disease at 53. The comorbidity of two polyglutamine (polyQ) diseases may complicate the phenotype and further assessment is warranted. Nevertheless, this is the first time that CAG expansion greater than 45 repeats in *THAP11* was found beyond Chinese ataxia population. Therefore, this patient further confirmed this new subtype of spinocerebellar ataxia (SCA). In addition, Fearnley et al also investigated the normal expansion ranging from 19 to 39 repeats in *THAP11* from 2504 individuals in 1000 Genomes. In general, these findings suggest that ATX-*THAP11* is not limited to Asian. As the latest SCA included in Online Mendelian Inheritance in Man (OMIM) is SCA50, the new SCA subtype caused by the CAG expansion in *THAP11* identified by us is appropriately named ATX-*THAP11*/SCA51.

The pathogenic mechanism of polyQ diseases is widely attributed to the toxicity of abnormal proteins containing stretches of polyQ residues.[1] We observed an increase in the length of glutamine correlates with a more severe phenotype in our ATX-*THAP11*/SCA51 family.[2] Interestingly, expansion of CAG repeats interrupted by CAA is common in polyQ disease, which encodes the same glutamine for polyQ stretch, but may lead to distinct RNA secondary structures.[1] In recent years, several reports have explored the impact of various DNA structures within CAG repeats on the severity of polyQ diseases. Three independent studies have reported that in individuals with Huntington's disease (HD), a reduction in CAA insertions and an increase in uninterrupted CAG repeats are associated with an earlier onset age of HD.[3,4] Conversely, increased CAA insertions are linked to delayed disease onset.[3] Further comparisons among different HD mouse models have demonstrated that longer uninterrupted CAG repeats are associated with striatum-selective transcriptionopathy, possibly because of CAG repeat instability and nuclear aggregation.[5] Another intriguing observation is that pure CAG repeats in *ATXN2* are associated with SCA2, whereas CAG repeats in *ATXN2* causing amyotrophic lateral sclerosis (ALS) and Parkinson's disease are predominantly interrupted by CAA.[6] In Drosophila models, ataxin-2 encoded by pure CAG repeats induces retinal and neurotoxicity, whereas CAA insertions or only CAA repeats do not induce toxicity, despite producing the same protein.[7] In our large family affected by ATX-*THAP11*/SCA51, we noted a decrease in CAA insertions and an increase in pure CAG length. However, there was no decrease in CAA insertions in our other family or in the patient reported by Fearnley et al. Currently, because of the limited number of ATX-*THAP11*/SCA51 patients, whether CAA interruption modulates pathogenesis of polyQ aggregation and influences clinical phenotypes needs further exploration.

In the future, it is necessary to detect CAG repeat expansion of *THAP11* across ethnically and geographically different ataxia populations to clarify its genotype–phenotype characteristics and pathogenesis of ATX-*THAP11*/SCA51.

Acknowledgments: The author declares that the research was conducted in the absence of other commercial or financial relationships that could be construed as a potential conflict of interest.

Data Availability Statement

Not applicable.

Cuijie Wei, MD,[1] Zhao Chen, MD,[2] Dandan Tan, PhD,[1,3] Hong Jiang, MD,[2,4,5,6*] Nanbert Zhong, PhD,[7*] and Hui Xiong, MD, PhD[1,8*]

Key Words: spinocerebellar ataxia, *THAP11* gene, polyglutamine

***Correspondence to:** Dr. Hui Xiong, Department of Pediatrics, Peking University First Hospital, 1 Xi anmen Street, West District, Beijing, 100034, China; E-mail: xh_bjbj@163.com; Dr. Hong Jiang, Department of Neurology, Xiangya Hospital, Central South University, No. 87 Xiangya Road, Changsha, 410008, China; E-mail: jianghong73868@126.com; Dr. Nanbert Zhong, New York State Institute for Basic Research in Developmental Disabilities, 1050 Forest Hill Road, Staten Island, NY 10314, USA; E-mail: nanbert.zhong@opwdd.ny.gov

Dandan Tan, Cuijie Wei, and Zhao Chen contributed equally to this work.

Hong Jiang, Nanbert Zhong, and Hui Xiong contributed equally to this work.

Relevant conflicts of interest/financial disclosures: Nothing to report.

Funding agencies: The study was supported by grants from the National High Level Hospital Clinical Research Funding (High Quality Clinical Research Project of Peking University First Hospital, 2022CR69), National Key R&D Program of China (2021YFA0805200), National Natural Science Foundation of China (82171393), Natural Science Foundation of Beijing Municipality (7212116), National Key Research and Development Program of China (2016YFC0901505), the Innovation Research Group Project of Natural Science Foundation of Hunan Province (2020JJ1008), the Science and Technology Innovation Group of Hunan Province (2020RC4043), and Natural Science Foundation of Hunan Province (2022JJ20094).

Received: 19 September 2023; **Accepted:** 4 October 2023

Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.29639