

BRIEF COMMUNICATION **OPEN**


Unexpected diagnosis of myotonic dystrophy type 2 repeat expansion by genome sequencing

Haloom Rafehi^{1,2}, Cherie Green³, Kiymet Bozaoglu^{4,5}, Greta Gillies⁴, Martin B. Delatycki^{4,5,6,7}, Paul J. Lockhart^{4,6}, Ingrid E. Scheffer^{1,5,6,8,9} and Melanie Bahlo^{1,2}✉

© The Author(s) 2022

Several neurological disorders, such as myotonic dystrophy are caused by expansions of short tandem repeats (STRs) which can be difficult to detect by molecular tools. Methodological advances have made repeat expansion (RE) detection with whole genome sequencing (WGS) feasible. We recruited a multi-generational family (family A) ascertained for genetic studies of autism spectrum disorder. WGS was performed on seven children from four nuclear families from family A and analyzed for REs of STRs known to cause neurological disorders. We detected an expansion of a heterozygous intronic CCTG STR in *CNBP* in two siblings. This STR causes myotonic dystrophy type 2 (DM2). The expansion did not segregate with the ASD phenotype. Repeat-primed PCR showed that the DM2 CCTG motif was expanded above the pathogenic threshold in both children and their mother. On subsequent examination, the mother had mild features of DM2. We show that screening of STRs in WGS datasets has diagnostic utility, both in the clinical and research domain, with potential management and genetic counseling implications.

European Journal of Human Genetics (2023) 31:122–124; <https://doi.org/10.1038/s41431-022-01166-y>

INTRODUCTION

The advent of Whole Exome and Whole Genome Sequencing (WES/WGS) has been instrumental in the rapid genetic diagnosis of many conditions, including neurological disorders in both clinical and research settings. More recently, there have been significant advances in the development of bioinformatic methods for the detection of repeat expansions (REs) of short tandem repeats (STRs) in short-read sequencing datasets, such as exSTRA and ExpansionHunter [1]. These tools have proven successful for the diagnosis of conditions caused by known REs, and in the discovery of new REs [2, 3]. In this study, we recruited a multi-generational family (Family A) with multiple incidences of autism spectrum disorder (ASD), in order to study generic determinates of ASD, including REs. Fragile X syndrome is caused by a RE in *FMR1* has a 50% incidence of ASD co-morbidity [4]. ASD has also been reported as a symptom of myotonic dystrophy type 1 (DM1), a rare genetic disorder caused by a RE in *DMPK*. Here we show that applying a standard RE screening pipeline to Family A resulted in an unexpected diagnosis of myotonic dystrophy type 2 (DM2), caused by a CCTG repeat expansion in *CNBP* [5], in a sub-branch of the family.

METHODS

We recruited an extended multi-generational family (Family A) ascertained for genetic studies of ASD (Royal Children's Hospital ethics approval

#25043, with written informed consent). Genomic DNA was isolated from saliva. Paired-end 150 bp WGS was performed on seven children, siblings or first-degree cousins, from four nuclear families from within Family A using TruSeq Nano Library Preparation Kit and the Illumina HiSeq X Ten platform. Alignment was performed based on the GATK best practice pipeline. Fastq files aligned to the hg19 reference genome using BWA-mem, then duplicate marking, local realignment, and recalibration were performed with GATK and analyzed for REs of STRs known to cause neurological disorders using RE detection tools exSTRA [6] and ExpansionHunter [7], as previously described [2]. A database of pre-defined pathogenic REs was used: https://github.com/bahlolab/exSTRA/blob/master/inst/extdata/repeat_expansion_disorders_hg19.txt (file version committed on Nov 21, 2019).

These tools utilize paired-end reads to detect expanded STRs: exSTRA uses an empirical Cumulative Distribution Function (ECDF) to determine outliers, while ExpansionHunter estimates the number of repeats in the STR on each allele. Repeat-primed PCR [8] was used to test for expansion of the RE and was performed by PathWest Laboratory Medicine WA. Twenty to 200 ng of DNA were amplified in a reaction volume of 30 µl, using 1.5 units of GoTaq Polymerase (Promega), supplied buffer 1×, MgCl₂ 1.25 mmol/L, DMSO 10%, 7-DeAZA-GTP (Roche) 0.2 mmol/L, primers P1 (Fam-5'-GCCTAGGGGACAAAGTGA-3') 3 pmol, P4 (5'-TACGCATCC-GAGTTTGAGACGCTGCCTGCCTGCCTGCCTG-3') 0.6 pmol, P3 (5'-TACGCATCCAGTTTGAGACG-3') 3 pmol, with the following cycling conditions: initial denaturation 5 min/94 °C, followed by 35 cycles of (30 s/94 °C; 30 s/56 °C; 2 min/72 °C) and final extension of 10 min/72 °C, then stored at +4 °C until analyzed. In the prospective study, every analysis included at least one DM2-positive and one DM2-negative sample as controls of PCR reaction.

¹The Walter and Eliza Hall Institute of Medical Research, Melbourne, VIC, Australia. ²Department of Medical Biology, University of Melbourne, Melbourne, VIC, Australia.

³Department of Psychology and Counselling, School of Psychology and Public Health, La Trobe University, Melbourne, VIC, Australia. ⁴Bruce Lefroy Centre for Genetic Health Research, Murdoch Children's Research Institute, Melbourne, VIC, Australia. ⁵Department of Medicine, The University of Melbourne, Austin Health, Melbourne, VIC, Australia.

⁶Department of Paediatrics, The University of Melbourne, Melbourne, VIC, Australia. ⁷Victorian Clinical Genetics Services, Melbourne, VIC, Australia. ⁸Florey Institute, Melbourne, VIC, Australia. ⁹Murdoch Children's Research Institute, Melbourne, VIC, Australia. ✉email: bahlo@wehi.edu.au

Received: 28 October 2021 Revised: 29 May 2022 Accepted: 21 July 2022

Published online: 9 August 2022

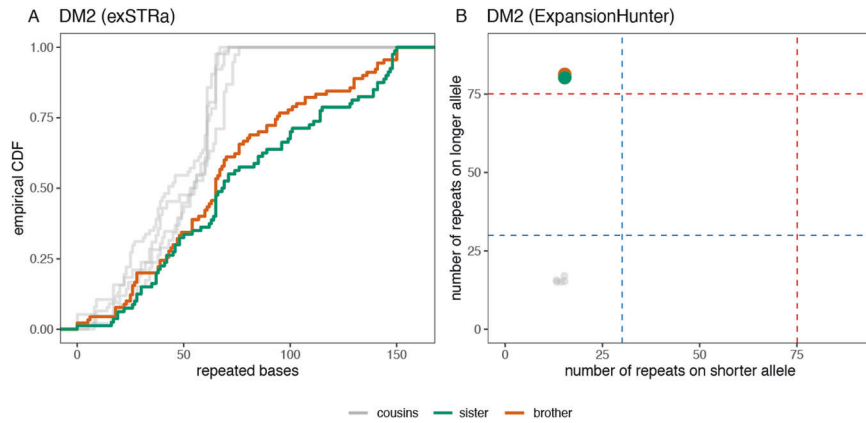


Fig. 1 Detection of a DM2 repeat expansion in two siblings using WGS. The DM2 repeat expansion was detected in two siblings (green and orange) using the tools (a) exSTRa (shown as ECDF: Empirical Cumulative Distribution Function) and (b) ExpansionHunter (red dotted line indicates pathogenic threshold of 75 repeats, blue dotted line: benign/stable STR threshold of 30 repeats), compared with their first cousins who do not carry the allele, shown in gray. The ECDF plots the number of base pairs with in each read at the locus that are comprised of the repeated motif, sorted by size. A shift to the right on the plot in a sample compared to other samples indicates a potential RE.

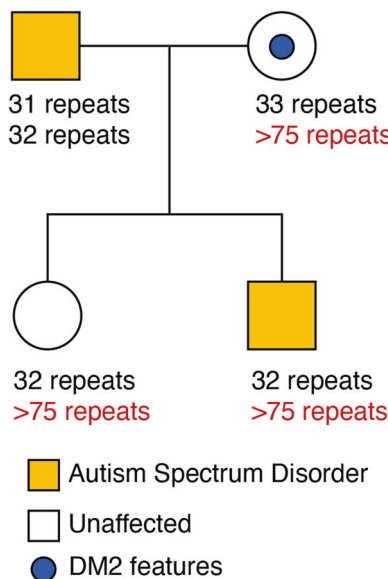


Fig. 2 Pedigree showing affected status for ASD (yellow). The estimated number of base pairs in the CCTG STR in *CNBP* on each allele is shown for each family member, noting that the father does not carry an expanded allele. The expanded alleles are highlighted in red. 300 base pairs corresponds to 75 repeat units, which is the pathogenic threshold for DM2.

RESULTS

We detected an expansion of a heterozygous intronic CCTG STR in the gene *CNBP*, which is known to cause DM2, in a brother and sister pair (Fig. 1). Results from exSTRa (Fig. 1a) identify excess reads containing the CCTG motif in the two siblings compared to their cousins, indicating an expanded allele. ExpansionHunter (Fig. 1b) was used to estimate the allele repeat to be approximately 80 repeats in size in both siblings, which is larger than the 75 repeat threshold for pathogenicity for DM2. Although this estimated RE size is small compared to the very large expansions of up to 10,000 repeats that have been reported in DM2, it is well established that *in silico* tools do not provide accurate estimates for larger expansions [9]. Repeat-primed PCR was performed and confirmed that the DM2 CCTG motif was expanded above the pathogenic threshold (>75 repeats) in both children. Their parent was also found to have a CCTG expansion

(>75 repeats) in *CNBP* (Fig. 2). On follow-up examination, the affected parent had mild features of DM2, including the classical finding of percussion myotonia [10]. This DM2 RE was not detected in any of their cousins in the broader family, regardless of ASD status, and notably did not segregate with their father who had ASD.

DISCUSSION

Here, through the implementation of an advanced bioinformatic pipeline, we have made an unexpected incidental diagnosis of DM2 in three related individuals. Our pipeline incorporated enhanced RE detection methods to increase the diagnostic utility of WGS in neurological disorders. DM2 is typically a late-onset condition. In contrast to myotonic dystrophy type 1 (DM1), it has a relatively mild phenotype and anticipation is rarely observed. The subtle features identified on reviewing the family, evident only in the mildly affected mother, suggest that the diagnosis of DM2 can be missed due to its relatively mild phenotype. Unlike DM1 [11], DM2 has not been associated with ASD, suggesting that DM2 is an incidental finding in this family. This is further supported by the clinical genetic pattern of inheritance of ASD which extends through the paternal, rather than the maternal, line.

The missed diagnosis of DM2 in this family, despite clinical symptoms in the affected mother, highlights that repeat expansions disorders remain underdiagnosed. A molecular diagnosis is important for optimal management of the symptoms of DM2 and for genetic counseling [4]. Critically, we identified this repeat expansion in a family that was recruited for ASD research, highlighting the utility of this pipeline for the diagnosis of incidental neurological disorders in cohorts not thought to have REs.

Although this RE represents a likely incidental finding, we originally pursued the discovery due to the potential association of DM1 and ASD [11], speculating that the even less common and more subtle features of DM2 might be a hidden contributor to ASD likelihood. However the DM2 RE did not segregate with ASD in this family. Furthermore, a recent screen of children with ASD ($N = 1812$) identified DM1 expansion in seven individuals with ASD ($OR = 1.37$), no expansions were identified in DM2, providing additional evidence that DM2 is not associated with ASD [12].

The use of bioinformatic tools to identify REs in WGS datasets has not been broadly applied in research and clinical settings. Indeed, recent publications still incorrectly state that WGS cannot be used to detect RE [13], despite multiple demonstrations of their utility across a wide range of RE disorders. We show that the

implementation of an STR analysis pipeline to screen all WGS datasets has considerable diagnostic utility, both in the clinical and research domain.

DATA AVAILABILITY

The datasets generated during and/or analyzed during the current study are not publicly available due to specific ethics that does not allow public sharing of genomic data, but are available from the corresponding author on reasonable request.

REFERENCES

- Bahlo M, Bennett MF, Degorski P, Tankard RM, Delatycki MB, Lockhart PJ. Recent advances in the detection of repeat expansions with short-read next-generation sequencing. *F1000Res*. 2018;7:F1000, Faculty Rev-736.
- Rafehi H, Szmulewicz DJ, Pope K, Wallis M, Christodoulou J, White SM, et al. Rapid Diagnosis of Spinocerebellar Ataxia 36 in a three-Generation Family Using Short-Read Whole-Genome Sequencing Data. *Mov Disord*. 2020;35:1675–9.
- Rafehi H, Szmulewicz DJ, Bennett MF, Sobreira NLM, Pope K, Smith KR, et al. Bioinformatics-based identification of expanded repeats: a non-reference intronic pentamer expansion in RFC1 causes CANVAS. *Am J Hum Genet*. 2019;105:151–65.
- Abbeduto L, McDuffie A, Thurman AJ. The fragile X syndrome-autism comorbidity: what do we really know? *Front Genet*. 2014;5:355.
- Liquori CL, Ricker K, Moseley ML, Jacobsen JF, Kress W, Naylor SL, et al. Myotonic dystrophy type 2 caused by a CCG expansion in intron 1 of ZNF9. *Science*. 2001;293:864–7.
- Tankard RM, Bennett MF, Degorski P, Delatycki MB, Lockhart PJ, Bahlo M. Detecting expansions of tandem repeats in cohorts sequenced with short-read sequencing data. *Am J Hum Genet*. 2018;103:858–73.
- Dolzhenko E, Deshpande V, Schlesinger F, Krusche P, Petrovski R, Chen S, et al. ExpansionHunter: a sequence-graph-based tool to analyze variation in short tandem repeat regions. *Bioinformatics*. 2019;35:4754–6.
- Warner JP, Barron LH, Goudie D, Kelly K, Dow D, Fitzpatrick DR, et al. A general method for the detection of large CAG repeat expansions by fluorescent PCR. *J Med Genet*. 1996;33:1022–6.
- Day JW, Ricker K, Jacobsen JF, Rasmussen LJ, Dick KA, Kress W, et al. Myotonic dystrophy type 2: molecular, diagnostic and clinical spectrum. *Neurology*. 2003;60:657–64.
- Johnson NE, Heatwole CR. Teaching video neuroimages: trapezius myotonia percussion sign in myotonic dystrophy type 2. *Neurology*. 2013;80:e251.
- Angeard N, Huerta E, Jacqueline A, Cohen D, Xavier J, Gargiulo M, et al. Childhood-onset form of myotonic dystrophy type 1 and autism spectrum disorder: Is there comorbidity? *Neuromuscul Disord*. 2018;28:216–21.
- Trost B, Engchuan W, Nguyen CM, Thiruvahindrapuram B, Dolzhenko E, Backstrom I, et al. Genome-wide detection of tandem DNA repeats that are expanded in autism. *Nature*. 2020.
- Boesch SM, Nance MA. Intronic pentanucleotide expansion in the replication factor 1 gene (RFC1) is a major cause of adult-onset ataxia. *Neurol Genet*. 2020;6:e436.

ACKNOWLEDGEMENTS

We thank the patients and their families for participating in our research.

AUTHOR CONTRIBUTIONS

CG, GG, MBD and IES contributed to patient recruitment and phenotyping. KB performed molecular work. All authors contributed to the study design. HR performed bioinformatics analysis. HR wrote the manuscript with support from MB and PL. All authors contributed to editing the manuscript.

FUNDING

This work was funded in part by National Health and Medical Research Council of Australia Project Grants (GNT1044175 and GNT1098255). HR was supported by an NHMRC EL1 Investigator Grant (GNT1194364). KB was supported by an EH Flack Fellowship. IES was supported by a NHMRC Senior Practitioner Fellowship (GNT1006110; GNT1104831). PJL was supported by a NHMRC Fellowship (GNT1032364) and the Vincent Chiodo Foundation. MB was supported by an NHMRC Senior Research Fellowship (GNT1102971). This work was made possible through the Victorian State Government Operational Infrastructure Support and Australian Government NHMRC IRIISS.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICAL APPROVAL

Royal Children's Hospital ethics approval #25043, with written informed consent

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Melanie Bahlo.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022