



OPEN Multi-omics analysis detail a submicroscopic inv(15)(q14q15) generating fusion transcripts and *MEIS2* and *NUSAP1* haploinsufficiency

Marlene Ek^{1,2}✉, Malin Kvarnung^{1,2}, Maria Pettersson^{1,2}, Maria Johansson Soller^{1,2}, Britt-Marie Anderlid^{1,2}, Håkan Thonberg^{1,2}, Jesper Eisfeldt^{1,2,3} & Anna Lindstrand^{1,2}

Inversions are balanced structural variants that often remain undetected in genetic diagnostics. We present a female proband with a *de novo* Chromosome 15 paracentric inversion, disrupting *MEIS2* and *NUSAP1*. The inversion was detected by short-read genome sequencing and confirmed with adaptive long-read sequencing. The breakpoint junction analysis revealed a 96 bp (bp) deletion and an 18 bp insertion in the two junctions, suggesting that the rearrangement arose through a replicative error. Transcriptome sequencing of cultured fibroblasts revealed normal *MEIS2* levels and 0.61-fold decreased expression of *NUSAP1*. Furthermore, three fusion transcripts were detected and confirmed by Sanger sequencing. Heterozygous loss of *MEIS2* (MIM# 600987) is associated with a cleft palate, heart malformations, and intellectual impairment, which overlap with the clinical symptoms observed in the proband. The observed fusion transcripts are likely non-functional, and *MEIS2* haploinsufficiency is the likely disease causative mechanism. Altogether, this study's findings illustrate the importance of including inversions in rare disease diagnostic testing and highlight the value of long read sequencing for the validation and characterization of such variants.

Keywords Inversion, Short-read whole genome sequencing, Adaptive long-read sequencing, Transcriptome sequencing, *MEIS2*, Gene disruption, Gene fusion

Pathogenic variants in *MEIS2* are associated with a syndrome known as “cleft palate, cardiac defects, and impaired intellectual development” (CPCMR, MIM# 600987). *MEIS2*, a homeodomain-containing protein, is part of the three amino acid loop extension (TALE) family, and functions as a transcription regulator. *Meis2* has been implicated in the development of limbs and brain in chicks, lens and retina formation in mice and *Mekada* fish, and is essential for both cardiac and neural crest development in mice^{1–5}. The gene is highly intolerant to loss-of-function mutations, with a pLI score greater than 0.9, and causes disease primarily through haploinsufficiency. Most of the described *MEIS2* variants are *de novo*, although cases of inherited variants have been described. Carrier parents tend to exhibit a similar phenotype, with mosaic carriers showing milder expression^{6–10}.

The previously reported disease-causing variants in *MEIS2* include both structural variants and sequence variants, such as single nucleotide variants (SNVs) and insertions/deletions (INDELs). The structural variants are predominantly copy number variants (CNVs), with deletions, ranging in size from 123 kb to 6.97 Mb, and one family reported with a 58 kb tandem duplication. One of the deletions was part of a *de novo* complex chromosomal rearrangement^{6–8,11–18}. Notably, one case involving a balanced structural variant has been reported: a balanced reciprocal translocation between chromosomes 11 and 15, t(11;15)(p14;q14), where the breakpoint on chromosome 15 disrupts *MEIS2* in intron 6¹⁹. The reported SNV/INDELs include truncating, splice site, missense variants or in-frame deletions mainly within the homeodomain (Fig. 1)^{9,10,18,20–27}.

We detail the genetic aberrations in a proband with cleft palate and high-functioning autism. Structural variant analysis of short-read genome sequencing revealed a 4.37 Mb inversion on chromosome 15 (Fig. 1), with breakpoints in intron 8 of *MEIS2* and intron 7 of *NUSAP1*, generating three fusion transcripts.

¹Department of Molecular Medicine and Surgery, Karolinska Institutet, 171 76 Stockholm, Sweden. ²Department of Clinical Genetics and Genomics, Karolinska University Hospital, 171 76 Stockholm, Sweden. ³Science for Life Laboratory, Karolinska Institutet Science Park, 171 65 Solna, Sweden. ✉email: marlene.ek.2@ki.se

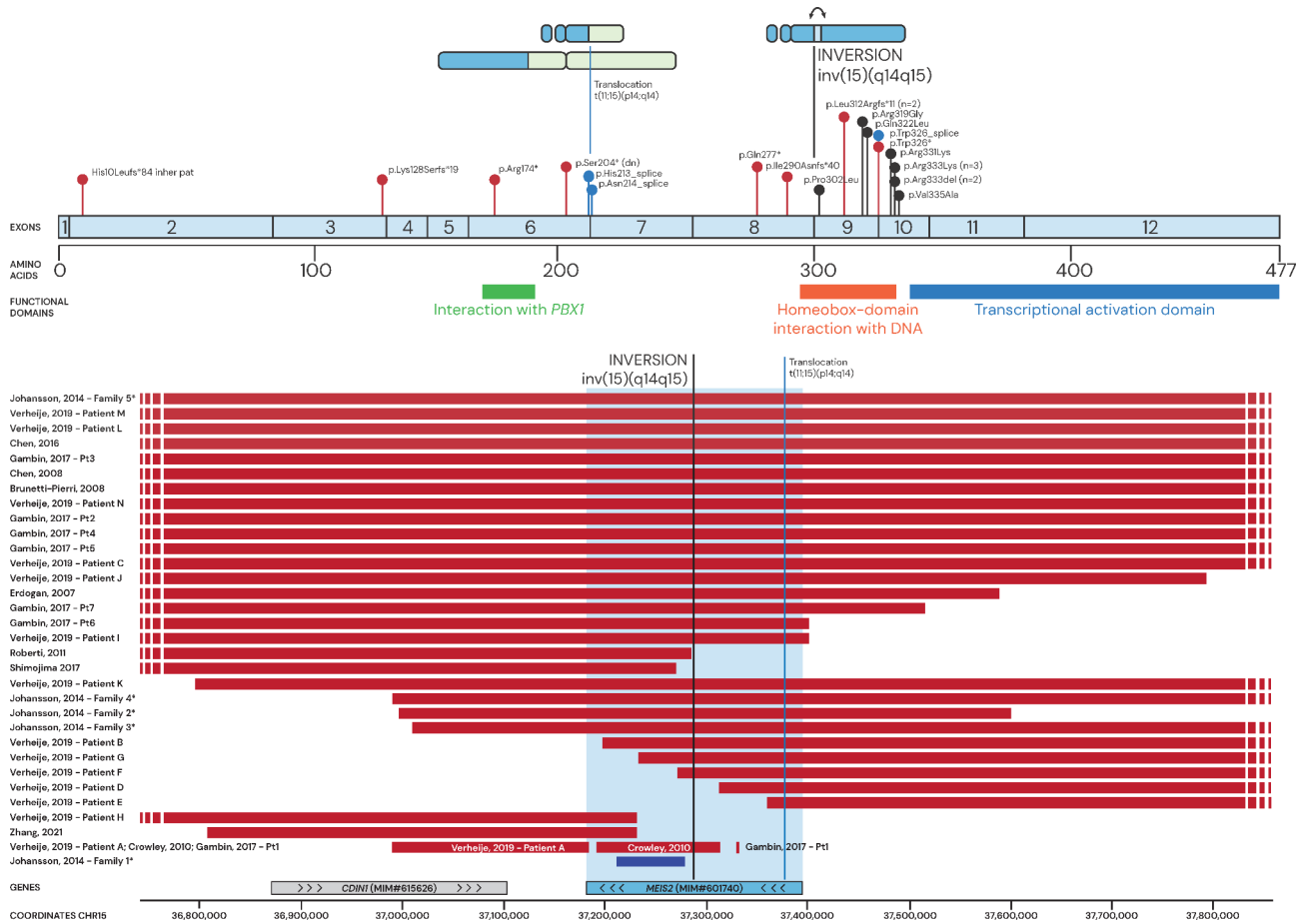


Fig. 1. Genetic alterations in *MEIS2* in previously reported cases and our study proband. (Top) Sequence variants identified across the *MEIS2* gene. Nonsense (red), missense (black) and splice variants (blue), as well as previously reported translocation breakpoint (blue line) and our proband’s inversion breakpoint in black. (Bottom) Copy number variants (CNVs) reported in *MEIS2*, showing deletions (red), duplications (blue) translocation’s breakpoint (blue line) and our proband’s inversion breakpoint (black line). * Approximated breakpoints.

Results

The proband is the first child of healthy, non-consanguineous parents. She was born after an uneventful pregnancy at 39 weeks + 6 days (birth weight 3670 g, birth length 53 cm and head circumference 35 cm). A posterior cleft palate was diagnosed early, and she experienced feeding difficulties during the first months of life. The cleft palate was surgically corrected twice, at 5 and 12 months of age. Her motor development was delayed; she sat unsupported at 9 months and walked independently at 19 months of age. She exhibited hypotonia and hypermobile joints, and her postnatal growth was slightly restricted, with a head circumference of -2 SD and a length of -0.5 SD at 3 years of age. Adhesions between labia majora initially raised suspicion of urethral duplication, but an ultrasound of the urinary tracts was normal. A cardiac ultrasound revealed a small atrial septal defect, which was clinically insignificant. A neuropsychological evaluation at the age of 5.5 years diagnosed her with high-function autism and her intellectual level was within the normal range. Some facial features were present, including penciled eyebrows, rounded eyes, large and protruding ears, and discrete fetal finger pads. She has two younger, healthy siblings. Genetic tests included array comparative genomic hybridization (aCGH) at 18 months of age, short-read genome sequencing analysis with in-silico gene panels for neuromuscular disorders (499 genes) and intellectual disability (885 genes) at age 7, all with normal results.

Genome-wide structural variant analysis of the short-read genome sequencing data identified a 4.37 Mb inversion on chromosome 15. The first breakpoint (BP1), was located in intron 8 of the MANE transcript NM_170675.5 of *MEIS2*, and the second breakpoint (BP2) was found in intron 7 of *NUSAP1* (NM_016359.5), resulting in a disruption of the homeodomain of *MEIS2*. Follow-up analysis of the parents confirmed that the variant had arisen *de novo*. The inversion was verified with adaptive long-read sequencing (Oxford Nanopore Technologies, ONT), with 34 reads supporting the breakpoints. Breakpoint junction analysis revealed a one nucleotide microhomology and a 96 bp deletion (GRCh37, chr15:g.37,288,170–37,288,265) in BP1. In BP2 (GRCh37, chr15:g.41,663,130), located within a MER104, there was an 18 bp insertion, of which 11 bp were templated from the adjacent sequence of the inverted segment (Fig. 2).

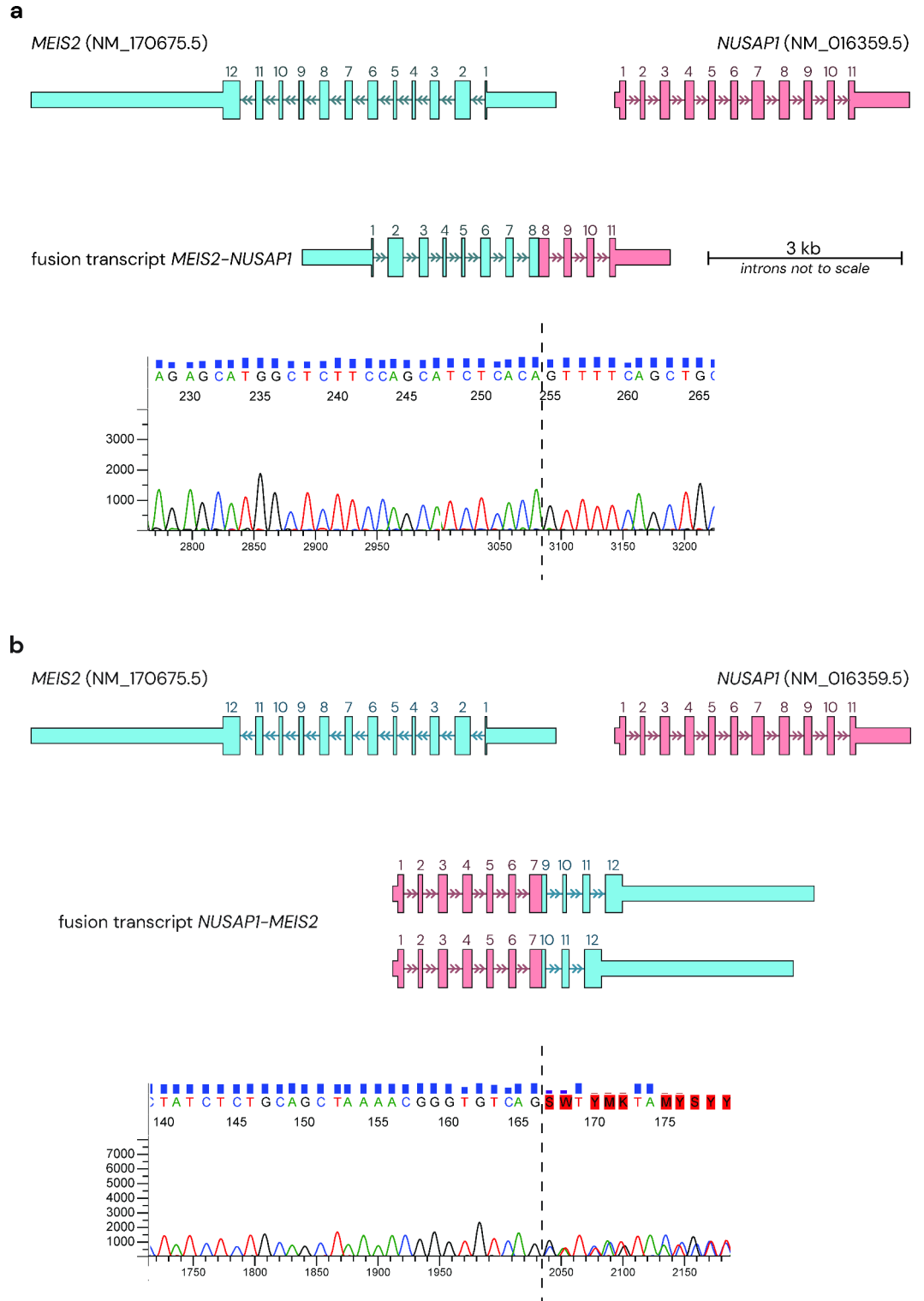


Fig. 3. Depiction of fusion transcripts identified in our proband. *MEIS2* transcript (NM_170675.5) depicted in blue, *NUSAP1* transcript (NM_016359.5) in red. (a) Fusion transcript of *MEIS2-NUSAP1* with Sanger sequencing across the fusion junction (dashed line) of exon 8 of *MEIS2* and exon 8 of *NUSAP1*. (b) Fusion transcripts of *NUSAP1-MEIS2* with Sanger sequencing across the fusion junction (dashed line) of exon 7 of *NUSAP1* and exon 9 of *MEIS2*, as well as exon 7 of *NUSAP1* and exon 10 of *MEIS2*.

Conversely, not all individuals with large CNVs and missense variants present with a severe phenotype, suggesting a more complex background contributing to disease expression (Supplementary Table S1). In the proband presented here, there is a clear clinical overlap with CPCMR, including cleft palate, hypotonia, delayed motor development, autism and facial features consistent with previously reported cases. The presence of a likely non-functional *MEIS2-NUSAPI* fusion transcript, along with the normal functional transcript, could be misinterpreted as two normal transcripts, potentially masking the effects of a functional haploinsufficiency. The milder phenotype in our proband, compared to previously reported *MEIS2* cases, may suggest some remaining function in the fusion gene product.

However, a recent publication reported an *NUSAPI* nonsense variant in two individuals with microcephaly, severe developmental delay, brain abnormalities, and seizures. A truncated transcript was detected that was hypothesized to evade nonsense mediated decay and have a toxic effect²⁹. The phenotype in our proband did not overlap the phenotype described, but we cannot rule out a contribution from *NUSAPI*. The identified *NUSAPI-MEIS2* fusion transcripts could have a similar toxic effect as the truncated *NUSAPI* transcript described above²⁹.

For the proband presented here, the diagnostic odyssey spanned over nine years, starting with aCGH and sequencing of *KMT2D* at 18 months of age. At the age of 8 years, our proband was analyzed with short-read genome sequencing without any pathogenic findings. Four years later, the genome analysis was reevaluated with a new pipeline that included structural variant detection, identifying the inversion on Chromosome 15. Currently, with genome sequencing now offered as a first-tier test to children with neurological diseases, this timeline would have been significantly shortened.

In conclusion, by applying short-read genome sequencing, adaptive long-read and transcriptome analysis a cryptic Chromosome 15 inversion was investigated. The disease-causing rearrangement, undetected by all standard clinical diagnostic tests, showcases how balanced structural variants need to be included in genetic diagnostics of syndromes and neurodevelopmental disorders. Further studies are needed to determine the clinical impact of fusion transcripts in rare diseases.

Materials and methods

Ethics declaration

The study was conducted in accordance with the Declaration of Helsinki and approved by the Regional Ethical Review Board in Stockholm, Sweden (protocol number 2019-04746).

Written informed consent was provided by the participant and her parents to publish this paper.

DNA and RNA extraction

Genomic DNA was extracted from whole blood using the QIAasympphony instrument (QIAGEN, Hilden, Germany) with the QIAasympphony DSP DNA Midi Kit (cat. no. 937255, QIAGEN, Hilden, Germany), following the manufacturer's standard protocol.

Fibroblasts were cultured from a skin biopsy in a medium composed of RPMI 1640 (1x) and Ham's F-10 Nutrient Mixture at a 1:1 ratio, supplemented with 10% fetal bovine serum, 1% L-glutamine, and 0.2% Penicillin-Streptomycin solution (5000 U/mL). The cells were harvested by mechanical scraping from the culture flask and transferred to a PAXgene tube, and RNA was extracted using the PAXgene Blood RNA Kit (Prenalytix, cat. no. 762174, QIAGEN, Hilden, Germany) on a QIAcube Connect MDx system, following the standard protocol.

The RNA was reverse transcribed to cDNA using SuperScript VILO Master Mix (QIAGEN, Hilden, Germany). The RNA template and nuclease-free water were added to the master mix and incubated in a thermal cycler at 25 °C for 10 min, 42 °C for 60 min, and 85 °C for 5 min, followed by a hold at 4 °C. Nuclease-free water was then added, and the cDNA was stored at – 20 °C.

Genetic analysis

Our clinical short-read genome sequencing workflow has been previously described^{30–32}. To verify the inversion call, we performed adaptive long-read sequencing using the PromethION platform (ONT). The target region, spanning 32.5 Mb ([GRCh38] chr15:22,874,354 – 55,370,932) encompassed the entire inversion along with a 14 Mb buffer zone upstream and downstream of the region. Libraries were prepared from 2.76 µg of genomic DNA, following the ONT protocol 'Ligation Sequencing gDNA (SQK-LSK114), with an average sample fragment size of 50,031 bp. The fragment size was estimated using Femto Pulse, following the protocol for the 'Genomic DNA 165 kb Kit'. A single PromethION R10.4 (FLO-PRO114M) flow cell was used for sequencing. The base calling was performed using the Dorado basecaller (<https://github.com/nanoporetech/dorado>), which was run in High Accuracy Mode (HAC). The resulting data was processed using PoorPipe (<https://github.com/J35P312/poorpipe>), which performs alignment using Minimap2³³, and SV calling using Sniffles 1³⁴. The inversion was manually inspected in the Integrative Genomics Viewer (IGV) using the hg19/GRCh37 reference genome. Repetitive elements were analyzed using RepeatMasker in the UCSC Genome Browser.

The effects on RNA were evaluated by whole-transcriptome sequencing of RNA isolated from cultured fibroblasts. Briefly, RNA was quantified and processed using a stranded, poly(A)-tailed kit (Illumina) before being subjected to 150 bp paired-end sequencing with approximately 150 million reads generated per sample on the Nova Seq X platform. The data was processed using the genomic medicine Sweden transcriptome pipeline Tomte (<https://github.com/genomic-medicine-sweden/tomte>). Briefly, the data was aligned to GRCh37 using STAR³⁵, next aberrant expression events were detected by Detection of RNA Outlier Pipeline (DROP)³⁶ using the default, recommended settings for OUTRIDER³⁷, and fusion transcripts were detected using STAR-Fusion³⁸.

PCR and Sanger sequencing of fusion transcripts

Primers targeting the inversion breakpoints in *MEIS2*(exons 7 and 11) and *NUSAPI*(exons 7 and 8) were designed (Supplementary Table S2). Breakpoint PCR was performed using AmpliTaq Gold (Fisher Scientific,

Waltham, MA, USA) with a master mix containing PCR buffer II (1x), MgCl₂ (2 mM), dNTPs (100 μM), and AmpliTaq Gold (1 U). The PCR conditions included 10 min at 96 °C, followed by 35 cycles of 96 °C (30 s), 62 °C (30 s), and 72 °C (2 min), with a final extension at 72 °C for 10 min.

Amplified products were detected using the FlashGel system (Lonza) with a 100 bp–3 kb DNA marker (cat. no. 57034, Lonza) and imaged on the GenoPlex system (VWR). Sanger sequencing was performed on the normal alleles of *MEIS2* and *NUSAP1*, and the fusion transcripts *MEIS2-NUSAP1* and *NUSAP1-MEIS2* (Supplementary Table S2).

The PCR products were purified using Illustra ExoProStar, and the sequencing reaction was conducted with BigDye Terminator v3.1 (Applied Biosystems). Sequencing was done on an ABI 3500xL Genetic Analyzer (Applied Biosystems, Waltham, MA, USA).

Data availability

The BAM files from the adaptive long-read sequencing are available at the European Genome-phenome Archive (EGA) web portal (<https://ega-archive.org/datasets/EGAS50000000632>) under project ID EGAD50000000632.

Received: 18 September 2024; Accepted: 27 November 2024

Published online: 05 December 2024

References

1. Capdevila, J., Tsukui, T., Rodríguez Esteban, C., Zappavigna, V. & Izpisua Belmonte, J. C. Control of vertebrate limb outgrowth by the proximal factor *Meis2* and distal antagonism of BMPs by *Gremlin*. *Mol. Cell* **4**(5), 839–849 (1999).
2. Mercader, N. et al. Opposing RA and FGF signals control proximodistal vertebrate limb development through regulation of *Meis* genes. *Development* **127**(18), 3961–3970 (2000).
3. Zhang, X., Friedman, A., Heaney, S., Purcell, P. & Maas, R. L. *Meis* homeoproteins directly regulate *Pax6* during vertebrate lens morphogenesis. *Genes Dev.* **16**(16), 2097–2107 (2002).
4. Conte, I. et al. miR-204 is required for lens and retinal development via *Meis2* targeting. *Proc. Natl. Acad. Sci. U.S.A.* **107**(35), 15491–15496 (2010).
5. Machon, O., Masek, J., Machonova, O., Krauss, S. & Kozmik, Z. *Meis2* is essential for cranial and cardiac neural crest development. *BMC Dev. Biol.* **15**, 40 (2015).
6. Johansson, S. et al. Haploinsufficiency of *MEIS2* is associated with orofacial clefting and learning disability. *Am. J. Med. Genet. A* **164a**(7), 1622–1626 (2014).
7. Gambin, T. et al. Identification of novel candidate disease genes from de novo exonic copy number variants. *Genome Med.* **9**(1), 83 (2017).
8. Zhang, B., Liu, M., Fong, C. T. & Iqbal, M. A. *MEIS2* (15q14) gene deletions in siblings with mild developmental phenotypes and bifid uvula: documentation of mosaicism in an unaffected parent. *Mol. Cytogenet.* **14**(1), 58 (2021).
9. Su, J. X., Velsher, L. S., Juusola, J. & Nezarati, M. M. *MEIS2* sequence variant in a child with intellectual disability and cardiac defects: expansion of the phenotypic spectrum and documentation of low-level mosaicism in an unaffected parent. *Am. J. Med. Genet. A* **185**(1), 300–303 (2021).
10. Santoro, C. et al. A novel *MEIS2* mutation explains the complex phenotype in a boy with a typical *NF1* microdeletion syndrome. *Eur. J. Med. Genet.* **64**(5), 104190 (2021).
11. Erdogan, F. et al. Characterization of a 5.3 mb deletion in 15q14 by comparative genomic hybridization using a whole genome tiling path BAC array in a girl with heart defect, cleft palate, and developmental delay. *Am. J. Med. Genet. A* **143a**(2), 172–178 (2007).
12. Chen, C. P. et al. A 5.6-Mb deletion in 15q14 in a boy with speech and language disorder, cleft palate, epilepsy, a ventricular septal defect, mental retardation and developmental delay. *Eur. J. Med. Genet.* **51**(4), 368–372 (2008).
13. Brunetti-Pierri, N. et al. 15q13q14 deletions: phenotypic characterization and molecular delineation by comparative genomic hybridization. *Am. J. Med. Genet. A* **146a**(15), 1933–1941 (2008).
14. Crowley, M. A. et al. Further evidence for the possible role of *MEIS2* in the development of cleft palate and cardiac septum. *Am. J. Med. Genet. A* **152a**(5), 1326–1327 (2010).
15. Roberti, M. C. et al. Complex chromosome rearrangements related 15q14 microdeletion plays a relevant role in phenotype expression and delineates a novel recurrent syndrome. *Orphanet J. Rare Dis.* **6**, 17 (2011).
16. Chen, C. P. et al. Prenatal diagnosis and molecular cytogenetic characterization of a de novo 4.858-Mb microdeletion in 15q14 associated with *ACTC1* and *MEIS2* haploinsufficiency and tetralogy of Fallot. *Taiwan. J. Obstet. Gynecol.* **55**(2), 270–274 (2016).
17. Shimojima, K., Ondo, Y., Okamoto, N. & Yamamoto, T. A 15q14 microdeletion involving *MEIS2* identified in a patient with autism spectrum disorder. *Hum. Genome Var.* **4**, 17029 (2017).
18. Verheije, R. et al. Heterozygous loss-of-function variants of *MEIS2* cause a triad of palatal defects, congenital heart defects, and intellectual disability. *Eur. J. Hum. Genet.* **27**(2), 278–290 (2019).
19. Schluth-Bolard, C. et al. Whole genome paired-end sequencing elucidates functional and phenotypic consequences of balanced chromosomal rearrangement in patients with developmental disorders. *J. Med. Genet.* **56**(8), 526–535 (2019).
20. Louw, J. J. et al. *MEIS2* involvement in cardiac development, cleft palate, and intellectual disability. *Am. J. Med. Genet. A* **167a**(5), 1142–1146 (2015).
21. Fujita, A. et al. De novo *MEIS2* mutation causes syndromic developmental delay with persistent gastro-esophageal reflux. *J. Hum. Genet.* **61**(9), 835–838 (2016).
22. Srivastava, S. et al. Monogenic disorders that mimic the phenotype of Rett syndrome. *Neurogenetics* **19**(1), 41–47 (2018).
23. Douglas, G. et al. De novo missense variants in *MEIS2* recapitulate the microdeletion phenotype of cardiac and palate abnormalities, developmental delay, intellectual disability and dysmorphic features. *Am. J. Med. Genet. A* **176**(9), 1845–1851 (2018).
24. Hildebrand, M. S. et al. Severe childhood speech disorder: gene discovery highlights transcriptional dysregulation. *Neurology* **94**(20), e2148–e2167 (2020).
25. Giliberti, A. et al. *MEIS2* gene is responsible for intellectual disability, cardiac defects and a distinct facial phenotype. *Eur. J. Med. Genet.* **63**(1), 103627 (2020).
26. Gangfuß, A. et al. Intellectual disability associated with craniofacial dysmorphism, cleft palate, and congenital heart defect due to a de novo *MEIS2* mutation: a clinical longitudinal study. *Am. J. Med. Genet. A* **185**(4), 1216–1221 (2021).
27. Barili, V. et al. Success and pitfalls of genetic testing in undiagnosed diseases: whole exome sequencing and Beyond. *Genes (Basel)* **14**(6) (2023).
28. Carvalho, C. M. et al. Replicative mechanisms for CNV formation are error prone. *Nat. Genet.* **45**(11), 1319–1326 (2013).
29. Mo, A. et al. A recurrent de novo variant in *NUSAP1* escapes nonsense-mediated decay and leads to microcephaly, epilepsy, and developmental delay. *Clin. Genet.* **104**(1), 73–80 (2023).

30. Ek, M. et al. Genome sequencing with comprehensive variant calling identifies structural variants and repeat expansions in a large fraction of individuals with ataxia and/or neuromuscular disorders. *Front. Neurol.* **14**, 1170005 (2023).
31. Stranneheim, H. et al. Integration of whole genome sequencing into a healthcare setting: high diagnostic rates across multiple clinical entities in 3219 rare disease patients. *Genome Med.* **13**(1), 40 (2021).
32. Lindstrand, A. et al. From cytogenetics to cytogenomics: whole-genome sequencing as a first-line test comprehensively captures the diverse spectrum of disease-causing genetic variation underlying intellectual disability. *Genome Med.* **11**(1), 68 (2019).
33. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**(18), 3094–3100 (2018).
34. Sedlazeck, F. J. et al. Accurate detection of complex structural variations using single-molecule sequencing. *Nat. Methods* **15**(6), 461–468 (2018).
35. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**(1), 15–21 (2013).
36. Yépez, V. A. et al. Detection of aberrant gene expression events in RNA sequencing data. *Nat. Protoc.* **16**(2), 1276–1296 (2021).
37. Brechtmann, F. et al. OUTRIDER: a statistical method for detecting aberrantly expressed genes in RNA sequencing data. *Am. J. Hum. Genet.* **103**(6), 907–917 (2018).
38. Haas, B. J. et al. Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol.* **20**(1), 213 (2019).

Acknowledgements

The authors acknowledge support from the National Genomics Infrastructure in Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council, and the NAISS/Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with massively parallel sequencing and access to the UPPMAX computational infrastructure.

Author contributions

A.L. conceptualized the study. B-M.A. wrote the phenotype description and performed the medical investigations. M.P., J.E. and M.E. did the genomic analysis and J.E., H.T. and M.E. did the analysis of the transcriptome. A.L., M.K. and M.J.S. evaluated the findings. M.E. designed the figures. M.E. wrote the original draft, and all the other authors contributed to the final manuscript. A.L. acquired funding for the project.

Funding

Open access funding provided by Karolinska Institute. This work was supported by the Swedish Research Council [2019–02078], the Swedish Brain Fund [FO2022-0256], the Stockholm City Council and the Swedish Rare Diseases Research Foundation (Sällsyntafonden). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-81507-7>.

Correspondence and requests for materials should be addressed to M.E.

Reprints and permissions information is available at 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024