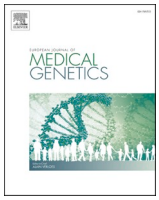
European Journal of Medical Genetics 69 (2024) 104949

Contents lists available at ScienceDirect

European Journal of Medical Genetics

journal homepage: www.elsevier.com/locate/ejmg



Clinical Report

Identification of a pathogenic deep intronic variant in *ATRX* ends a diagnostic odyssey

Jasper J. van der Smagt [a], Angeliki P. Lampri [a], Iris de Lange [a], Mariëlle Alders [b], Michiel L. Houben [c], Marco J. Koudijs [a], Richard H. van Jaarsveld [a,*]

[a] *Department of Genetics, University Medical Center Utrecht, Utrecht, the Netherlands*
[b] *Amsterdam UMC, Department of Human Genetics, Genome Diagnostics Laboratory, Amsterdam Reproduction & Development, University of Amsterdam, Amsterdam, the Netherlands*
[c] *Department of Pediatrics, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, the Netherlands*

ARTICLE INFO

Handling Editor: A. Verloes

ABSTRACT

Variation in the non-coding genome is being increasingly recognized to be involved in monogenic disease etiology. However, the interpretation of non-coding variation is complicated by a lack of understanding of how non-coding genetic elements function. Additional lines of evidence are therefore needed to recognize non-coding variants as pathogenic. We here present a case where a collective body of evidence resulted in the identification and conclusive classification of a pathogenic deep intronic variant in *ATRX*. This report demonstrates the utility of a multi-platform approach in aiding the identification of pathogenic variants outside coding regions. Furthermore, it marks the first reported instance of a deep intronic pathogenic variant in *ATRX*.

1. Introduction

Variation in the non-coding genome is being increasingly recognized to underlie monogenic disorders. However, diagnosing patients based on non-coding variants remains challenging, mainly due to our limited capacity to accurately predict the functional consequences of such variants. Additional lines of evidence are therefore needed to support a pathogenic classification of most non-coding variants. Typically, this evidence entails the detection of aberrant transcriptomic events, making RNA-sequencing (RNA-seq) a powerful complementary approach in genome diagnostics.

ATRX-syndrome (OMIM: #309580) is an X-linked syndromic form of usually severe to profound mental retardation. The vast majority (>90%) has a characteristic recognizable facial gestalt during infancy, ~80% has some degree of genital abnormality and ~90% has hematological signs of Alpha-Thalassemia (Leon and Harley, 2021). Female carriers are generally unaffected but show marked skewing of X-chromosome inactivation to silence the pathogenic allele (Wada et al., 2005).

ATRX-syndrome is caused by pathogenic variants in *ATRX*, a chromatin remodeler involved in regulation of transcription and replication. Reported pathogenic variants are primarily missense variants, generally resulting in reduced protein levels (Badens et al., 2006; Villard and Fontes, 2002). However, other types of variants have been identified as well, including splice site variants and larger copy number variants (CNVs) (Leon and Harley, 2021).

In line with its role in epigenetic regulation, pathogenic variants in *ATRX* are associated with a specific genome-wide DNA-methylation pattern (Schenkel et al., 2017). Such methylation patterns (a.k.a. episignatures) have been identified for >65 monogenic disease genes and are emerging as powerful diagnostic tools (Levy et al., 2022; Sadikovic et al., 2021).

We here report the identification of a novel pathogenic variant causing ATRX-syndrome, which is located deep within an intron and results in inclusion of a cryptic exon. The classification of this variant is supported by results from multiple platforms, including Episign testing and RNA-seq. This study demonstrates the power of such a multi-platform approach and reports the first deep intronic single nucleotide variant (SNV) to be associated with ATRX-syndrome.

* Corresponding author.
E-mail address: r.h.vanjaarsveld@umcutrecht.nl (R.H. van Jaarsveld).

https://doi.org/10.1016/j.ejmg.2024.104949
Received 15 February 2024; Received in revised form 16 May 2024; Accepted 23 May 2024
Available online 24 May 2024

2. Results

2.1. Case presentation

The proband is a male (III:1, Fig. 1A), currently 9 years of age, without speech and unable to walk without support. He is the first and only child of healthy, highly educated, non-consanguineous Dutch parents. The boy was born following an uneventful pregnancy (delivery at 38+5 weeks) with a birthweight of 3520g. He was borderline microcephalic at birth and his head circumference (OFC) has continued to grow at −2,5 to −3,0SD. He failed his postnatal hearing exam due to a bilateral sensorineural loss of ~50Db. He was also shown to have limited vision (abnormal VEP and ERG tests), which results in poor fixation. His development was delayed from the start, accompanied by severe axial hypotonia; however, no signs of developmental regression were noted. Upon dysmorphological examination, several features were noted, including: plagiocephaly (likely due to prolonged periods of lying on this back in the same position); mild hypertelorism with down slanting palpebral fissures; short nose with anteverted nares; a hypotonic face with triangular mouth and full lips; eversion of the lower lip; and tongue protrusion. The teeth appeared small and widely spaced and one element with synodontia was noted. The boy has normal male genitalia with descended testes. He usually has a happy demeanor with frequent bouts of laughter. He has had seizures associated with fever but does not have epilepsy.

The proband remains under the care of a pediatrician because of excessive drooling (treated with yearly botox injections), constipation, and recurrent urinary tract infections. He does eat but has a gastrostoma to ensure sufficient intake. Hematological testing found mild anemia on several occasions, however no microcytic anemia was detected. Furthermore, Alpha-Thalassemia testing, including testing for HBH inclusions, was negative on two occasions. Family history was non-

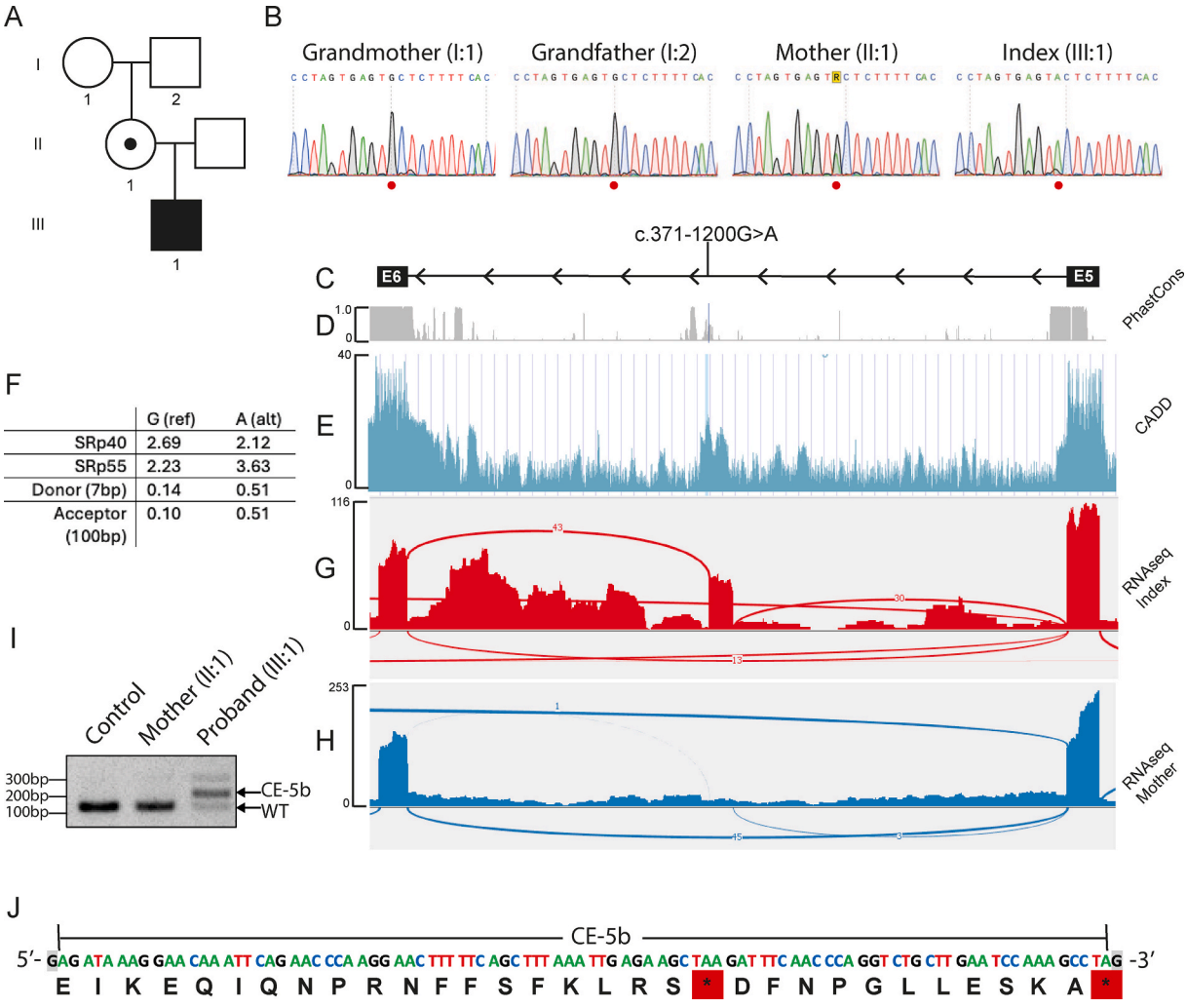


Fig. 1. Identification of a pathogenic deep intronic variant in *ATRX*. A) Pedigree of the presented family. Siblings have been left out of this pedigree for privacy reasons. B) Sanger sequencing results demonstrating the *de novo* occurrence of the c.371-1201G>A (NM_000489.6) variant. Red dots indicate the variant location. C) Schematic representation of the genetic locus harboring the c.371-1201G>A variant. *ATRX* is encoded by the minus strand. Shown are intron 5 (black line, 2639bp) and its flanking exons exon 5 (E5) and exon 6 (E6) from *ATRX* (NM_000489.6). Arrows indicate reading frame direction. D) Histogram showing phastCons scores (Y-axis) for the respective locations. Blue line indicates location of the variant. E) Histogram showing CADD scores (Y-axis) for respective locations. Blue line indicates location of the variant. F) Table depicting splice enhancer binding consensus scores from ESE finder and SpliceAI scores (+7 and + 100bp on the genomic sequence) for reference (ref) and the variant (alt) allele. G and H) Sashimi plots summarizing the results from RNA-seq analysis of blood samples of the index (III:1, F) and mother (II:1, G). Y-axis: read depth. I) RT-PCR results indicating the inclusion of CE-5b in the proband (III:1), but not the mother (II:1) and a random unrelated control sample. Primers were designed to target exon 5 and 6 to generate a 168bp product (WT). CE-5b is 94bp long, corresponding to the size shift as seen in the proband. J) Coding sequence of CE-5b (top row) including the nucleotides from the flanking exons to complete the first and last codon (highlighted in grey), and the translated protein sequences (bottom row). In-frame stopcodons are depicted in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

contributory, except for multiple miscarriages with normal parental karyotypes. Based on the combination of severe intellectual disability, hypotonia and characteristic facial features, a clinical diagnosis of ATRX-syndrome was expected.

2.2. Clinical genetic testing

Diagnostic whole-exome-sequencing (WES) identified no suspect *ATRX* SNVs or indels after analysis according to the ACMG/AMP guidelines (Richards et al., 2015). In addition, no suspect CNVs were detected by SNP-array or WES-based CNV analysis.

Because initial genetic testing did not identify a pathogenic *ATRX* variant, we looked for further evidence to support the suspicion of ATRX-syndrome. First, we tested X-inactivation in the mother (II:1, Fig. 1A) and found a fully (100%) skewed X-inactivation pattern (data not shown). X-inactivation testing in the grandmother (I:1) was inconclusive, as she is homozygous for the target repeat region of the assay (data not shown).

Next, we performed diagnostic Episign testing for the proband (III:1), which showed a methylation signature consistent with ATRX-syndrome (data not shown). These results, combined with the clinical presentation of the proband, unequivocally supported the diagnosis of ATRX-syndrome, prompting us to continue our search for the underlying pathogenic variant.

2.3. Identification of a pathogenic deep intronic variant in ATRX

As previous WES-based efforts did not identify any potential pathogenic variant, we hypothesized a non-coding variant to underlie the phenotype. Since the family history showed no indication of a hereditary disease (i.e. no other affected family members and no recurrent miscarriages besides in the mother (II:1)), we expected the variant to be absent in the grandmother (I:1) and to have formed *de novo* in the mother (II:1). Under this hypothesis, we performed WGS for the boy (III:1), his mother (II:1) and grandmother (I:1) and searched for shared SNVs between the proband (III:1) and his mother (II:1) but absent in the grandmother (I:1). Only one SNV fitted our hypothesis: c.371-1201G>A (NM_000489.6). Sanger sequencing confirmed the genotypes and demonstrated absence of the variant in the grandfather (I:2) (Fig. 1B), confirming the variant formed *de novo* in the mother (II:1).

The variant is absent from the gnomAD database (Karczewski et al., 2020) and resides in a conserved intronic element according to PhastCons scores (Fig. 1C and D) (Siepel et al., 2005). In addition, the CADD score of this nucleotide is 21 (Kircher et al., 2014; Rentzsch et al., 2019), and the region shows elevated CADD scores as compared to the direct surroundings (Fig. 1E). Moreover, the variant was predicted to shift the affinity for splice enhancers from SRp40 to SRp55, and SpliceAI predicts the introduction of a splice donor and an acceptor site, 7bp and 100bp upstream in the transcript, respectively (Fig. 1F) (Cartegni et al., 2003; Jaganathan et al., 2019). These data collectively suggested that the variant might affect a functional non-coding element residing deep within intron 5 (NM_000489.6).

We next performed RNA-seq on peripheral blood samples to detect potential splicing impact of the variant. We observed inclusion of a cryptic exon spanning 94nt (GRCh38 chrX:77695145-77695238), located directly upstream of the c.371-1201G>A variant and corresponding to the splice sites as predicted by SpliceAI (Fig. 1F). Inclusion of this cryptic exon (hereafter referred to as cryptic exon 5b, CE-5b) was commonly observed in the proband (III:1, Fig. 1G), while inclusion of CE-5b was rare in the mother (II:1, Fig. 1H), consistent with skewed X-inactivation and silencing of the allele. The RNA-seq results were confirmed by RT-PCR (Fig. 1I).

CE-5b introduces an in-frame premature stop codon (Fig. 1J) and is therefore likely to severely impact on gene function, either by producing a truncated protein missing all the downstream exons (6–35), or by complete loss of expression due to nonsense-mediated decay (NMD).

In conclusion, we identified a novel pathogenic deep intronic variant in *ATRX* which formed *de novo* in an unaffected mother. The variant causes inclusion of CE-5b, a cryptic exon which introduces an in-frame stopcodon and therefore likely impacts on ATRX function. This variant has been submitted to ClinVar (Variation ID: 3233402).

3. Discussion

The contribution of non-coding variation to monogenic disease is gaining recognition in clinical genetics. We here demonstrate how multiple platforms can be utilized to guide the search for pathogenic non-coding variation, which is still difficult to detect due to a lack of understanding of the functional impact of such variants. Motivated by a collective body of evidence pointing towards a single candidate gene (*ATRX*), we employed WGS and RNA-seq to detect and confirm a pathogenic variant residing deep within an intron. This variant (c.371-1201G>A) causes the inclusion of a cryptic exon (CE-5b), which introduces an in-frame premature stop codon and as such is expected to severely impact on ATRX function.

Although premature stopcodons are known to induce NMD, the abundant detection of transcripts including CE-5b suggests that transcripts containing this exon might escape NMD. The expected protein product would lack most of the C-terminal part of the protein, as it will contain only the first five exons, which encode for only 123aa of the otherwise 2429aa full length ATRX protein. Therefore, CE-5b is expected to severely impact on ATRX function also in absence of NMD. Further studies will be necessary to elucidate the precise nature of the final gene product.

The introduction of premature stopcodons by inclusion of cryptic exons has been recognized as a regulatory mechanism for gene expression and are commonly referred to as poison exons (Carvill and Mefford, 2020). One striking example of the clinical relevance is a mutational hotspot in *SCN1A*, where pathogenic variants activate a poison exon and cause Dravett Syndrome (Carvill et al., 2018). Interestingly, antisense oligonucleotides targeting these SNVs silence the poison exon and relieve Dravett associated phenotypes in mice (Han et al., 2020). It is tempting to speculate such a therapeutic approach can be translated to many more monogenic disorders, and this report might indicate ATRX-syndrome is such a disorder.

We demonstrate how the identification of a pathogenic deep-intronic variant ended a diagnostic odyssey. By employing different platforms, both classic and novel, a collective body of evidence pointing towards a pathogenic variant in *ATRX* served to strengthen the search for genetic variants outside of the normal diagnostic spectrum. The most convincing argument to focus on *ATRX* was an *ATRX*-positive episignature, demonstrating the value of Episign testing in diagnostics.

The identification of pathogenic variation in non-coding elements remains challenging, mainly due to our inability to adequately predict the functional consequences of such variants on gene expression, splicing, or other biological processes they might influence. RNA-seq provides an unbiased approach to detect aberrant transcriptomic events and as such serves as an outstanding method to provide functional evidence to support the classification of non-coding variants. We here employed this technique to confirm functional impact of a deep intronic variant. Of note, although some publicly available metrics to assist in the classification and/or prioritization of non-coding variants (including CADD, PhastCons and SpliceAI) correctly reflected the functional importance of the genetic element, other splice site prediction tools failed to predict an impact on splicing. An integrated approach using advanced diagnostic tools (such as episignatures and RNA-seq) combined with continuously improving classification metrics, is expected to greatly increase diagnostic yield of clinical WGS efforts and will prevent or shorten diagnostic odysseys for many cases like presented here.

4. Materials and methods

4.1. Diagnostic genetic testing and consent

WES, WGS and X-chromosome silencing assays were performed by the genome diagnostic laboratory at the UMC Utrecht according to local diagnostic protocols and standards. Episign testing was performed through the genome diagnostic laboratory at the Amsterdam UMC, location AMC, as described previously (Sadikovic et al., 2021).

This study is performed within the ethical framework of the UMC Utrecht. Consent forms, signed by the mother, that allow for publication of the presented data, make part of the electronic health record stored at the UMC Utrecht.

4.2. RNA sequencing and RT-PCR

RNA was isolated from total blood. Sequencing libraries were prepared using the KAPA RNA HyperPrep with RiboErase kit (Roche) and TrueSeq RNA sample preparation kit (Illumina) and sequenced on a NovaSeq 6000 (Illumina). Quality of the library was checked with FastQC v0.11.8 and reads were aligned to the human genome reference (GRCh38) by STAR v2.7.9a. To visualize splice junctions from RNAseq data, IGV (2.13.2) was used.

For RT-PCR, cDNA was generated using the High-Capacity cDNA Reverse Transcription Kit (Thermofisher). RT-PCR was performed using the Q5® High-Fidelity DNA Polymerase (New England Biolabs) and primers for the exons 5 and 6 of ATRX with the following sequence: 5′-AAAACCTTTGGATGATGAAACTGT-3′ and 5′-TCTGCTTCTAAATTCAGGCCC-3′, respectively. RT-PCR products were analyzed by agarose gel electrophoresis.

4.3. Variant analysis

Predicted variant impact was analyzed using the following tools: Alamut Visual (v2.15) to access PhasCons data (44 vertebrates) and splice predictor software (SpliceSiteFinder-like, MaxEntScan, NNSPLICE and GeneSplicer), the UCSC genome browser for CADD (v1.6) and gnomAD V4.1.0.

Accession numbers

ClinVar Variation ID: 3233402.

CRediT authorship contribution statement

Jasper J. van der Smagt: Investigation, Writing – original draft. **Angeliki P. Lampri:** Investigation, Writing – review & editing, Visualization. **Iris de Lange:** Investigation, Writing – review & editing. **Mariëlle Alders:** Investigation, Resources, Writing – review & editing. **Michiel L. Houben:** Investigation, Writing – review & editing. **Marco J. Koudijs:** Validation, Resources, Writing – review & editing, Supervision. **Richard H. van Jaarsveld:** Conceptualization, Writing – original draft, Visualization, Supervision, Funding acquisition, Project administration, Formal analysis.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

The data that has been used is confidential.

Acknowledgements

We thank the presented family for their contributions to and permission for this report; the genome diagnostics and bioinformatics sections at the department of genetics, UMC Utrecht, for support and access to facilities; and the Laboratory for Childhood Cancer Pathology of the Princess Máxima Centre Utrecht for RNA-sequencing and bioinformatic analyses. RHvJ is supported by the Wilhelmina Kinderziekenhuis Onderzoeksfonds.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmg.2024.104949.

References

Badens, C., Lacoste, C., Philip, N., Martini, N., Courrier, S., Giuliano, F., Levy, N., 2006. Mutations in PHD-like domain of the ATRX gene correlate with severe psychomotor impairment and severe urogenital abnormalities in patients with ATRX syndrome. Clin. Genet. 70 (1), 57–62. https://doi.org/10.1111/j.1399-0004.2006.00641.x.

Cartegni, L., Wang, J., Zhu, Z., Zhang, M.Q., Krainer, A.R., 2003. ESEfinder: a web resource to identify exonic splicing enhancers. Nucleic Acids Res. 31 (13), 3568–3571. https://doi.org/10.1093/nar/gkg616.

Carvill, G.L., Engel, K.L., Ramamurthy, A., Cochran, J.N., Roovers, J., Stamberger, H., Mefford, H.C., 2018. Aberrant inclusion of a poison exon causes Dravet syndrome and related SCN1A-associated genetic epilepsies. Am. J. Hum. Genet. 103 (6), 1022–1029. https://doi.org/10.1016/j.ajhg.2018.10.023.

Carvill, G.L., Mefford, H.C., 2020. Poison exons in neurodevelopment and disease. Curr. Opin. Genet. Dev. 65, 98–102. https://doi.org/10.1016/j.gde.2020.05.030.

Han, Z., Chen, C., Christiansen, A., Ji, S., Lin, Q., Anumonwo, C., Isom, L.L., 2020. Antisense oligonucleotides increase Scn1a expression and reduce seizures and SUDEP incidence in a mouse model of Dravet syndrome. Sci. Transl. Med. 12 (558) https://doi.org/10.1126/scitranslmed.aaz6100.

Jaganathan, K., Kyriazopoulou Panagiotopoulou, S., McRae, J.F., Darbandi, S.F., Knowles, D., Li, Y.I., Farh, K.K., 2019. Predicting splicing from primary sequence with deep learning. Cell 176 (3), 535–548 e524. https://doi.org/10.1016/j.cell.2018.12.015.

Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alfoldi, J., Wang, Q., MacArthur, D.G., 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581 (7809), 434–443. https://doi.org/10.1038/s41586-020-2308-7.

Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., Shendure, J., 2014. A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. 46 (3), 310–315. https://doi.org/10.1038/ng.2892.

Leon, N.Y., Harley, V.R., 2021. ATR-X syndrome: genetics, clinical spectrum, and management. Hum. Genet. 140 (12), 1625–1634. https://doi.org/10.1007/s00439-021-02361-5.

Levy, M.A., McConkey, H., Kerkhof, J., Barat-Houari, M., Bargiacchi, S., Biamino, E., Sadikovic, B., 2022. Novel diagnostic DNA methylation episignatures expand and refine the epigenetic landscapes of Mendelian disorders. HGG Adv 3 (1), 100075. https://doi.org/10.1016/j.xhgg.2021.100075.

Rentzsch, P., Witten, D., Cooper, G.M., Shendure, J., Kircher, M., 2019. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res. 47 (D1), D886–D894. https://doi.org/10.1093/nar/gky1016.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Committee, A.L.Q.A., 2015. 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. 17 (5), 405–424. https://doi.org/10.1038/gim.2015.30.

Sadikovic, B., Levy, M.A., Kerkhof, J., Aref-Eshghi, E., Schenkel, L., Stuart, A., Alders, M., 2021. Clinical epigenomics: genome-wide DNA methylation analysis for the diagnosis of Mendelian disorders. Genet. Med. 23 (6), 1065–1074. https://doi.org/10.1038/s41436-020-01096-4.

Schenkel, L.C., Kernohan, K.D., McBride, A., Reina, D., Hodge, A., Ainsworth, P.J., Sadikovic, B., 2017. Identification of epigenetic signature associated with alpha thalassemia/mental retardation X-linked syndrome. Epigenet. Chromatin 10, 10. https://doi.org/10.1186/s13072-017-0118-4.

Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Haussler, D., 2005. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 15 (8), 1034–1050. https://doi.org/10.1101/gr.3715005.

Villard, L., Fontes, M., 2002. Alpha-thalassemia/mental retardation syndrome, X-Linked (ATR-X, MIM #301040, ATR-X/XNP/XH2 gene MIM #300032). Eur. J. Hum. Genet. 10 (4), 223–225. https://doi.org/10.1038/sj.ejhg.5200800.

Wada, T., Sugie, H., Fukushima, Y., Saitoh, S., 2005. Non-skewed X-inactivation may cause mental retardation in a female carrier of X-linked alpha-thalassemia/mental retardation syndrome (ATR-X): X-inactivation study of nine female carriers of ATR-X. Am. J. Med. Genet. 138 (1), 18–20. https://doi.org/10.1002/ajmg.a.30901.