




CLINICAL REPORT

Novel frameshift mutation in PURA gene causes severe encephalopathy of unclear cause

Lucía Spangenberg^{1,2}  | Rosario Guecaimburú³ | Alejandra Tapié⁴ | Susana Vivas³ | Soledad Rodríguez⁴ | Martín Graña¹  | Hugo Naya^{1,5}  | Víctor Raggio⁴

¹Unidad de Bioinformática, Institut Pasteur de Montevideo, Montevideo, Uruguay

²Departamento de Informática y Ciencias de la computación, Facultad de Ingeniería, Universidad Católica del Uruguay, Montevideo, Uruguay

³Equipo de Enfermedades Raras, CRENADECER, BPS, Montevideo, Uruguay

⁴Departamento de Genética, Facultad de Medicina, Universidad de la República, Montevideo, Uruguay

⁵Departamento de Producción Animal y Pasturas, Facultad de Agronomía, Universidad de la República, Montevideo, Uruguay

Correspondence

Víctor Raggio, Departamento de Genética, Facultad de Medicina, Universidad de la República, Gral. Flores 2125, Montevideo, Uruguay.
Email: vraggio@fmed.edu.uy

Funding information

URUGENOMES, Grant/Award Number: IP011-17/CEI/LC/MB

Abstract

Background: The etiology of many genetic diseases is challenging. This is especially true for developmental disorders of the central nervous system, since several genes can be involved. Many of such pathologies are considered rare diseases, since they affect less than 1 in 2000 people. Due to their low frequency, they present several difficulties for patients, from the delay in the diagnosis to the lack of treatments. Next-generation sequencing techniques have improved the search for diagnosis in several pathologies. Many studies have shown that the use of whole-exome/genome sequencing in rare Mendelian diseases has a diagnostic yield between 30% and 50% depending on the disease.

Methods: Here, we present the case of an undiagnosed 6-year-old boy with severe encephalopathy of unclear cause, whose etiological diagnosis was achieved by whole-genome sequencing.

Results: We found a novel variant that has not been previously reported in patients nor it has been described in GnomAD. Segregation analysis supports a de novo mutation, since it is not present in healthy parents. The change is predicted to be harmful to protein function, since it falls in the first quarter of the protein producing an altered reading frame and generating a premature stop codon. Additionally, the variant is classified as pathogenic according to ACMG criteria (PVS1, PM2, and PP3). Furthermore, there are several reported frameshift mutations in nearby codons as well as nonsense mutations that are predicted as pathogenic in other studies.

Conclusion: We found a novel de novo frameshift mutation in the PURA gene (MIM number 600473), c.151_161del, with sufficient evidence of its pathogenicity.

1 | INTRODUCTION

Specific diagnosis of many genetic diseases is challenging. One area where greater difficulties are met is that of developmental disorders of the central nervous system, since many genes can be involved (van Loo & Martens,

2007) with many more to be discovered probably. In more broad terms, rare diseases (RD) are pathologies that affect less than 1 in 2000 people (Commission & – European Commission, 2020). Due to their low frequency, they present diverse difficulties, from the delay in the diagnosis to the lack of specific treatments. Most of them have a

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *Molecular Genetics & Genomic Medicine* published by Wiley Periodicals LLC.

substantial impact on quality and life expectancy, affecting children and young people.

Since the diagnosis itself is the first obstacle that RD patients face, from which most medical decisions will rely on, it becomes of crucial importance to have a precise and accurate diagnosis. Molecular genomics approaches have helped in this sense. Next-generation sequencing (NGS) techniques have highly improved diagnosis of RD. Many studies have shown that the use of whole-exome sequencing (WES) in rare Mendelian diseases has a diagnostic yield between 30% and 50% depending on the disease (Clark et al., 2018; Yang et al., 2014). The use of whole-genome sequencing (WGS) in a cohort of pediatric patients has found causative genetic variants in 34% of the cases (Stavropoulos et al. 2016). These results encouraged our group to use these tools (both WES/WGS) to help with the diagnosis of pediatric patients with RD.

Here, we present the case of a 6-year-old boy presenting with severe encephalopathy with unclear cause, which started in the first months of life, whose etiological diagnosis was made by WGS.

2 | METHODS

2.1 | Ethical compliance

This project (URUGENOMES Ref IP011-17/CEI/LC/MB) was approved by the ethics committee of the Institut Pasteur de Montevideo.

A written informed consent was elaborated for this study and was signed by the patients' parents.

2.2 | NGS sequencing and bioinformatics analysis

Genomic DNA was extracted from 100 μ l of whole blood using QIAamp[®] DNA Blood Mini kit (Qiagen, Germany) according to manufacturer instructions.

We did a whole-genome sequencing of the patient with 30X in a HiSeq X Ten Illumina sequencer. Quality of reads was analyzed using FastQC (Andrews, 2010), and they were mapped onto the human genome (GRCh37) using BWA (Li & Durbin, 2009). Variant calling was performed using GATK (best practices) (McKenna et al., 2010). Annotation of found variants was done with ANNOVAR (Wang et al., 2010). Series of filters were applied in order to detect relevant mutations. Candidate mutation (frameshift deletion) was further evaluated with the SIFT Indel tool [9], in order to address pathogenicity.

This relevant mutation found in the patient was evaluated via Sanger sequencing.

Additionally, the mitochondrial genome was analyzed using MToolBox (Calabrese et al., 2012).

2.3 | Variants filtering scheme

In order to filter and prioritize the variants found, we used the following rationale:

1. Homozygous or hemizygous mutations with a frequency lower than 1% in coding/splicing region;
2. Heterozygous mutations with at least two variants in the same gene with frequency lower than 1% (compound heterozygous) in coding/splicing region;
3. Heterozygous mutations with frequency less than 0.5% in coding/splicing region;
4. Mitochondrial mutations with high heteroplasmy (>10%) and in coding regions or tRNA and rRNA genes (and not part of the definition of the haplogroup), not in D-Loop region;
5. Noncoding variants, either with "uncertain significance" (VUS) or "pathogenic/Likely pathogenic" or "conflicting interpretations of pathogenicity" classifications, as determined by ClinVar (Landrum et al., 2018).

3 | RESULTS

3.1 | Case report

A 6-year-old boy was referred to the medical genetics' unit for evaluation for neurodevelopmental delay and epilepsy. He is the son of a non-consanguineous healthy couple who has three healthy children from other marriages, two from the mother's and one from the father's side. The child is the younger one and the product of a third uncomplicated pregnancy. The mother did not take alcohol or illegal drugs and was not exposed to teratogens during pregnancy. Fetal movements were normal as well as fetal ultrasounds. A C-section delivery was performed at week 40 due to cephalopelvic disproportion. The weight at birth was 4150 grams, length 50.5 cm, Apgar score: 3/8. He presented involuntary rhythmic movements of the diaphragm, difficulties in sucking, weak cry, and jaundice that required phototherapy. From the examination of this period we can highlight: myopathic facies, no spontaneous eye opening with intermittent eye fixation, severe hypotonia with quadriparesis, and increased tendon reflexes.

Excessive somnolence and lethargy were present in the neonatal period as well as in the first year of life. This feature showed slow improvement since the age of 2 years.

Generalized neurodevelopmental delay was evident afterwards. Regarding motor skills, he reached cephalic support at

the age of one year, and he was able to sit unsupported at the age of two. He is still not able to walk. No verbal language was developed so far.

At 2 years old, he began to suffer from seizures which are multifocal and generalized. Some of these crises are nocturnal. Currently, he is under treatment with valproic acid and levetiracetam with amelioration of the number of crises but not achieving a complete remission. Sleep apneas and dystonic movements appeared in the last year. Apart from “hypotonic face” no other dysmorphic signs were evident. Growth was normal in all somatic and cranial parameters.

The exams performed showed additional findings as follows:

Abnormal (prechiasmatic) visual evoked potentials. Auditory evoked potentials were normal.

Electroencephalograms were normal at the beginning. The last one (induced sleep) showed bifocal epileptic activity: intense activity consisting of spike-and-slow-wave complexes in temporal and parietal regions of the left hemisphere, and moderate activity in temporal regions of the right hemisphere. Background rhythms were normal.

Polysomnography showed severe sleep apnea. Gastroesophageal reflux is present (coincident with apneas) as measured by esophageal pH.

Magnetic resonance imaging was performed twice (age 2 and 4), being normal in both cases. Previous genetic studies were performed with normal results: karyotype (46, XY), CGH array, and methylation studies for PWS/ANG region (MS-MLPA). Additionally, amino acid dosification in plasma and organic acids in urine were measured with normal results.

3.2 | GS results

From a whole blood sample from the patient whole-genome sequencing (nuclear and mitochondrial DNA) was performed with a sequencing depth of 30x. We obtained 709.502.196 reads that passed QC-controls (according to samtools flagstat) and ~94% (669945761) were mapped onto the reference genome (GRCh37). Variant calling analysis detected 4925231 variants that were further annotated and prioritized (see 2.3). For the mitochondrial genome a high sequencing depth (2999x) and 100% coverage were obtained. A total of 273 variants were detected in the mtDNA.

3.3 | Frameshift variant in the *PURA* gene is a candidate pathogenic mutation

Results of our filtering rationale (see methods 4.2), are found in Table 1. An autosomal recessive inheritance for

TABLE 1 Results of variant filtering

	1% hom	1% comp. Het	0.5% het	chrM
# variants	24	477	427	29
# genes	14	75	196	12
Missense	19	273	373	9
Nonsense	1	10	14	0
Stoploss	0	1	1	0
fs indel	1	15	17	0
Non-fs indel	0	2	0	0
Silent	3	162	0	18
Splicing	0	3	5	0
Others	.	Unknown 10	Unknown 17	2 tRNA

Columns show each of the four filtering categories: homozygous variants with a population frequency of less than 1%; heterozygous variants with a population frequency of less than 1%, with at least two variants in the same gene; heterozygous variants with a population frequency of less than 0.5%. Rows have the total number of variants found and the number of genes involved, and below the different types of variants found in each category. Unknown category corresponds to various ambiguities in the gene structure definition in the database file. fs, frameshift; non-fs, non-frameshift. # symbol represents number.

the patient's phenotype was plausible, since both parents were unaffected. For this reason, on the one hand, coding variants in homozygosity with low frequency were considered. Out of the 24 (21 non-synonymous variants) none of them were related to a gene with a concordant phenotype. Additionally, compound heterozygous variants were considered. Out of the 75 genes, neither of them was previously associated with disease, or the phenotype associated was not concordant to the patient's. Another possibility was a de novo mutation, appearing for the first time in a heterozygous state in the patient. For this reason, we investigated all heterozygous mutations in coding regions with very low frequency (<0.5%). Here, we found a frameshift deletion of 11 nucleotides (CCAGGGGGGCT, c.151_161del) (Figure 1a) in the only exon of the *PURA* gene (chr5:139493917–139493927, GRCh37, *PURA*:NM_005859:exon1:c.151_161del:p.P51fs) not previously reported in any of the population genomic projects (such as 1000 Genomes (The 1000 Genomes Project Consortium, 2015)/GnomAD (Karczewski et al., 2020) /ExAC (Lek et al. 2016)). The deletion occurs between the 51th and the 54th amino acid (the first 16% of the protein), it causes a stop codon at codon 147 after the deletion. Therefore, not only the reading frame is altered for 147 amino acids, but also it results in a 130 amino acid shorter protein. Additionally, the variant was evaluated with SIFT Indel 2 (Pauline, 2003) to predict its pathogenicity from an in silico point of view. It was predicted as “damaging” with a confident score of 0.858.

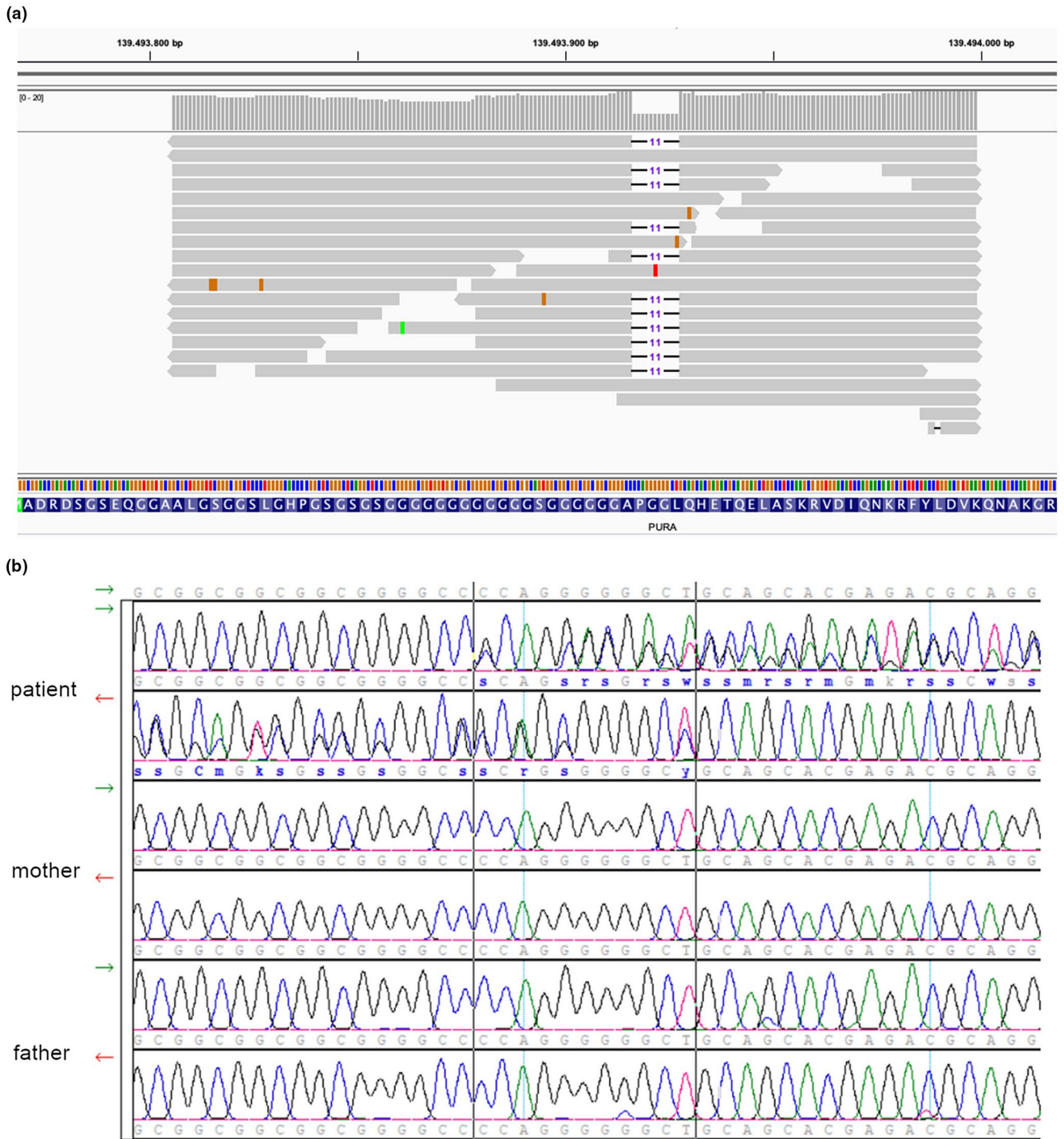


FIGURE 1 (a) IGV view of NGS result of patient. This view shows the reads mapped onto the reference PURA gene. On top, a ruler marks absolute genomic positions on chromosome 5 (from 139.493.800 to 139.494.000 bp). Below that, on a middle small panel, horizontal bars mark the sequencing depth of each base in the region. Positions with low sequencing depth have small horizontal bars (as is the case with the positions covering the deletion). Below, as horizontal bars, the reads as they were mapped onto the genome are shown. Deletions within reads are represented with black lines with a number denoting the length of the deletion. Note that only the region of interest is shown (not the whole gene). (b) Sanger sequencing validation of patient, mother and father

3.4 | De novo status and confirmation of the variant

Even though the quality of the sequencing in the variant position was good (total coverage of the region is 19x: 11 reads

with the deletion and 8 reads without), the deletion was validated and confirmed in the patient via Sanger sequencing. Additionally, both biological parents were analyzed for the presence of the variant with Sanger sequencing. The variant was not found in any of them in white blood cells (Figure 1b).

4 | DISCUSSION

The *PURA* gene is a DNA/RNA-binding protein which functions as a transcription regulator also involved in mRNA localization (Weber, 2016). A number of pathologies, ranging from encephalopathies and mental retardation to leukemia and ALS, have been associated with heterozygous mutations in the *PURA* gene (Daniel, 2018; Johnson & Gordon, 2013). Its expression is ubiquitous with no clear tissue preference (human protein atlas, www.proteinatlas.org). The gene product is a protein with three domain repeats: repeats I and II are binding domains and repeat III is a dimerization domain (Johnson & Gordon, 2013; Weber, 2016). The crystal structure of the first two domains involved in nucleic acid binding from *D. melanogaster* complexed with ssDNA, has provided definitive evidence into the molecular interactions of the *PURA* protein (Weber, 2016), proving its purine base preference allows it for binding both DNA and RNA.

Mutations in the *PURA* gene were reported as a cause of dominant form of severe delayed psychomotor development and seizures a few years ago (Hunt et al., 2014; Lalani et al., 2014; Tanaka et al., 2015) and the *PURA* protein was shown to be a key player in brain development. *PURA* gene knockdown (−/−) mice show severe neurological and brain developmental anomalies and die within 4 weeks after birth (Hokkanen, 2012; Khalili et al., 2003).

Several lines of evidence point to the variant we found as a causative mutation. The change is predicted to be harmful to protein function, since it falls in the first quarter of the protein (AA 51 to 54 of 323 AA) producing an altered reading frame (for 147 AA following the deletion) and generating a premature stop codon (product is 130AA shorter), which leads to two possibly overlapping scenarios. On the one hand, the production of prematurely truncated mRNAs to be processed by the nonsense-mediated mRNA decay pathway (Brogna, & Wen, 2009). On the other hand, the production of a different protein, which we posit would be 195 amino acids: *PURA* gene's first 50 residues followed by a different sequence starting position 51, and therefore, affecting all of the three functional domains of the protein. This predicted protein has no detectable eukaryotic homologs, as checked by an HHpred (Söding, 2005) search over the Protein Data Bank or two iterations of Jackhammer (Potter et al., 2018) over the Reference Proteome database. The production of an abnormal protein may result in multiple cellular problems, including accumulation of misfolded proteins. In this sense, the most probable transcript, if translated, is predicted to have high disorder propensity. Any of those mechanisms probably causes a total loss of function (LoF) of the product encoded by this allele. LoF has been demonstrated to be the main pathogenic mechanism in diseases associated with this gene (Bonaglia et al., 2015 and see below). Furthermore, microdeletions involving the *PURA* gene are also associated with a form of dominant

epileptic encephalopathy (Hosoki et al., 2012; Shimojima et al., 2011). In most of the cases, these are de novo mutations with apparently complete penetrance.

Additionally, the variant fulfills following ACMG criteria for pathogenesis: (a) PVS1 (pathogenic very strong), since the deletion is a null variant in a gene for which LoF is a known mechanism of disease. According to the Varsome database (Kopanos et al., 2018) *PURA* gene has 70 pathogenic LoF variants and a significant LOF Z-Score of 2.77, associated with Mental retardation, autosomal dominant. (b) PS2 (pathogenic strong), since it is a de novo mutation (both maternally and paternally confirmed) in a patient with the disease and no family history. (c) PM2 (pathogenic moderate), since the variant has not been previously reported in patients nor it has been described in GnomAD and/or other large-scale population studies involving healthy individuals (the region has a good coverage in GnomAD (>25x) according to Varsome analysis). (iv) PP3 (supporting evidence of pathogenicity), since an in silico score (GERP) predicted the variant as pathogenic (and no score predicted it as benign). Hence, all together the variant classifies as pathogenic according to ACMG.

There are several reported frameshift mutations in nearby codons (40, 46, 54, 59, and 83) as well as nonsense mutations (e.g., p. Gln55Ter, see <https://www.ncbi.nlm.nih.gov/clinvar/variation/426145/>) in that region of the protein, all of them reported as pathogenic. Other nonsense variants downstream in the protein (p. Gln163*, p. Gln186*, and p. Tyr261*) have also been reported as pathogenic (see: <https://www.ncbi.nlm.nih.gov/clinvar/variation/582890/>).

Finally, segregation analysis in the family (parents) supports a de novo mutation hypothesis which is consistent with both parents and half-siblings being healthy. In addition, no other variants in the *PURA* gene of less than 1% population frequency have been detected in this patient.

There are several patients reported with gene disrupting mutations which likely cause haploinsufficiency (nonsense and frameshift indels). In a larger series of patients reported so far (Reijnders et al., 2018 and Lee et al., 2018), no significant differences in clinical severity between mutation classes was found. As expected, in our patient, whose mutation causes a complete loss of function of the affected allele, whether caused by protein primary structure disruption or/and truncation or nonsense-mediated mRNA decay, the clinical presentation is severe with serious defect in motor skills, absence of language, respiratory problems, and intense epileptic activity.

Therefore, we think that the variant we found is a strong candidate pathogenic mutation probably explaining the patients' phenotype marked by severe central neurological dysfunction.

ACKNOWLEDGMENTS

Funded by URUGENOMES Project IP011-17/CEI/LC/MB.

CONFLICT OF INTERESTS

Authors declare no conflict of interests.

DATA AVAILABILITY STATEMENT

WGS is available upon request.

ORCID

Lucía Spangenberg  <https://orcid.org/0000-0001-5124-9250>

Martín Graña  <https://orcid.org/0000-0003-1182-8236>

Hugo Naya  <https://orcid.org/0000-0001-6982-4399>

REFERENCES

- Andrews. (2010). FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Bonaglia, M. C., Zanotta, N., Giorda, R., Dangelo, G., & Zucca, D. (2015). Long-term follow-up of a patient with 5q31.3 microdeletion syndrome and the smallest de novo 5q31.2q31.3 deletion involving PURA. *Molecular Cytogenetics*, 8, 89.
- Brogna, S., & Wen, J. (2009). Nonsense-mediated mRNA decay (NMD) mechanisms. *Nature Structural & Molecular Biology*, 16, 107.
- Calabrese, C., Simone, D., Diroma, M. A., Santorsola, M., Gutta, C., Gasparre, G., Picardi, E., Pesole, G., & Attimonelli, M. (2012). MToolBox: A highly automated pipeline for heteroplasmy annotation and prioritization analysis of human mitochondrial variants in high-throughput sequencing. *Bioinformatics*, 30(21), 3115–3117.
- Clark, M. M., Stark, Z., Farnaes, L., Tan, T. Y., White, S. M., Dimmock, D., & Kingsmore, S. F. (2018). Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ Genomic Medicine*, 3, 16.
- Daniel, J. (2018). PURA, the gene encoding Pur-alpha, member of an ancient nucleic acid-binding protein family with mammalian neurological functions. *Gene*, 643, 133–143.
- European Commission – European Commission. (2020). Rare diseases. https://ec.europa.eu/info/research-and-innovation/research-area/health/rare-diseases_en
- Hokkanen, S., Feldmann, H. M., Ding, H., Jung, C. K., Bojarski, L., Renner-Müller, I., Schüller, U., Kretzschmar, H., Wolf, E., & Herms, J. (2012). Lack of Pur-alpha alters postnatal brain development and causes megalencephaly. *Human Molecular Genetics*, 21, 473–484.
- Hosoki, K., Ohta, T., Natsume, J., Imai, S., Okumura, A., Matsui, T., Harada, N., Bacino, C. A., Scaglia, F., Jones, J. Y., & Niikawa, N. (2012). Clinical phenotype and candidate genes for the 5q31.3 microdeletion syndrome. *American Journal of Medical Genetics Part A*, 158, 1891–1896.
- Hunt, D., Leventer, R. J., Simons, C., Taft, R., Swoboda, K. J., Gawnecain, M., Magee, A. C., Turnpenny, P. D., & Baralle, D. Whole exome sequencing in family trios reveals de novo mutations in PURA as a cause of severe neurodevelopmental delay and learning disability. *Journal of Medical Genetics*, 51, 806–813.
- Johnson, E. M., Daniel, D. C., & Gordon, J. (2013). The pur protein family: Genetic and structural features in development and disease. *Journal of Cellular Physiology*, 228, 930–937.
- Karczewski, K. J., Francioli, L. C., & 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/nature25262>
- Khalili, D. V., Muralidharan, G., Darbinian, O., Meier, J., Daniel, K., Amini, S., & Gordon, J. (2003). Pur-alpha is essential for postnatal brain development and developmentally coupled cellular proliferation as revealed by genetic inactivation in the mouse. *Molecular and Cellular Biology*, 23, 6857–6875.
- Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C. E., Albarca Aguilera, M., Meyer, R., & Massouras, A. (2018). VarSome: the human genomic variant search engine. *Bioinformatics*, 35(11):1978. <https://doi.org/10.1093/bioinformatics/bty897>
- Lalani, S. R., Zhang, J., Schaaf, C. P., Brown, C. W., Magoulas, P., Tsai, A. C., El-Gharbawy, A., Wierenga, K. J., Bartholomew, D., Fong, C. T., Barbaro-Dieber, T., Kukulich, M. K., Burrage, L. C., Austin, E., Keller, K., Pastore, M., Fernandez, F., Lotze, T., Wilfong, A., ... Xia, F. (2014). Mutations in PURA cause profound neonatal hypotonia, seizures, and encephalopathy in 5q31.3 microdeletion syndrome. *American Journal of Human Genetics*, 95(5), 579–83. <https://doi.org/10.1016/j.ajhg.2014.09.014>. Epub 2014 Oct 16. PMID: 25439098; PMCID: PMC4225583.
- Landrum, L., Benson, B., Chao, C., Gu, H., Hoffman, J., Karapetyan, K., Liu, M., Malheiro, A., McDaniel, K., Ovetsky, M., Riley, G., Zhou, G., Holmes, J. B., Kattman, B. L., & Maglott, D. R. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Research*, 46(D1), D1062–D1067. <https://doi.org/10.1093/nar/gkx1153>
- Lee, B. H., Reijnders, M. R. F., Abubakare, O., Tuttle, E., Lape, B., Minks, K. Q., Stodgell, C., Bennetto, L., Kwon, J., Fong, C. T., & Gripp, K. W. (2018). Expanding the neurodevelopmental phenotype of PURA syndrome. *American Journal of Medical Genetics. Part A*, 176(1), 56–67. <https://doi.org/10.1002/ajmg.a.38521>
- Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., O'Donnell-Luria, A. H., Ware, J. S., Hill, A. J., Cummings, B. B., & Tukiainen, T. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, 536, 285–291.
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, 25(14), 1754–1760.
- Martens, G. J., & van Loo, K. M. (2007). Genetic and environmental factors in complex neurodevelopmental disorders. *Current Genomics*, 8, 429–444.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The genome analysis toolkit: A mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), 1297–1303.
- Pauline, H. (2003). SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Research*, 31, 3812–3814.
- Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., & Finn, R. D. (2018). HMMER web server: 2018 update. *Nucleic Acids Research*, 46(W1), W200–W204. <https://doi.org/10.1093/nar/gky448>
- Reijnders, M. R. F., Janowski, R., Alvi, M., Self, J. E., van Essen, T. J., Vreeburg, M., Rouhl, R. P. W., Stevens, S. J. C., Stegmann, A. P. A., Schieving, J., Pfundt, R., van Dijk, K., Smeets, E., Stumpel, C. T. R. M., Bok, L. A., Cobben, J. M., Engelen, M., Mansour, S., Whiteford, M., ... Baralle, D. (2018). PURA syndrome: Clinical delineation and genotype-phenotype study in 32 individuals with review of published literature. *Journal*

- of *Medical Genetics*, 55(2), 104–113. <https://doi.org/10.1136/jmedgenet-2017-104946>
- Shimajima, K., Isidor, B., Le Caignec, C., Kondo, A., Sakata, S., Ohno, K., & Yamamoto, T. (2011). A new microdeletion syndrome of 5q31.3 characterized by severe developmental delays, distinctive facial features, and delayed myelination. *American Journal of Medical Genetics. Part A*, 155A, 732–736.
- Söding, J. (2005). Protein homology detection by HMM–HMM comparison. *Bioinformatics*, 21, 951–960.
- Stavropoulos, D. J., Merico, D., Jobling, R., Bowdin, S., Monfared, N., Thiruvahindrapuram, B., Nalpathamkalam, T., Pellicchia, G., Yuen, R. K., Szego, M. J., & Hayeems, R. Z. (2016). Whole-genome sequencing expands diagnostic utility and improves clinical management in paediatric medicine. *NPJ Genomic Medicine*, 1(1):1-9.
- Tanaka, A. J., Bai, R., Cho, M. T., Anyane-Yeboa, K., Ahimaz, P., Wilson, A. L., Kendall, F., Hay, B., Moss, T., Nardini, M., & Bauer, M. (2015). De novo mutations in PURA are associated with hypotonia and developmental delay. *Molecular Case Studies*, 1, a000356.
- The 1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature*, 526, 68–74.
- Wang, K., Li, M., & Hakonarson, H. (2010). Annovar: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, 38(16):e164.
- Weber, J., Bao, H., Hartmüller, C., Wang, Z., Windhager, A., Janowski, R., Madl, T., Jin, P., & Niessing, D. (2016). Structural basis of nucleic-acid recognition and double-strand unwinding by the essential neuronal protein Pur-alpha. *Elife*, 5, 11297.
- Yang, Y., Muzny, D. M., Xia, F., Niu, Z., Person, R., Ding, Y., Ward, P., Braxton, A., Wang, M., Buhay, C., & Veeraghavan, N. (2014). Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA*, 312(18), 1870–1879.

How to cite this article: Spangenberg L, Guecaimburú R, Tapié A, et al. Novel frameshift mutation in PURA gene causes severe encephalopathy of unclear cause. *Mol Genet Genomic Med*. 2021;9:e1622. <https://doi.org/10.1002/mgg3.1622>