Expansion of the neurodevelopmental phenotype of individuals with EEF1A2 variants and genotype-phenotype study

Expansion of the neurodevelopmental phenotype of individuals with *EEF1A2* variants and genotype-phenotype study.

Alix Paulet[1], Cavan Benett-Ness[2], Faustine Ageorges[1], Detlef Trost[3], Andrew Green[4], David Goudie[5], Rosalyn Jewell[6], Minna Kraatari-Tiri[7,8], Juliette PIARD[9], Christine Coubes[10], Wayne Lam[11], Sally Ann Lynch[12], Groeschel Samuel[13], Francis Ramond[14], Joël Fluss[15], Christina Fagerberg[16], Charlotte Brasch-Andersen[17], Konstantinos Varvagiannis[18], Tjitske Kleefstra[19,20,21,22], Bénédicte Gérard[23], Mélanie Fradin[24], Antonio Vitobello[25], Romano Tenconi[26], Anne-Sophie Denommé-Pichon[27,28], Aline Vincent[29], Tobias Haack[30], Joseph Marsh[31], Lone Walentin Laulund[32], Mona Grimmel[30], Angelika Riess[30], Elke de Boer[19,20,21], Sergio Padilla Lopez[33], Somayeh Bakhtiari[33], Michael Kruer[33], Jonathan Levy[1], Alain Verloes[1], Catherine Abbott[2], Lyse Ruaud[1].

Correspondence to: Alix Paulet, Département de Génétique, Hôpital Robert-Debré, Paris, France.

Email: alix.paulet@hotmail.fr

1. Département de Génétique, Hôpital Robert-Debré, Paris, France.

2. Centre for Genomic and Experimental Medicine and Simons Initiative for the Developing Brain, Institute of Genetics and Cancer, Edinburgh, Scotland, UK.

3. Laboratoire Cerba, Saint-Ouen l'Aumône, France.

4. UCD School of Medicine and Medical Science Consultant in Clinical Genetics, Dublin, Ireland.

5. Regional Genetics Service, NHS Tayside, Dundee, Scotland, UK.

6. Yorkshire Regional Genetics Service, Leeds Teaching Hospitals NHS Trust, Leeds, England, UK.

7. Department of Clinical Genetics, Research unit of Clinical Medicine, Medical Research Center Oulu, Finland.

8. Oulu University Hospital and University of Oulu, Finland.

9. Centre de Génétique Humaine, CHU Besançon, Besançon, France.

10. Service de Génétique Médicale, CHU de Montpellier, Montpellier, France.

11. South-East of Scotland Clinical Genetics Service, General Hospital, Edinburgh, Scotland, UK.

12. Clinical Genetics, Children's Health Ireland, Dublin, Ireland.

13. Department of Neuropediatrics, University Children's Hospital, Tuebingen, Germany.

14. Service de Génétique, CHU Saint-Etienne - Hôpital Nord, Saint-Etienne, France.

15. University Hospitals of Geneva, Geneva, Switzerland.

16. Department of Clinical Genetics, Odense University Hospital, Odense, Denmark.

17. Human Genetik, Syddansk Universitet, Odense, Denmark.

18. Service de Médecine Génétique, Hôpitaux universitaires de Genève, Geneva, Switzerland.

19. Department of Clinical Genetics, Erasmus MC University Medical Center, Rotterdam, The Netherlands.

20. Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands.
21. Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands.
22. Center of Excellence for Neuropsychiatry, Vincent van Gogh Institute for Psychiatry, Venray, The Netherlands.
23. Molecular Biology, Nouvel Hôpital Civil, Strasbourg, France.
24. Service de Génétique Médicale, Hôpital Sud, CHU de Rennes.
25. UMR-Inserm, Génétique des Anomalies du développement, Université de Bourgogne Franche-Comté, Dijon.
26. Servizio di Genetica Medica, Dipartimento di Pediatra, Padova, Italia.
27. Unité Fonctionnelle Innovation en Diagnostic génomique des maladies rares, FHU-TRANSLAD, CHU Dijon Bourgogne, Dijon, France.
28. UMR1231 GAD, Inserm - Université Bourgogne-Franche Comté, Dijon, France.
29. Génétique Médicale, CHU de Caen, Caen France.
30. Institute of Medical Genetics and Applied Genomics, University of Tübingen, Tübingen, Germany.
31. MRC Human Genetics Unit, Western General Hospital, University of Edinburgh, Edinburgh, Scotland, UK.
32. H C Andersen Children's Hospital, Odense University Hospital, Odense, Denmark.
33. Pediatric Movement Disorders Program, Division of Pediatric Neurology, Barrow Neurological Institute, Phoenix Children's Hospital, Phoenix, Arizona, USA.

Keywords: *EEF1A2*, Intellectual disability, regression, epilepsy, neurodevelopment.

Abstract

Translation elongation factor eEF1A2 constitutes the alpha subunit of the elongation factor-1 complex, responsible for the enzymatic binding of aminoacyl-tRNA to the ribosome. Since 2012, 21 pathogenic missense variants affecting *EEF1A2* have been described in 42 individuals with a severe neurodevelopmental phenotype including epileptic encephalopathy and moderate to profound intellectual disability (ID), with neurological regression in some patients.

Through international collaborative call, we collected 26 patients with *EEF1A2* variants and compared them to the literature.

Our cohort shows a significantly milder phenotype. 83% of the patients are walking (vs. 29% in the literature), and 84% of the patients have language skills (vs. 15%). Three of our patients do not have ID. Epilepsy is present in 63% (vs. 93%). Neurological examination shows a less severe phenotype with significantly less hypotonia (58% vs. 96%), and pyramidal signs (24% vs. 68%). Cognitive regression was noted in 4% (vs. 56% in the literature).

Among individuals over 10 years, 56% disclosed neurocognitive regression, with a mean age of onset at 2 years.

We describe 8 novel missense variants of *EEF1A2*. Modelling of the different amino-acid sites shows that the variants associated with a severe phenotype, and the majority of those associated with a moderate phenotype, cluster within the switch II region of the protein and thus may affect GTP exchange. In contrast, variants associated with milder phenotypes may impact secondary functions such as actin bundling. We report the largest cohort of individuals with *EEF1A2* variants thus far, allowing us to expand the phenotype spectrum and reveal genotype-phenotype correlations.

Abbreviations:

ACMG: American College of Medical Genetics

ADHD: Attention Deficit and Hyperactivity Disorder

ASD: Autism Spectrum Disorder

CADD: Combined Annotation Dependent Depletion

CLB: Clobazam

FISQ : Full-Scale Intelligence Quotient

ID: Intellectual Disability

IQ: Intellectual Quotient

KETO DIET: Ketogenic Diet

LAM: Lamotrigin

LORAZEP: Lorazepam

MICROPAK: Micropakine

MRI: Magnetic Resonance Imaging

OFC: Occipito-Frontal Circumference

SD: Standard Deviation

TIQ: Total Intellectual Quotient

VA: Valproic Acid

WISC: Wechsler Intelligence Scale for Children

128 WM and GM: White Matter And Grey Matter

129 WPPSI-IV: Wechsler Preschool and Primary Scale of Intelligence

Introduction

The eEF1 family of eukaryotic elongation factor genes, which comprises the two paralog eEF1A proteins eEF1A1 and eEF1A2 and the 3 subunits of the eEF1B complex (eEF1Bα, eEF1Bβ, and eEF1Bγ) encodes integral components of the translation elongation factor complexes whose function is delivery of aminoacyl tRNA to ribosome during the elongation step of protein synthesis[1]. The two eEF1A proteins, the second most abundant protein in the cell[2], share 92% identity. eEF1A binds aa-tRNAs in a GTP-dependent manner, relying on its cognate guanine exchange factor (GEF), eEF1B, to recycle the inactive eEF1A-GDP complex into the active GTP-bound form. *EEF1A1* gene is expressed almost ubiquitously. *EEF1A2* is expressed mainly in muscle (including cardiac muscle) and in neurons[3,4]. During development, eEF1A1 is down-regulated in muscle and neurons and is undetectable in mouse muscle by 3 weeks post-natal[5,6].

No pathogenic *EEF1A1* variants have been described, presumably because they would not be compatible with life[7]. In contrast, in 1972, a deletion of 15.8 kb that abolishes expression of eEF1A2 (MIM_ 01958) was discovered in mice that developed motor neuron degeneration, muscle atrophy, gait abnormalities and then died by four weeks[3,5]. A trio-based exome in 2012 revealed a *EEF1A2* variant in a patient with early onset epilepsy, severely delayed psychomotor development, and autistic behavior[8]. Subsequent individual reports and a series of 14 patients[9] allowed to further delineate the neurodevelopmental phenotype, combining moderate to severe

155 Intellectual Disability (ID), epilepsy, Autism Spectrum Disorder (ASD),
156 and sleep disorders neurodegeneration and movement disorders[9-13].

157 Since 2012, 21 *EEF1A2* variants have been reported[8-27]. All variants
158 are missense. In one pedigree, a variant was inherited from a parent
159 with less than 25% of mosaicism[9]. Other variants were *de novo*. To
160 date, no genotype-phenotype correlation has been studied.

161 Through an international collaboration, we collected data on 26
162 unreported *EEF1A2* patients, the largest cohort to date, and
163 investigated possible genotype-phenotype correlations.

Material And Methods:

Individual ascertainment:

Between December 2020 and December 2021, 26 individuals with variants in *EEF1A2* were identified through European Reference Network ERN-ITHACA (https://ern-ithaca.eu/) using its collaborative call system. Patients were evaluated by a geneticist. Written informed consents for DNA and data analyses were obtained from individuals or their legal guardians.

Cognitive assessment:

Out of the 26 patients included in our study, we assessed cognitive abilities for 13 individuals (Table 1): six underwent a Wechsler Intelligence Scales for Children according to age (Scales IV and V: 6 years to 16 years 11 months old, or Wechsler Preschool and Primary Scale of Intelligence III or IV: 4 years to 7 years 3 months old) and 7 patients were evaluated by clinicians (Table S1).

Among 6 patients which performed Wechsler tests, only 4 of the 6 patients had a computable Total Intelligence Quotient (TIQ). For the remaining two, they had too heterogenous profiles to allow TIQ calculation. We therefore utilized the available indices data (VCI, FRI, VSI, WMI, PSI) to estimate their intellectual levels.

For the 7 patients evaluated by clinicians, cognitive assessments were provided but we did not have quantitative data (Table S1).

Patients were divided into 4 categories such as “not ID” for TIQ ≥70, “mild ID” for 50≤IQ<70, “moderate ID” for 35≤IQ<50, and “severe ID” for IQ<35, according to their IQ scores especially TIQ. In total, we classified 13 patients according to their level of intellectual efficiency (Table 1).

We excluded 13 individuals of cognitive analysis because they were too young at the last follow-up appointment or they have been lost to follow-up.

To categorize individuals from the literature when Weschler’s FSIQ was not available, we took into account the authors' assigned category ID for each individual or we used the information on individual developmental milestones (walking age, capacity for language, verbal abilities…) to estimate the developmental delay as “mild”, “moderate” or “severe” phenotype (Table S2). The individuals whose information was insufficient to be to be classified into these categories were not scored.

Genetic investigations:

DNA was extracted from the peripheral blood leukocytes of the patients and parents (whenever possible) using standard procedures.

In 24/26 patients, genotyping was performed by exome sequencing (ES) (single or trio) using routine methods. Confirmation and segregation of variants in single ES were carried out by Sanger sequencing. Variant prioritization was conducted according to the transmission mode (de novo, autosomal recessive and X-linked), and the frequency of the variants in the gnomAD database. Those variants were classified according to American College of Medical Genetics (ACMG) (ACMG and Combined Annotation Dependent Depletion (CADD) score in Supplementary data, Table S3). There were no other pathogenic variants in ClinVar and HGMD or loss-of-function variants which could explain the patient phenotypes. Two variants were identified on a NGS panel of 119 ID genes using standard NGS procedures.

Protein Modelling:

The structures of eEF1A2*GDP (4C0S), eEF1Ay*eEF1Bα (pdb: 1IJF), aEF1A*GTP (pdb: 3AGJ), and EF-TU*EF-TS (pdb:1EFU) were used for analysis using Pymol (The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.) and CCP4MG[28]. Structure-based variant predictions made using Missense 3D[29] and dimer predictions using FoldX[30]. Briefly, eEF1A2 (pdb: 4C0S) was repaired, then the protein

stability was estimated for each amino acid change in both monomeric and the dimeric eEF1A2.

Statistical analysis

We compared our patients to the previous reported ones. We used Student Test with bilateral hypothesis, have tolerated a risk of error of less than 5% (*p-value* <0.05).

Literature review:

PubMed was searched for peer-reviewed articles published in English using the following keywords: « *EEF1A2* », « epileptic encephalopathy gene », « *EEF1A* » «phenotype of *EEF1A2* » « *EEF1A2* with no intellectual disability » « epilepsy gene ».

Results

In our series, the mean age at last examination is 10.67 years vs. 9.6 years for literature patients (Table S4). The mean Occipito-frontal circumference (OFC) is -0.67 SD (vs. -1.33 SD) (Table S4). Most patients can walk on independently. The mean age for acquiring walking is 30 months. For the 4 individuals in the literature this mean is 39 months, *p-value*= 0.41 (Table S4).

Among 18 individuals aged more than 2 years, 83% are verbal (Figure 1) with a mean age of first words at 29 months (Table S4) (data not available for published cases). The average age at onset of epilepsy is 3.5 years (Table S4) (data not available for previously reported patients).

Neuropsychologist assessments were performed in six patients (Table 1). For the 4 patients for whom a TIQ could be calculated, the Full-Scale Intelligence Quotient (FSIQ) ranged from 40 to 77 (mean= 60.2; median= 58.5) (Table S1).

Patient 1 has a FSIQ in the low normal range (77) and the two others (patients 6 and 19) had several indexes > 70 despite their heterogeneous profile did not allow to compute FSIQ. Patient 19 has a FSIQ of 69, but with low normal verbal comprehension index (76) and fluid reasoning (71) (Table S1).

In our cohort of 13 for whom we have developmental information, 4 patients have a phenotype described as severe (31%) compared to

those from the literature (76%) *p-value*=0.008 (-0.77;-0.13) (Table 1). Among published patients, 5/33 (15%) had profound ID (Table 2), whereas there was no patient classified as having profound ID in our cohort. There was no difference observed concerning mild and moderate ID (Table 1).

The individuals of this series are more ambulant than previous reported ones: 82% (18/22) vs. 29% (10/34), *p-value*=0.00004 (0.29; 0.76). They are more likely to have acquired language (83% vs. 14% *p-value*=10E-7 (0.47;0.91)) (Figure 1).

In our cohort, less patients had hypotonia and pyramidal syndrome (Figure 1). There was no difference in the incidence of ataxia and ASD. A majority of our patients have sleep disorders (11/24), with limited effectiveness of melatonin treatment (Table S1). Half of our patients (53%-9/17) had ADHD (Table S1). More than a half of our patients (14/25) have gross motor issues with mainly unstable walk, and 77% (17/22) have fine motor issues particularly concerning coordination (Table S1).

We also compared our patients to the previously described ones regarding several items. The percentage of patients with epilepsy refractory to therapy, the percentage with epileptic encephalopathy were compared so as our series brain MRI features. The results are presented in Table 2.

We report 8 new *EEF1A2* variants (T24M, R96H, D97N, T104R, G356S, D362N, P420L, V437F) with detailed phenotyping. Phenotypes

observed varies from mild to severe (Figure 2a). The patients with variants T24M and R96C have no ID but are symptomatic: have epilepsy (R96C) or speech delay with firsts sentences at age 6 (T24M). For patients with P420L and V437F we could not document ID because of the lack of psychometric evaluation (Table S1).

Molecular Modelling

We explored the impact of the variants through 3D-modelling of eEF1A2, comparing the GDP-bound form of eEF1A2 (pdb:4C0S) with eEF1Ay*eEF1Bα (pdb:1IJF), aEF1A*GTP (pdb:3AGJ), and EF-TU*EF-TS (pdb:1EFU). All variants have CADD scores above 20 (Table S3).

Mapping of the variants onto the structure of eEF1A2 (Figure 2b) showed that several of the variants cluster around the switch II region, a flexible motif key for GTP hydrolysis and GEF-mediated GDP dissociation. R96C, R96H, and D97N are found in a loop in the Switch II region, whilst T104R, R381W, and V437F are situated nearby. These variations are anticipated to contribute towards switch II destabilization and interfere with GTP recycling, whilst another variant, T24M, is located at the GTP-binding pocket, and is predicted to directly disrupt nucleotide binding (Supplementary modelling in Supplemental data).

Further from the GDP-binding or switch regions, several variations (E297K, G356S, D362N, and P420L) are located near the actin binding region (Figure 2b) and are anticipated to disrupt actin-related functions (Supplementary modelling). Three of these variants, E297K, D362N and P420L are located near the tRNA-binding site, and may affect eEF1A2 interactions with aminoacyl-tRNA. Given the overlap between the binding sites for actin and aa-tRNAs and the eEF1B-binding region, some of these variants may also impact eEF1B binding. Additionally, the presence of variants T104R and D362N on the dimer interface, and the change in amino acid charge, suggest

313 they may influence eEF1A2 dimer formation (Supplementary
314 modelling).

Discussion

Widening *EEF1A2*-related spectrum

Our results show that patients with *EEF1A2* variants can have efficient cognitive abilities. Indeed, in our cohort, 3 individuals do not have an intellectual disability (IQ > 70), while in the literature, no patients are reported with preserved intellectual efficiency. In addition, within our cohort, the severity of ID is lower, with a majority of patients with mild to moderate ID.

In terms of development, the majority of our individuals have access to language and walking, contrary to what has already been described.

Taken together, the phenotypic spectrum associated with *EEF1A2* variants is wider than previously suggested. All previously reported patients had intellectual disabilities and the majority (93%) were epileptic and nonverbal. We have shown that individuals with *EEF1A2* variants may have milder phenotypes and that nearly half of them do not have epilepsy at time of report. Epilepsy can be late-onset, as illustrated by patient 7 (Ref. 20) with the D252H variant who did not develop epilepsy until age 8 and the average age in the cohort was 10.7 years.

Young individuals are difficult to assess using the WPPSI-IV or WISC scales. There are 13 individuals who have received cognitive evaluation but only 4/13 have a valid TIQ because profiles are too heterogenous. For this study, we chose to take the heterogenous

profiles into account because even if partial Weschler 'scales remain an objective way to evaluate individuals' cognitive capacities.

To date, the only reported patient of autosomal dominant inherited *EEF1A2* variant was from a mother with <25% of mosaicism[9] but there was no clinical information available. We newly report 2 patients (patient 6 (whose father is patient 7) and patient 20) with inherited variants from an asymptomatic parent (in terms of neurodevelopmental disorder) which suggests an incomplete penetrance of these variants (G356S and R96C). G356S is not reported in the database and the position of the residue is highly conserved. For G356S, the asymptomatic parent has a history of irregular heartbeat, and he has at least 3 relatives who had seizures (the relatives have not been tested for the variant yet). One individual (patient 6) has the variant R96C inherited from his affected father (patient 7), this variant had been shown to be associated with (Genetic Generalized epilepsies) GGE[27]. No mosaicism detected in unaffected parents.

These two examples of inherited *EEF1A2* variants strongly suggest incomplete penetrance. Although only 3 patients (including our 2) are reported so far, incomplete penetrance should be considered in *EEF1A2* variants. More individuals are needed to confirm this hypothesis and we recommend to test proband's parents even if there are apparently not affected. This incomplete penetrance is relevant to genetic counselling.

We can notice a fairly variability of expression in patients with the same variation. For example, in our series, patients 15, 16 and 17 have the same E124K variant already reported by Kaur et al[15]. Whereas patients 15,16 and Kaur et al patient acquired independent walk before 2 years with good language abilities and epilepsy, our patient 17 acquired walk at 4 years and had no epilepsy.

We also report 2 patients (11 and 5) of de novo *EEF1A2* variant mosaicism which did not cause milder phenotypes. Patient 11 presents epileptic encephalopathy and has TIQ of 52 at 3y. Patient 5 can speak in sentences. Nevertheless, they have similarities: they both have white and grey matter abnormalities on brain MRI and focal epilepsy, controlled with antiepileptic treatment. Concerning patient 5, the variation D91N has been reported three times in the Literature in patients with severe to profound ID and early onset epilepsy[9,15,18], whereas our patient is obviously less severe affected than the three others. This mosaicism (estimated to 23%) could have caused bias in comparison because this milder phenotype is possibly only a result of mosaicism. So, mosaicism state should be taken into account while considering an *EFF1A2* variant, although the impossibility to estimate percentage of mosaicism in brain makes the phenotype difficult to be predicted.

Ten patients[9-10,12,22-23] have been reported to have regression in childhood. In our cohort, only patient 22 (R381W) has shown regressive traits which began in her third decade (but she is also the only individual to have reached this age). She has started to show less

388 interest in activities she used to like such as writing, swimming and
389 she lost some abilities. This individual suggest that regression may
390 appear later than we thought while reviewing the literature patients
391 (where most of patients showed regression in infancy[9]). It's clear that
392 patients must be followed up and supported to prevent
393 complications during time.

Molecular insights

Plotting the patients' variants and those previously described on the surface of the eEF1A2 protein (Figure 2b) suggests that variants, resulting in severe ID are generally clustered around the switch I and switch II regions, or nucleotide/GEF binding sites for GTP and GTP exchange factor eEF1B. The growing cluster of variants around the switch II region or GDP-binding site (Figure 2b) suggests that disruption of GDP-binding and GEF-induced GDP dissociation may be a key mechanism for the NDD-causing *EEF1A2* missense variants, adding to the evidence from Carvill et al[9]. Four of the 19 variants classified as severe directly coincide with defined eEF1B binding sites (Figure 2b).

Many of the milder variants map further away but could affect tRNA or actin binding (Figure 2b). There is evidence that E295K, the equivalent E297K mutation in yeast, affects translational fidelity. Variants affecting translational fidelity might be less likely to affect neurodevelopment, but would be anticipated to lead to neurodegeneration, which is observed in a subset of individuals with *EEF1A2* variants. There is also the possibility that these variants are not pathogenic and the cause of the individual phenotype was misunderstood.

Conclusion

We have described a cohort of individuals with a less severe phenotype, though they share characteristics such as developmental delay (especially speech delay), mild to severe intellectual disability, ASD, ADHD, early onset epilepsy, hypotonia, ataxia and sleep disorder which are concordant with the features related to variants in this gene in the literature. We suggest the existence of incomplete penetrance of certain variants which was not described so far with *EEF1A2*. Our series illustrate how the evolution of diagnostic strategies may lead to redefine the phenotypic spectrum of known genes that have been initially reported with a homogeneous and usually “severe” phenotype. There was probably an ascertainment bias in older patients that were more likely to be reported because they were severe, and were perhaps selected on this basis. The widespread adoption of the whole genome and whole exome sequencing, which results in an agnostic pan-genomic evaluation lead to the diagnosis of patients that would not be sent to targeted panel diagnosis. The first genotype-phenotype correlations are emerging, but new patients will be necessary to confirm these correlations.

To conclude, we expanded the phenotype spectrum and described new *EEF1A2* variants.

All anonymized data and related documentation from this study are available on reasonable request.

Declaration to public database: All variants are reported and annotated in ClinVar website (accession number SCV004171535-SCV004171551): https://www.ncbi.nlm.nih.gov/clinvar/

Bibliographie

1. Kahns S, Lund A, Kristensen P, et al. The elongation factor 1 A-2 isoform from rabbit: cloning of the cDNA and characterization of the protein. *Nucleic Acids Res*. 1998;26(8):1884-1890. doi:10.1093/nar/26.8.1884

2. Abbott CM, Newbery HJ, Squires CE, Brownstein D, Griffiths LA, Soares DC. eEF1A2 and neuronal degeneration. *Biochem Soc Trans*. 2009;37(Pt 6):1293-1297. doi:10.1042/BST0371293

3. Newbery HJ, Loh DH, O'Donoghue JE, et al. Translation elongation factor eEF1A2 is essential for post-weaning survival in mice. *J Biol Chem*. 2007;282(39):28951-28959. doi:10.1074/jbc.M703962200

4. Knudsen SM, Frydenberg J, Clark BF, Leffers H. Tissue-dependent variation in the expression of elongation factor-1 alpha isoforms: isolation and characterisation of a cDNA encoding a novel variant of human elongation-factor 1 alpha. *Eur J Biochem*. 1993;215(3):549-554. doi:10.1111/j.1432-1033.1993.tb18064.x

5. Chambers DM, Peters J, Abbott CM. The lethal mutation of the mouse wasted (wst) is a deletion that abolishes expression of a

tissue-specific isoform of translation elongation factor 1alpha, encoded by the Eef1a2 gene. Proc Natl Acad Sci U S A. 1998 Apr 14;95(8):4463-8. doi: 10.1073/pnas.95.8.4463. PMID: 9539760; PMCID: PMC22512.

6. Khalyfa A, Carlson BM, Dedkov EI, Wang E. Changes in protein levels of elongation factors, eEF1A-1 and eEF1A-2/S1, in long-term denervated rat muscle. Restor Neurol Neurosci. 2003;21(1-2):47-53. PMID: 12808202.

7. McLachlan F, Sires AM, Abbott CM. The role of translation elongation factor eEF1 subunits in neurodevelopmental disorders. *Hum Mutat*. 2019;40(2):131-141. doi:10.1002/humu.23677

8. De Ligt J, Willemsen MH, van Bon BWM, et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med*. 2012;367(20):1921-1929. doi:10.1056/NEJMoa1206524

9. Carvill GL, Helbig KL, Myers CT, et al. Damaging de novo missense variants in EEF1A2 lead to a developmental and degenerative epileptic-dyskinetic encephalopathy. *Hum Mutat*. 2020;41(7):1263-1279. doi:10.1002/humu.24015

10. Veeramah KR, Johnstone L, Karafet TM, et al. Exome sequencing reveals new causal mutations in children with epileptic encephalopathies. *Epilepsia*. 2013;54(7):1270-1281. doi:10.1111/epi.12201

11. Long K, Wang H, Song Z, Yin X, Wang Y. EEF1A2 mutations in epileptic encephalopathy/intellectual disability: Understanding the potential mechanism of phenotypic variation. *Epilepsy Behav EB*. 2020;105:106955. doi:10.1016/j.yebeh.2020.106955

12. Inui T, Kobayashi S, Ashikari Y, et al. Two cases of early-onset myoclonic seizures with continuous parietal delta activity caused by EEF1A2 mutations. *Brain Dev*. 2016;38(5):520-524. doi:10.1016/j.braindev.2015.11.003

13. Cao S, Smith LL, Padilla-Lopez SR, et al. Homozygous EEF1A2 mutation causes dilated cardiomyopathy, failure to thrive, global developmental delay, epilepsy and early death. *Hum Mol Genet*. 2017;26(18):3545-3552. doi:10.1093/hmg/ddx239

14. Kaneko M, Rosser T, Raca G. Dilated cardiomyopathy in a patient with autosomal dominant EEF1A2-related neurodevelopmental disorder. *Eur J Med Genet*. 2021;64(1):104121. doi:10.1016/j.ejmg.2020.104121

15. Kaur S, Van Bergen NJ, Gold WA, et al. Whole exome sequencing reveals a de novo missense variant in EEF1A2 in a Rett syndrome-like patient. *Clin Case Rep*. 2019;7(12):2476-2482. doi:10.1002/ccr3.2511

16. Zacher P, Mayer T, Brandhoff F, et al. The genetic landscape of intellectual disability and epilepsy in adults and the elderly: a systematic genetic work-up of 150 individuals. *Genet Med Off J*

Am Coll Med Genet. 2021;23(8):1492-1497. doi:10.1038/s41436-021-01153-6

17. Lopes F, Barbosa M, Ameur A, et al. Identification of novel genetic causes of Rett syndrome-like phenotypes. *J Med Genet*. 2016;53(3):190-199. doi:10.1136/jmedgenet-2015-103568

18. Lam WWK, Millichap JJ, Soares DC, et al. Novel de novo EEF1A2 missense mutations causing epilepsy and intellectual disability. *Mol Genet Genomic Med*. 2016;4(4):465-474. doi:10.1002/mgg3.219

19. de Kovel CGF, Brilstra EH, van Kempen MJA, et al. Targeted sequencing of 351 candidate genes for epileptic encephalopathy in a large cohort of patients. *Mol Genet Genomic Med*. 2016;4(5):568-580. doi:10.1002/mgg3.235

20. Nakajima J, Okamoto N, Tohyama J, et al. De novo EEF1A2 mutations in patients with characteristic facial features, intellectual disability, autistic behaviors and epilepsy. *Clin Genet*. 2015;87(4):356-361. doi:10.1111/cge.12394

21. Ostrander BEP, Butterfield RJ, Pedersen BS, et al. Whole-genome analysis for effective clinical diagnosis and gene discovery in early infantile epileptic encephalopathy. *NPJ Genomic Med*. 2018;3:22. doi:10.1038/s41525-018-0061-8

22. De Rinaldis M, Giorda R, Trabacca A. Mild epileptic phenotype associates with de novo eef1a2 mutation: Case report and

review. *Brain Dev*. 2020;42(1):77-82. doi:10.1016/j.braindev.2019.08.001

23. Lance EI, Kronenbuerger M, Cohen JS, Furmanski O, Singer HS, Fatemi A. Successful treatment of choreo-athetotic movements in a patient with an EEF1A2 gene variant. *SAGE Open Med Case Rep*. 2018;6:2050313X18807622. doi:10.1177/2050313X18807622

24. O'Roak BJ, Stessman HA, Boyle EA, et al. Recurrent de novo mutations implicate novel genes underlying simplex autism risk. *Nat Commun*. 2014;5:5595. doi:10.1038/ncomms6595

25. Helbig KL, Farwell Hagman KD, Shinde DN, et al. Diagnostic exome sequencing provides a molecular diagnosis for a significant proportion of patients with epilepsy. *Genet Med*. 2016;18(9):898-905. doi:10.1038/gim.2015.186

26. Vogt LM, Lorenzo M, B Prendergast D, Jobling R, Gill PJ. EEF1A2 pathogenic variant presenting in an infant with failure to thrive and frequent apneas requiring respiratory support. Am J Med Genet A. 2022 Oct;188(10):3106-3109. doi: 10.1002/ajmg.a.62932. Epub 2022 Aug 8. PMID: 35938194.

27. Epi25 Collaborative. Electronic address: jm4279@cumc.columbia.edu; Epi25 Collaborative. Sub-genic intolerance, ClinVar, and the epilepsies: A whole-exome sequencing study of 29,165 individuals. Am J Hum Genet. 2021 Jun 3;108(6):965-982. doi: 10.1016/j.ajhg.2021.04.009. Epub

2021 Apr 30. Erratum in: Am J Hum Genet. 2021 Oct 7;108(10):2024. PMID: 33932343; PMCID: PMC8206159.

28. Presenting your structures: the CCP4mg molecular-graphics software S. McNicholas, E. Potterton, K. S. Wilson and M. E. M. Noble Acta Cryst. (2011). D67, 386-394

29. Ittisoponpisan, S., Islam, S.A., Khanna, T., Alhuzimi, E., David, A. & Sternberg, M.J.E. (2019)_Can Predicted Protein 3D Structures Provide Reliable Insights into whether Missense Variants Are Disease Associated? J. Mol. Biol. 431, 2197-2212._https://doi.org/10.1016/j.jmb.2019.04.009

30. Van Durme J, Delgado J, Stricher F, Serrano L, Schymkowitz J and Rousseau F. A graphical interface for the FoldX forcefield. Bioinformatics. 2011 Jun 15;27(12):1711-2.

Figure Legends:

Table 1. Cognitive abilities of patients in our series compared to those previous reported.

This table shows comparison in terms of cognitive abilities between our patients and the previously reported ones. The 4 categories classification strategy has been explained in the Material and Method section. The number of patients are in parentheses. * means the comparison is statistically significant, p-value <0.05. CI : Confidence Interval.

Table 2. Cohort and literature comparison regarding epilepsy and brain MRI characteristics

This table shows comparison between our cohort and the previous reported cases regarding epilepsy and brain MRI features. *P-value* and CI 95% are mentioned. CI : Confidence Interval. * when significant.

Figure 1. Our series main features compared with the patients from the literature.

The figure 1 shows the proportion of our patients (in black) in comparison to those from the literature (featured in grey) according to the main criteria studied. Walking abilities is represented by the motor category. Concerning the speech abilities, “verbal” means the patient can speak and be understandable, “words” means he can express himself in words but not in sentences, “sentences” stands for the ability of the patient to make proper sentences. On the right

590 part of the diagram, the neurological features are shown withe the
591 presence of hypotonia, pyramidal syndrome, epilepsy and the
592 presence of regression. * Stands for statistically significance $p<0.05$.
593 *P-values* are framed at the top.

594 **Figure 2. Distribution of pathogenic *EEF1A2* variants.**

595 Figure 2a: distribution of variants from our cohort (up) and the
596 previous reported ones (down). Novel variants are underlined.
597 Variants are classified by their associated phenotype in terms of ID,
598 classification is as described as above with associated symbols (circle
599 for not deficient, triangle for mild ID, square for moderate ID and star
600 for severe). Blank when the variant is not clearly classified (for
601 example when associated with 2 ID categories). Figure 2b: variants
602 mapped onto the crystal structure of GDP-bound eEF1A2 (PDB:4C0S).
603 New variants are shown in black, with labels in bold, previously
604 described mutations are in white. T24M and V437F are buried. The
605 binding site of eEF1B is highlighted in white, and the GTP binding site
606 in dark grey.

607 **Table S1. Cohort's patients characteristics**

608 The table S1 shows the characteristics of our 26 patients ordered by
609 the position of their variation along the *EEF1A2* gene. Each column
610 represents a criterion of interest. "M" means Male, "F" means
611 Female. The ages at last examination are presented in years "y". The
612 ages at first steps achievement without falling and first
613 comprehensive words achievement are presented in months "m".
614 Concerning the languages abilities at last examination, classification

into 3 categories (absent/words/sentences) are presented, with details of language level. About the neuropsychological assessment column, Total Intellectual Quotients (TIQ) are presented when they can be calculated. In case the TIQ could not be calculated the different QI indicators are presented with their respective scores: VCI, Verbal Comprehension Index, VSI, Visual Spatial Index, FRI, Fluid Reasoning index, WMI, Working Memory Index, PSI, Processing Speed Index. The age of the tests were carried out are in parentheses ("NA" when non available), followed by the type of test performed: WISC-V or WPPSI-IV. For the therapy and efficacity column, the drugs used are in parentheses. Regarding the weight, height and Occipito-frontal circumference (OFC) columns; the results are presented in standard deviation (SD). The gross motor feature indicates walking disorder. The fine motor features entail coordination disorder. The proportion of patients are in parentheses.

Y: Yes // N: No // NA: Non available // NC: Non concerned //ADHD: Attention Deficit Hyperactivity Disorder // VA: Valproic Acid // LAM: Lamotrigin // CLB: Clobazam // Lorazep: Lorazepam // Micropak: micropakine // Keto diet: Ketogenic Diet // TIQ: Total Intellectual Quotient.

Table S2. Literature patients characteristics

The table S2 shows the characteristics of the 42 patients reported in literature. Each column represents a criterion of interest. The ages at last examination are presented in years. "y" stands for "years", "m" for "months". "M" means Male, "F" means Female. "NA": Non

available. “Y” stands for yes, “N” stands for no. Regarding the Occipito-frontal circumference (OFC) at examination column, the results are presented in SD. The proportion of patients for each category are in parentheses.

Table S3. ACMG classification and CADD Score

WES :Whole Exome sequencing; ACMG: American College of Medical Genetics; CADD: Combined Annotation Dependent Depletion. This table presents the classification according to ACMG guidelines and the CADD score for each variant ordered by their position along the *EEF1A2* gene. The aim of ACMG classification is to divide the variants into 5 categories, classes “1” and “2” are “benign classes”. The class “3” is for “unknown significance variants”, the class “4” represents “probably pathogenic variants” and the 5^{th} is for “pathogenic variants”. The CADD score evaluates the pathogenicity : if it is > 20, it supports the pathogenicity of the variant. The ”mosaic state” column indicates if there is mosaicism for the variant and the associated percentage found in the patient’s blood sample in case of mosaicism.

Table S4. Our series and previously reported patients description

This table shows the number of patients and the mean age at last examination (in years), the mean age at first steps achievement without falling (in months), the mean age at onset of epilepsy and the meanoccipito-frontal circumference (OFC) in standard deviation (SD). The comparison, when possible, was not significant, *p-value* <0.05.

Supplementary modelling figures:

Figure S1. Variations in the switch II loop

Crystal structure of eEF1A2:GDP (4C0S), with wild-type residues in blue, and variant residues in red. Variants R96C, R96H, D97N, and R381W highlighted in the switch II region.

Figure S2. R96H and R96C disrupt hydrogen bond network in eEF1Bα-bound form.

Crystal structure of $eEF1A_y$*eEF1Bα (1IJF), with wild-type residues in blue and mutant residues in red. R96 forms hydrogen bonds with E132 and E135 of helix C in eEF1B-bound form. Both variants, R96C and R96H, orient away from helix C, likely preventing hydrogen bond formation and potentially destabilising disrupting eEF1Ba-mediated release of GDP.

Figure S3. Variants impact eEF1A2 dimerisation.

Crystal structure of eEF1A2 homodimer (4C0S), with chain A in blue and chain A in gold. WT residues shown in blue, mutant residues in red. T104 interacts with Thr71 from chain B on the dimer interface. Modelling the T104R variant suggests that the bulky Arg residue will clash with chain B and disrupt dimer formation. D362 forms a salt bridge with K62 across the dimer interface, and the loss of charge in the D362N variant suggests that this interaction will be lost.

Figure S4. Variant at the nucleotide-binding site.

Crystal structure of eEF1A2:GDP (4C0S), showing the GDP-binding site. The T24M variant is depicted, adjacent to the GDP-binding

residues. The Thr residue is buried within a helix, so the Met substitution disrupts side-chain hydrogen bonds and sterically clashes. The variant likely destabilises the helix and displaces Lys20, disrupting GDP-binding.

Acknowledgements:

We want to acknowledge AnDDI-Rares and the families for their participation. In addition, the collaboration in this study were facilitated by ERN ITHACA, one of the 24 European Reference Networks (ERNs) approved by the ERN Board of Member States, cofounded by European Commission. [EU Framework Partnership Agreement ID: 3HP-HP-FPA ERN-01-2016/739516]

The aims of this study contribute to the Solve-RD project, which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement number 779257.

Author Contributions:

Conceptualization: AP, LR, CA, AV.

Formal Analysis: AP.

Investigation: AP, FA, LR, MCK, SB, JL, DT, SB, MCK, AV, RT, CBA, MG, BG, SP.

Supervision: LR, AV, CA.

Visualization: AP, CA, CBN.

Writing - Original Draft: AP, LR, CA, CBN.

Writing - Review & Editing: AP, LR, CA, CBN, AV, FM, CC, RF, JP, SJ, CF, LW, JF, AV, ASDP, TH, JM, SAL, WL, RJ, MK, DG, AG, EDB, JL, DT, SB, MCK, AV, RT, CBA, MG, BG, SP.

Conflicts of interest:

The authors declare no conflicts of interest.

Ethical Approval:

Ethical approval was not required because it was a retrospective observational study and patients have already signed consent for genetics analyses for diagnosis. Every patient has been anonymized by the clinician before collecting data (a number has been assigned to each patient).

Funding:

The authors received no financial support for the research, authorship, and/or publication of this article.

Cognitive Evaluation	Cohort (13 patients)	Literature (33 patients)	*P-value*	CI 95%
No ID	23% (3)	0%	0.082	[-0.03 ; 0.49]
Mild	31% (4)	3% (1)	0.063	[-0.02 ; 0.57]
Moderate	15% (2)	6% (2)	0.418	[-0.14 ; 0.33]
Severe*	31% (4)	76% (25)	0.008	[-0.77 ; -0.13]
Profound*	0	15% (5)	0.023	[-0.28 ; -0.02]

		Cohort	Literature	*p-value*	CI 95%
Epilepsy	**Epilepsy refractory to therapy**	25% 3/12	50% 15/30	0.13	[-0.58 ; 0.08]
	Epileptic encephalopathy	7% 1/15	24% 9/38	0.085	[-0.36 ; 0.02]
Brain MRI	**Normal**	50% 11/22	47% 15/32	0.8258	[-0.25 ; 0.31]
	Thin corpus callosum	14% 3/22	17% 5/30	0.767	[-0.23 ; 0.17]
	Delayed myelinization	9% 2/22	13% 4/30	0.636	[-0.22 ; 0.13]
	White and grey matter abnormalities	27% 6/22	9% 3/32	0.114	[-0.04 ; 0.40]
	Cerebellar and cortical atrophy*	9% 2/22	34% 11/32	0.0207*	[-0.46 ; -0.04]

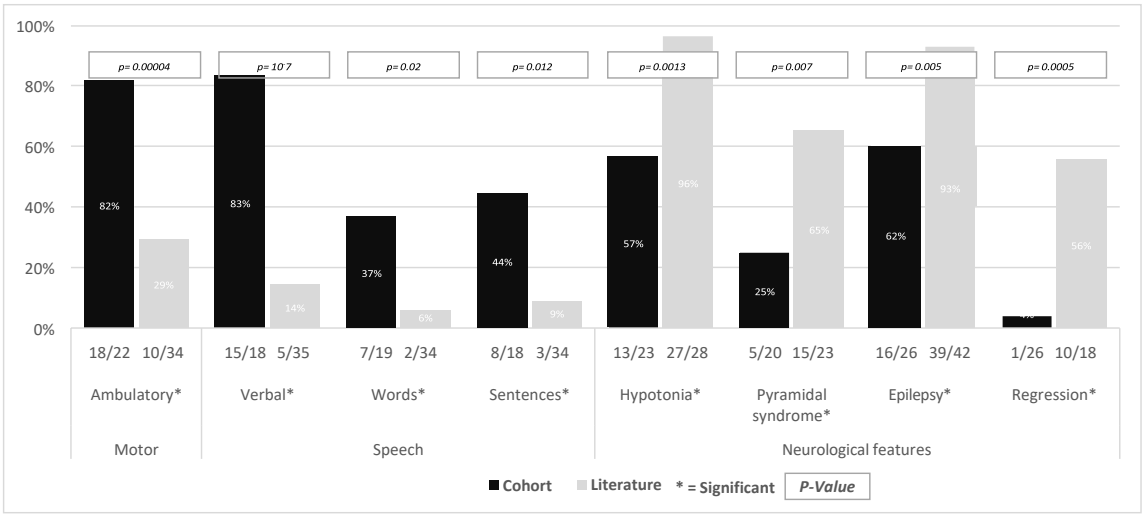
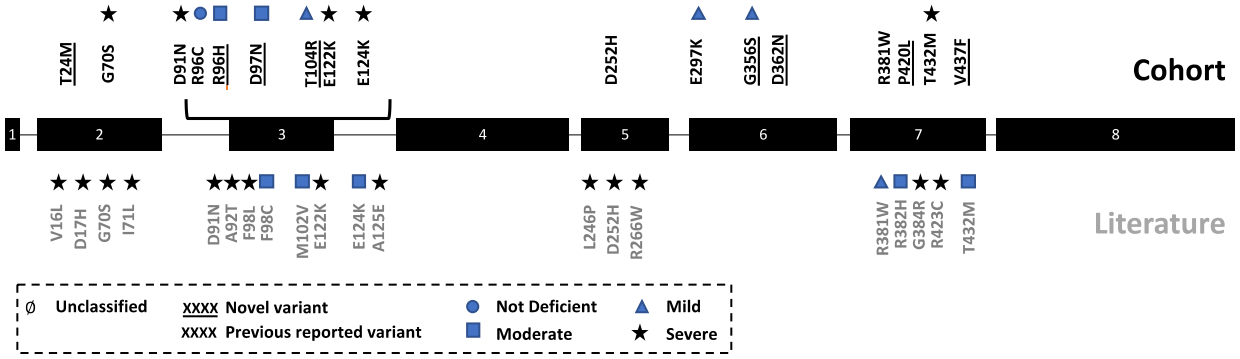


Figure 1. Our series main features compared with the patients from the literature. The figure 1 shows the proportion of our patients (in black) in comparison to those from the literature (featured in grey) according to the main criteria studied. Walking abilities is represented by the motor category. Concerning the speech abilities, “verbal” means the patient can speak and be understandable, “words” means he can express himself in words but not in sentences, “sentences” stands for the ability of the patient to make proper sentences. On the right part of the diagram, the neurological features are shown withe the presence of hypotonia, pyramidal syndrome, epilepsy and the presence of regression. * Stands for statistically significance $p<0.05$. *P-values* are framed at the top.

2a



2b

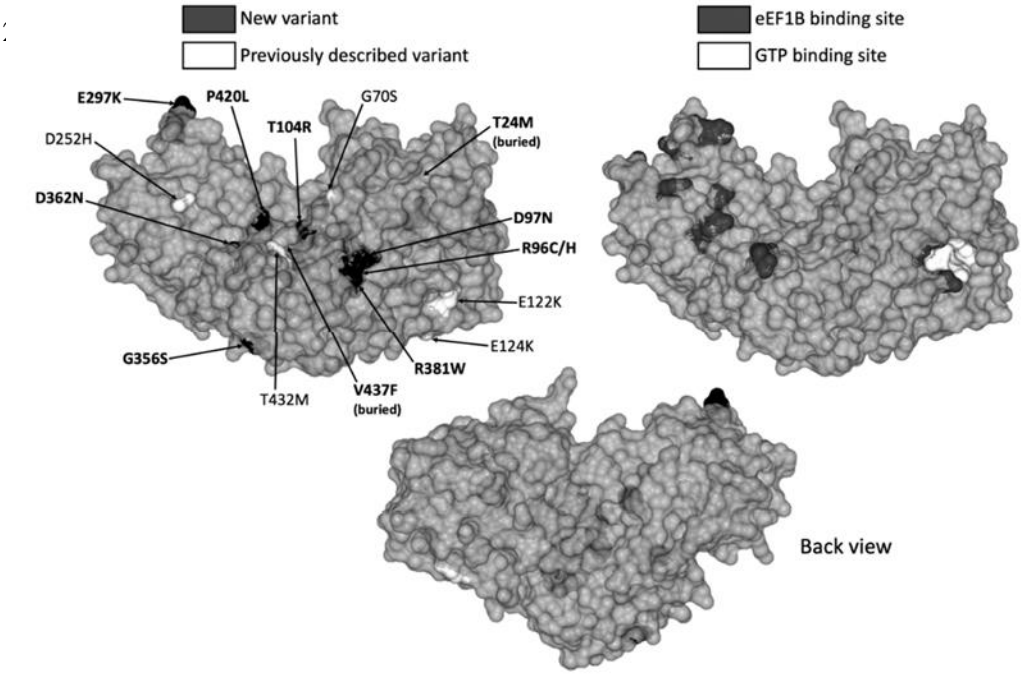


Figure 2. Distribution of pathogenic *EEF1A2* variants. Figure 2a: distribution of variants from our cohort (up) and the previous reported ones (down). Novel variants are underlined. Variants are classified by their associated phenotype in terms of ID, classification is as described as above with associated

symbols (circle for not deficient, triangle for mild ID, square for moderate ID and star for severe). Blank when the variant is not clearly classified (for example when associated with 2 ID categories). Figure 2b: variants mapped onto the crystal structure of GDP-bound eEF1A2 (PDB:4C0S). New variants are shown in black, with labels in bold, previously described mutations are in white. T24M and V437F are buried. The binding site of eEF1B is highlighted in white, and the GTP binding site in dark grey.