Intermediate repeat expansions of TBP and STUB1: genetic modifier or pure digenic inheritance in spinocerebellar ataxias?

Mathieu Barbier, Claire-Sophie Davoine, Emilien Petit, Maximilien Porche, Léna Guillot-Noel, Sabrina Sayah, Anne-Laure Fauret, Jean-Philippe Neau, Lucie Guyant- Marechal, Didier Deffond, et al.

▶ To cite this version:

Mathieu Barbier, Claire-Sophie Davoine, Emilien Petit, Maximilien Porche, Léna Guillot-Noel, et al.. Intermediate repeat expansions of TBP and STUB1: genetic modifier or pure digenic inheritance in spinocerebellar ataxias?. Genetics in medicine : official journal of the American College of Medical Genetics, 2022, 10.1016/j.gim.2022.10.009. . hal-03997720

HAL Id: hal-03997720
https://hal.sorbonne-universite.fr/hal-03997720v1

Submitted on 20 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Intermediate repeat expansions of *TBP* and *STUB1*: genetic modifier or pure digenic inheritance in spinocerebellar ataxias?

Authors

Mathieu Barbier[1]*, Claire-Sophie Davoine[1]*, Emilien Petit[1], Maximilien Porche[2], Léna Guillot-Noel[1], Sabrina Sayah[1], Anne-Laure Fauret[3], Jean-Philippe Neau[4], Lucie Guyant-Marechal[5], Didier Deffond[6], Christine Tranchant[7], Cyril Goizet[8], Giulia Coarelli[1], Anna Castrioto[9], Stephan Klebe[10], Claire Ewenczyk[1], Anna Heinzmann[1], Perrine Charles[1], Maya Tchikviladzé[1], Christine Van Broeckhoven[11,12], Alexis Brice[1] and Alexandra Durr[1].

Affiliations

[1]Sorbonne Université, Paris Brain Institute - ICM, Inserm, CNRS, APHP, Paris, France.

[2]Department of Neurology, Hôpital Lariboisière, Paris, France.

[3]Unité Fonctionnelle de Neurogénétique Moléculaire et Cellulaire, Département de Génétique et Cytogénétique, AP-HP, Hôpitaux Universitaires La Pitié Salpêtrière-Charles Foix, Paris, France.

[4]Centre Hospitalier Universitaire de Poitiers, Service de Neurologie, 86021 Poitiers, France.

[5]Department of Neurophysiology, Rouen University Hospital, 76000 Rouen, France.

[6]Department of Neurology, Gabriel Montpied Hospital, 63000 Clermont-Ferrand, France.

[7]Service de Neurologie, Hôpitaux Universitaires de Strasbourg, 67000 Strasbourg, France.

[8]Department of Medical Genetics, National Reference Centre for Neurogenetic Rare Diseases, Pellegrin Hospital, Bordeaux, and Translational Neurogenetic Team (NRGEN), INCIA CNRS UMR5287 Univ. Bordeaux, Bordeaux, France.

[9]Inserm, U1216, Grenoble Alpes University Hospital Center, Grenoble Institute of Neurosciences, Grenoble Alpes University, Grenoble, France.

[10]Department of Neurology, University Hospital Essen, Essen, Germany.

[11]Neurodegenerative Brain Diseases, VIB Center for Molecular Neurology, Antwerp, Belgium.

[12]Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium.

*These authors contributed equally.

Corresponding Author

Pr. Alexandra Durr, ICM, Hôpital de la Pitié-Salpêtrière, 47, boulevard de l'Hôpital CS21414, 75646 PARIS CEDEX 13 ; mail : alexandra.durr@icm-institute.org

Keywords: SCA48; SCA17; *TBP*; *STUB1*; spinocerebellar ataxia; cognitive impairment.

Abstract

Purpose: CAG/CAA repeat expansions in *TBP*$_{>49}$ are responsible for spinocerebellar ataxia type 17 (SCA17). We previously detected co-segregation of *STUB1* variants causing SCA48 with intermediate alleles of *TBP* in two families. This co-segregation questions the existence of SCA48 as a monogenic disease.

Methods: We systematically sequenced *TBP* repeats in 34 probands of dominant ataxia families with *STUB1* variants. In addition, we searched for pathogenic *STUB1* variants in probands with expanded alleles of *TBP*$_{>49}$ (n=2) or intermediate alleles of *TBP*$_{\geq 40}$ (n=47).

Results: *STUB1* variants were found in half of the *TBP*$_{40\text{-}49}$ cohort. Mirroring this finding, *TBP*$_{40\text{-}49}$ alleles were detected in 40% of *STUB1* probands. The longer the *TBP* repeat length the more likely the occurrence of cognitive impairment (p = 0.0129) and the faster the disease progression until death (p = 0.0003). Importantly, thirteen *STUB1* probands presenting with the full SCA48 clinical phenotype had normal *TBP*$_{37\text{-}39}$ alleles, excluding digenic inheritance as the sole mode.

Conclusion: We show that intermediate *TBP*$_{40\text{-}49}$ alleles act as disease modifiers of SCA48 rather than a *STUB1*/*TBP* digenic model. This distinction from what has been proposed before has crucial consequences for genetic counseling in SCA48.

Introduction

Spinocerebellar ataxia type 17 (SCA17) is a rare form of SCA caused by a pathogenic expansion of a CAG/CAA repeat in the *TBP* gene[1]. SCA17 is characterized mainly by cerebellar ataxia with predominant cognitive impairment. Up until now, the normal CAG/CAA repeat number in the polyglutamine tract was considered to be 25 to 40 units[2]. It has been established that all individuals with CAG/CAA repeats greater than 49 have the fully penetrant phenotype. By contrast, the pathogenicity and penetrance of rare *TBP* alleles with 41-49 repeats, known as "intermediate expanded alleles" (IE), are much more debated[2,3]. In previous reports, a repeat number of 43 or more has been found to be pathological[4], but healthy individuals with 45 and 49 repeats have also been identified[5,6]. Moreover, SCA17 CAG/CAA repeat length is difficult to assess due to a complex polyglutamine tract structure that can vary, the most frequent one being $(CAG)_3(CAA)_3(CAG)_n$-CAA-CAG-CAA$(CAG)_n$-CAA-CAG, with "n" referring to the polymorphic CAG repeats. Thus, the absence of consensus regarding the pathological and clinical consequences of *TBP*-IE hampers genetic diagnostic and counseling, as clinicians are unable to predict with any certainty the impact of these alleles.

The neurological manifestations of SCA17, in addition to the cerebellar ataxia, are pyramidal syndrome, parkinsonism, chorea, epilepsy, and cognitive impairement[2]. The association between cerebellar ataxia and cognitive impairment is also a hallmark of SCA48 due to heterozygous pathogenic variants in *STUB1*[7].

We have previously found that pathogenic variants of *STUB1* were a frequent cause of ataxia in a series of 440 probands[8]. Interestingly, the screening of this large series also revealed two probands with *TBP*-IE alleles (41 and 46 CAG/CAA repeats respectively), in addition to the variants detected in *STUB1*[8]. This suggests that *TBP*-IE may be more frequent among *STUB1* patients and affect disease penetrance, and/or that *STUB1* variants could explain the disease in patients with *TBP*-IE. This relationship was recently confirmed in a study concluding that digenic inheritance of *TBP*-IE / *STUB1* variants is necessary for the full SCA48 clinical phenotype to develop[9]. This finding calls into question the postulate that SCA48 is a monogenic disease which in turn, has important consequences on genetic counseling[9].

Here we systematically screened for *STUB1* variants in a cohort of index cases with *TBP*-IE$_{41-49}$ or pathogenic expanded alleles. We also extended our screening to individuals with *TBP*$_{=40}$ to delineate the model of interaction between *TBP* and *STUB1* as precisely as possible. In addition, we genotyped the *TBP* CAG/CAA repeats in a larger cohort of *STUB1* index cases. Our observations exclude systematic digenic inheritance in explaining full penetrance of SCA48. *TBP*-IE seem rather to act as modifiers of the SCA48 phenotype, increasing the severity of the disease while *STUB1* variants can solely explain the full disease spectrum in some cases.

Patients & methods

Patients

Thirty-four probands with a dominant cerebellar ataxia who also carried a heterozygous *STUB1* pathogenic variant were selected as followed: seven through the SPATAX research network (https://spatax.wordpress.com) following exome sequencing (n=253) plus 27 previously published cases including the two proband heterozygous for *TBP*-IE[8] (Figure 1). These 34 *STUB1* probands were then sequenced for the *TBP* CAG/CAA expansion. We selected 49 individuals with intermediate *TBP* expansion (with a predicted CAG/CAA repeat number ≥ 40 +/- 1 repeat) who had been addressed to the diagnostic unit at the Pitié-Salpêtrière University Hospital in Paris (n=6) and from the SPATAX database (n=43). The precise number of *TBP* CAG/CAA repeats was systematically checked by Sanger sequencing. Two probands with pathogenic repeat numbers were also included after the resequencing of the *TBP* expansion (53 and 55 CAG/CAA repeats respectively). The *STUB1* gene was also Sanger sequenced among these 51 *TBP* probands(Figure 1). DNA from relatives was also analyzed when available. Available clinical and imaging data were retrieved, and patients were invited to come for neuropsychological evaluation when possible. Patients' neurologists were contacted for relevant information if the patient was not followed at the national reference center for rare diseases.

Genetic analyses

The pathogenic *TBP* expansion was identified by Sanger sequencing after the amplification of a PCR-fragment containing the repeat as has been previously described[10]. Two of the participants were heterozygous for pathogenic *TBP* expanded alleles (53 and 55 repeats). They both had the following alternative structure showing a duplication of the CAA-CAG-CAA domain III separated by a stretch of CAG, making the repeat configuration for the pathogenic allele as follows: $(CAG)_3$-$(CAA)_3$-$(CAG)_9$-CAA-CAG-CAA-$(CAG)_{15}$-CAA-CAG-CAA-$(CAG)_{15or17}$-CAA-CAG. This less frequent repeat configuration has been described among SCA17 patients elsewhere and reflects a polymorphic repeat configuration[11,12]. The other index cases have the same structure of the *TBP* repeat: $(CAG)_3$-$(CAA)_3$-$(CAG)_n$-CAA-CAG-CAA-$(CAG)_n$-CAA-CAG (Supplementary Figure S1).

Whole-exome sequencing (WES) was performed on the 253 index cases with cerebellar ataxia with an unknown genetic cause. DNA libraries were prepared using KAPA HyperPrep Kits (Roche). Exome sequencing was achieved using the Twist Human RefSeq Panel (Twist Bioscience) on a NovaSeq 6000 (Illumina). Standard quality control, alignment and variant calling were performed using the Illumina DRAGEN Bio-IT Platform (Illumina). The correct coverage of *STUB1* was verified manually from BAM files. *STUB1* variants were then filtered as previously described using the VarAFT 2.17-2 software[8,13]. All candidate *STUB1* variants were further confirmed by Sanger sequencing.

Neuropsychological examination

A neuropsychological evaluation was performed for 12 patients and included: the Mini Mental State Examination (MMSE) score[14] and the FAB Score (Frontal Assessment Battery)[15]. An assessment of executive functioning included the Stroop Color Word Test[16], the Verbal Fluency test (animals and letters), the Trail Making Test (TMT) and the Modified Wisconsin Card Sorting Test (mWCST)[17]. Digit and spatial span (backward and forward) and bell's test were used to evaluate attention and working memory. To evaluate episodic memory, we used the Free and Cued Selective Reminding Test (FCSRT)[18,19]. We also evaluated visuospatial skills using subtests from the Visual Object and Space Perception (VOSP) and visuo-construction with the Rey Complex Figure. To assess language, we used the naming test from the GRECO Neuropsychological Semantic Battery as well as standard semantic and letter fluencies.

Statistical analyses

Effect of *TBP*-IE on the occurrence of dementia was estimated with logistic regression. A Receiver Operator Characteristic (ROC) curve to differentiate patients with and without clinical dementia was performed using *TBP*-IE as a continuous variable and Youden's statistic was used to determine the best cut-off. Survival analysis was performed using the Kaplan-Meier estimator to plot survival curves and the effect of having *TBP*-IE was estimated using Cox models (Hazard ratio HR, with a 95% confidence interval). The Student t-test was used to compare the age of onset of patients with or without having *TBP*-IE. Values are described as mean ± SD. Statistical tests were performed at the conventional two-tailed type I error of 0.05. Data were analyzed using R version 3.6.2.

Results

High frequency of STUB1 pathogenic variants in patients with TBP intermediate alleles

The sequencing of the *TBP* CAG/CAA expansion in *STUB1* patients and the screening of *STUB1* variants in the *TBP* cohort ultimately revealed a total of 43 probands with *STUB1* variants and 37 probands with either pathogenic (n=2) or intermediate TBP_{40-49} expansions (n = 35). Seventeen probands both had *STUB1* variants and *TBP*-IE. These *STUB1*/*TBP*-IE individuals represent 40% of all *STUB1* probands and 49% of all patients with *TBP*-IE (Figure 1). We did not find *STUB1* variants in the individuals with pathogenic $TBP_{>49}$ expansions. Of note, in two families for whom probands had *TBP*-IE and who were negative for *STUB1*, the intermediate alleles of *TBP* did not segregate with the disease (Supplementary Figure S2).

TBP intermediate alleles as a disease modifier in SCA48 patients

We analyzed the impact of carrying *TBP*-IE on the phenotype in *STUB1* patients, in particular survival and the variable presence of relevant cognitive impairment (n = 63 patients out of 43 families, Table 1). *TBP* CAA/CAG repeats on both alleles ranged from 29 to 47 in the whole *STUB1* cohort. Having *TBP*-IE was significantly associated with the risk of developing dementia among SCA48 patients (n = 24 without dementia and n = 39 with dementia). This risk increased with the number of CAA/CAG repeats (OR = 1.8 [1.2-3.2]; p = 0.0129) (Figure 2A). The unsupervised analysis of the ROC curve showed that the most efficient threshold to correctly classify *STUB1* patients with or without dementia was 40 repeats (Figure 2B). These results suggest that the adverse effect of *TBP*-IE on *STUB1* patients begins at 40 repeats. Having *TBP*-IE for *STUB1* patients also affected survival, with significantly shorter lifespans when $TBP_{\geq 40}$ (HR = 5.1 [2.1-12.4], $p = 0.0003$) (Figure 3). Age at onset did not differ between patients with or without *TBP*-IE (43.05 ± 9.37 vs 44.54 ± 12.48 respectively; $p = 0.62$).

Cognitive impairment was clinically evident in 65% of *STUB1* patients (41/63, Table 1). Importantly, 24 *STUB1* patients with $TBP_{37\text{-}39}$ had cognitive features, some of them carrying known pathogenic and recurrent *STUB1* variants. Detailed neuropsychological evaluations were available for 12 patients: one with SCA17 (TBP_{53}), seven patients carrying *STUB1* variants with $TBP_{38\text{-}39}$ and four *STUB1* patients with $TBP\text{-IE}_{40\text{-}49}$ (Table 2). All had a cognitive profile involving multiple cognitive domains except for episodic memory which was preserved in all. This corresponds to the presence of a cerebellar cognitive affective syndrome (CCAS) consisting of cognitive and affective deficits due to cerebellar disease only[20]. CCAS presents as a heterogenous picture of executive and attentional deficits, visual-spatial disorganization, and language difficulties. This also affects socio-emotional processing with deficits in social and behavioral cognition, present in all patients in our study. The primary impaired domain was executive function, all patients had difficulties in processing speed, and/or cognitive flexibility and/or inhibition and/or planning. Language impairment with difficulties in verbal fluencies, difficulties in naming, and mild anomia was present in all patients with both *TBP*-IE and *STUB1* variants. Conversely, those who carried *STUB1* alone showed less severe difficulties in this domain. This could point towards more frontal involvement in *STUB1* patients with additional *TBP*-IE. In addition, we had the clinical impression of a more rapid cognitive decline in patients also carrying *TBP*-IE, but the number of participants with longitudinal neuropsychological evaluations is too few to conclude.

Interestingly, we detected known pathogenic and/or recurrent *STUB1* variants in families some associated with *TBP*-IE segregating and some not allowing us to compare the two genotypes (Figure 4). The aforementioned individuals were heterozygous for the p.P243L variant, a recurrent substitution previously associated with both SCAR16 and SCA48[8,21]. Dementia was detected only in two relatives carrying both the *STUB1* p.P243L variant and *TBP*-IE (both with 43 repeats), and not in families with normal alleles of *TBP* (Figure 4). In a subgroup of six *STUB1* families (three families carrying either the *STUB1* p.K143del (this amino-acid deletion affects an amino-acid position already associated with SCA48 in two independent reports[8,22]), and three families with the *STUB1* p.C69Y segregating) the variable presence of cognitive impairment was observed both in families which segregated normal and intermediate alleles of *TBP*. This finding shows that carrying *TBP*-IE (e.g digenism) is not mandatory for producing the SCA48 phenotype.

Discussion

Considering a previous observation by our team, notably identifying patients with *STUB1* variants initially diagnosed as SCA17, we screened for *STUB1* variants in a larger cohort of SCA17 patients with the hypothesis that this gene may be responsible for the disease alone or act in concert with *TBP* expansions. We report rare, pathogenic, and recurrent variants of *STUB1* found in nearly half of a cohort of patients with dominantly inherited ataxia and with alleles ranging from 40-47 repeats. Together with the description of two families in which *TBP*-IE do not segregate with the disease, these observations argue against *TBP*-IE acting alone and suggest that other genetic events like *STUB1* variants (or when *STUB1* variants are not present, pathogenic variants in other genes, intermediate alleles in other expansion loci etc.) explain the disease in probands with *TBP*-IE. We did not find a common mutated gene in these remaining patients. Thus, *STUB1* appears as the most frequently mutated gene among symptomatic patients with *TBP*-IE. We previously shown that there no genotype-phenotype correlation in *STUB1* patients[8]. Here *TBP*-IEs were observed in patients with no specific location of *STUB1* variants, excluding again a particular genotype-phenotype correlation in *STUB1* / *TBP*-IE individuals.

The discussed role of “intermediate alleles” in the pathogenesis in the same or other phenotypes is common in expansion diseases. For almost all of them, there is a gap in thresholds between normal and pathogenic expanded alleles[4]. The concept of low penetrance, sometimes used to explain this gray zone, leaves the door open to additional modifying events including genetic ones, or alternatively that other causal genes may be responsible for the resulting phenotype.

Indeed, numerous publications report the existence of healthy adult with *TBP*-IE ranging from 43 to 49 repeats in SCA17 families[23–28]. In addition, in these studies *TBP*-IE were observed in cohorts of healthy controls, although with a low frequency (<5% regardless of the ethnicity). This illustrates the complexity of establishing a threshold repeat number between normal and pathological for *TBP*.

In addition, several SCA subtypes present with highly overlapping clinical presentations. A typical example is the comparison between SCA17 and SCA48 which are both characterized by ataxia with predominant cognitive impairment compatible with CCAS[1,7].Considering the numerous previous studies reporting inconsistencies between carrying *TBP*-IE or segregation with disease status, and in accordance with the high proportion of *STUB1* variants in patients with *TBP*-IE detected in our study (49%), it is very likely that carrying *TBP*-IE may not be sufficient to develop the disease.

The next key question is how *STUB1* pathogenic variants and *TBP*-IE act together to define a model of inheritance. In other words, can we consider these patients as SCA17 with *TBP*-IE plus *STUB1* variants, or are they SCA48 patients with a particular phenotype influenced by carrying small-expanded alleles of *TBP*?

A recent publication reported that all except one (30/31) index cases with *TBP*$_{41\text{-}46}$ were also positive for *STUB1* variants[9]. The unexpectedly high proportion of *STUB1* variants in individuals carrying *TBP*-IE made them conclude that *STUB1* alone does not trigger ataxia with cognitive impairment. The conclusions drawn from our own observations diverge about the proposed model. In our cohort 60% of probands carried *STUB1* variants with normal *TBP* alleles (<40 repeats) but shared the same phenotype of adult-onset ataxia. Although there was a clear enrichment of *STUB1* patients carrying *TBP*-IE in our sample cohort versus the frequency of *TBP*-IE$_{40\text{-}49}$ in control databases (~1.2% in the gnomAD database v3.1.2, non-Finnish population), the existence of *STUB1* patients with normal *TBP* alleles questions the digenic model. Thirteen probands who presented with both ataxia and dementia, compatible with SCA48 diagnosis, were positive for *STUB1* variants and carried normal *TBP* alleles (37-39 repeats). No common mutated gene or candidate pathological variant was evidenced by whole exome or gene panel sequencing to sustain a digenic model among these thirteen patients. Three recurrent *STUB1* pathogenic variants were detected in ten individuals. First, the *STUB1* p.Y49C and p.C69Y that we found in unrelated patients and which were also detected in the Italian study[8,9]. In the Italian cohort, the p.Y49C variant of interest was reported in one patient with cognitive impairment carrying a *TBP*$_{42}$. In our cohort, the same variant was found in four unrelated probands. Three of them developed cognitive impairment, and all were carrying normal *TBP* alleles (38 repeats). This shows that the development of cognitive deficits in addition to the ataxic phenotype is not entirely linked to the length of *TBP* alleles. Second, all three of the probands that have the *STUB1* p.C69Y variant had cognitive impairment, but only one carried an additional TBP_{40} allele. Third, we detected the *STUB1* p.N65S in five patients from three families, among two sibs with a discordant cognitive phenotype, as well as between families one with and another without dementia, both carrying *TBP*$_{38}$. This *STUB1* variant was originally associated with SCAR16 in a loss of function study, and more recently associated with SCA48[8,29]. Thus, these patients, presenting with ataxia and cognitive impairment, can be considered as "pure" *STUB1* patients with well-known pathogenic variants, excluding a digenic *STUB1*/*TBP*-IE model as the genetic basis of SCA48.

Alternatively, we proposed that *TBP*-IE may act as a disease modifier of SCA48. Patients with *STUB1* variants and *TBP*-IE died significantly earlier than patients carrying *STUB1* variants and normal *TBP* alleles. The comparison of probands from different families carrying the same *STUB1* variants, but with or without *TBP*-IE showed that dementia tended to be more frequent in the presence of *TBP*-IE. Statistical analysis of the entire *STUB1* cohort highlighted the significant association of *TBP*-IE with cognitive impairment. In addition, a more in-depth clinical analysis of *STUB1* patients with and without *TBP*-IE showed that *TBP*-IE was associated with a cognitive deficit, considered as CCAS but with more language impairment than in patients without *TBP*-IE. This could mean a more widespread disease in the presence of *TBP*-IE. Our data also showed a modifying effect of *TBP*-IE on cognition with a threshold at 40 repeats, as established by an unsupervised classification of *STUB1* patients according to their genotype for *TBP* and their level of cognitive deterioration. Of note, one *STUB1* patient was also homozygous for *TBP*-IE (42/43). This woman was one of the most severely affected patients. First, disease onset was very early (around 24y). Second, her clinical presentation was very severe with mutism and behavior abnormalities at age 33 after 9 years of disease duration. In addition to the cerebellar ataxia there was predominant myoclonus, pyramidal signs, optic atrophy and retinitis. However to definitely conclude to an additive effect of *TBP*-IE, the analysis of larger cohorts is warranted.

The role of intermediate alleles in expansion disorders is a source of debate regarding their theoretical pathogenic effects. It has already been suggested that intermediate alleles, could constitute susceptibility factors for other diseases[4]. A well-known example is the role of CAG repeats in *ATXN2*. Pathogenic uninterrupted CAG repeat expansions in this gene are responsible for SCA2, while intermediate alleles are risk factors for developing ALS[30,31]. Intermediate CAG repeats in *ATXN1, ATXN2* and *HTT* were also suspected to influence disease in frontotemporal dementia as well as in Alzheimer's disease[32]. Thus, our study confirms that intermediate CAG repeat expansions may act as genetic modifiers in neurodegenerative disorders other than the primary disease associated with the pathological repeat expansion.

Our work, in addition to independent reports[9,33] both highlight an unexpected link between *STUB1* variants and *TBP*-IE. To date, it's not clear how these two protagonists interact. The CHIP protein encoded by *STUB1* functions as an E3 ubiquitin ligase/cochaperone, involved in the cellular protein quality control system. The loss-of-function theory has been proposed to explain the pathogenicity of *STUB1* variants. The TATA-binding protein encoded by *TBP* is an essential component of transcription factor IID complex which binds to the core promoter, to initiate transcription by RNA polymerase II[34]. The polyglutamine tract, abnormally expanded in SCA17, regulates DNA-binding of TBP[35,36]. The abnormal structure of CAG-expanded in TBP may be recognized by CHIP to be eliminated. Thus, the *STUB1* variant could favor a pathogenic accumulation of TBP protein carrying an intermediate-expanded tract of polyglutamines. Interestingly we previously published the neuropathological analysis of a patient with *TBP*$_{46}$[10] also with a *STUB1* variant showing neuronal intranuclear inclusions (NII) containing expanded polyglutamine. We previously did not find NII in a *STUB1* patient with *TBP*$_{38}$ in line with previous neuropathological reports on *STUB1* patients except in one study[37–39]. The variable presence of NII containing expanded polyglutamine in *STUB1* patients may reflect different sizes of the *TBP* expansion. The presence of NII in brains of *STUB1* / *TBP*-IE individuals may contribute to the higher diffusivity of the pathology in the brains of these patients. Another hypothesis would be that the CHIP proteins may regulate TBP recruitment to the promotor. Recently in a yeast model, an E3 ubiquitin ligase, San1, was shown to interact with TBP and this interaction is necessary for the formation of the pre-initiation complex[40]. To our knowledge there is no ortholog for *San1* in humans. However, as CHIP is also an E3 ubiquitin ligase, and considering our results, it is reasonable to suggest that CHIP proteins may interact with TBP. The effect of pathogenic *STUB1* variants could be synergized by *TBP*-IE, altering the formation of the pre-initiation complex and recruitment of RNA Polymerase II for gene activation and influence possibly the clinical outcome and penetrance in *STUB1* / *TBP*-IE patients. In any event, functional investigations are now warranted to better understand the relationship between *STUB1* and *TBP*-IE.

Conclusion

We propose that cerebellar ataxia associated with cognitive impairment, the predominant clinical features of SCA48, should be considered as a phenotypical continuum in which *TBP*-IE (including alleles with 40 repeats) are an important risk factor influencing both, the development of dementia with frontal involvement and survival, in SCA48. This has important consequences for genetic counseling.

Data Availability

Data supporting this work are available upon request from qualified principal investigator to the corresponding author. Clinical data sets have been de-identified.

Acknowledgments

Part of this work was carried out on the iGenSeq sequencing and the Data Analysis Core bioinformatic facilities of the ICM.

The research leading to these results has received funding through the "Investissements d'avenir" ANR-10- IAIHU-06 and ANR-11- INBS-0011 – NeurATRIS: Translational Research Infrastructure for Biotherapies in Neurosciences program.

We deeply thank the French clinical network, patients and their referral neurologists and in particular: Rabab Debs, Jérôme Honnorat, Pierre Labauge, Alain Creange, Fabienne Prieur. All of these contributors have agreed to this manuscript.

Author Contributions

Conceptualization: MB, C-SD, AB, AD; Data curation: AD; Formal analysis: MB, C-SD, EP, MP, LG-N, AD; Funding acquisition: AD; Investigation: MP, SS, A-LF, J-PN, LG-M, DD, CT, CG, GC, AC, SK, CE, AH, PC, MT, CVB, AB, AD; Methodology: MB, C-SD, EP, AD; Supervision: AD; Visualization: MB, C-SD, EP; Writing-original draft: MB, C-SD, EP, AB, AD.

Ethics declaration

The study was conducted as was presented to and approved by the Ethics Review Board (Paris Necker ethics committee approval (RBM 01-29 and RBM 03-48) to A.B. and A.D). Written informed consent was obtained from all participants.

Conflicts of Interest

The authors declare no conflict of interest related to this work.

Figure and Table legends

Figure 1. Flow chart of the study design, genetic analyses, and summary of results.

Figure 2. *TBP* intermediate alleles are associated with the risk of developing cognitive impairment in SCA48/*STUB1* patients. (**A**) Violin plot representing the distribution of *TBP* CAG/CAA repeat length in conjunction with cognitive state in *STUB1* patients. The red dotted line indicates the most efficient threshold (40 repeats) to correctly classify *STUB1* patients with or without dementia. (**B**) Receiver Operator Characteristic (ROC) curve separating patients with and without dementia using *TBP* repeat length. The best cut-off value is 39.5 (value which gave the curve closer to the top-left corner), with a high specificity (92%) but low sensitivity (44%).

Figure 3. The carrying of *TBP* intermediate alleles reduces survival of *STUB1* patients. Kaplan-Meier estimator for patients with and without *TBP* CAG/CAA repeat length < 40. The cross on the estimator represents censoring times.

Figure 4. Pedigrees comparing patients with cerebellar ataxia carrying the same pathogenic *STUB1* variant, with *TBP* allele, intermediate CAG repeats, ≥ 40 on the left side and normal <40 on the right side of the figure. The numbers indicate the size of both sequenced alleles of *TBP*. All patients (females as circles and males as squares) in black presented with an ataxic phenotype. The letter “D” highlights the presence of clinically meaningful cognitive impairment. ? = clinical details not available.

Supplementary Figure S1. Structure of the *TBP* CAG/CAA repeats: (A) Structure found in patients with either normal or intermediate alleles. (B) Structure found in patients with pathogenic alleles (53 and 55 repeats). Variable numbers of CAG repeats among individuals are indicated.

Supplementary Figure S2. *TBP* intermediate alleles do not segregate with the disease in two families negative for *STUB1*. *TBP* CAG/CAA repeat number is mentioned. AAD-566: patients had predominately cerebellar features with intellectual deficiency; FSP-948: patients had a predominantly spastic paraparesis with intellectual deficiency and extrapyramidal features.

Table 1 Clinical characteristics of 63 *STUB1* variant (NM_005861.4, GRCh38) patients listed according to the size of the intermediate allele of *TBP* decreasing from 47 to 36 CAG/CAA repeats.

Table 2. Cognitive evaluation of one patient with confirmed SCA17/*TBP* CAG/CAA repeat expansion >49 and eleven SCA48/*STUB1* patients with non-pathogenic expansion repeat number (seven without *TBP*-IE and four with *TBP*-IE). Cognitive domains: - = not affected, + = mildly affected, ++ = strongly affected, NA = not available.

References

1. Koide R, Kobayashi S, Shimohata T, et al. A neurological disease caused by an expanded CAG trinucleotide repeat in the TATA-binding protein gene: a new polyglutamine disease? *Hum Mol Genet*. 1999;8(11):2047-2053. doi:10.1093/hmg/8.11.2047

2. Toyoshima Y, Onodera O, Yamada M, Tsuji S, Takahashi H. Spinocerebellar Ataxia Type 17. In: Adam MP, Ardinger HH, Pagon RA, et al., eds. *GeneReviews®*. University of Washington, Seattle; 1993. Accessed June 23, 2021. http://www.ncbi.nlm.nih.gov/books/NBK1438/

3. Gao R, Matsuura T, Coolbaugh M, et al. Instability of expanded CAG/CAA repeats in spinocerebellar ataxia type 17. *Eur J Hum Genet*. 2008;16(2):215-222. doi:10.1038/sj.ejhg.5201954

4. Depienne C, Mandel JL. 30 years of repeat expansion disorders: What have we learned and what are the remaining challenges? *Am J Hum Genet*. 2021;108(5):764-785. doi:10.1016/j.ajhg.2021.03.011

5. Oda M, Maruyama H, Komure O, et al. Possible reduced penetrance of expansion of 44 to 47 CAG/CAA repeats in the TATA-binding protein gene in spinocerebellar ataxia type 17. *Arch Neurol*. 2004;61(2):209-212. doi:10.1001/archneur.61.2.209

6. Zühlke C, Dalski A, Schwinger E, Finckh U. Spinocerebellar ataxia type 17: report of a family with reduced penetrance of an unstable Gln49 TBP allele, haplotype analysis supporting a founder effect for unstable alleles and comparative analysis of SCA17 genotypes. *BMC Med Genet*. 2005;6:27. doi:10.1186/1471-2350-6-27

7. Genis D, Ortega-Cubero S, San Nicolás H, et al. Heterozygous STUB1 mutation causes familial ataxia with cognitive affective syndrome (SCA48). *Neurology*. 2018;91(21):e1988-e1998. doi:10.1212/WNL.0000000000006550

8. Roux T, Barbier M, Papin M, et al. Clinical, neuropathological, and genetic characterization of STUB1 variants in cerebellar ataxias: a frequent cause of predominant cognitive impairment. *Genet Med*. 2020;22(11):1851-1862. doi:10.1038/s41436-020-0899-x

9. Magri S, Nanetti L, Gellera C, et al. Digenic inheritance of STUB1 variants and TBP polyglutamine expansions explains the incomplete penetrance of SCA17 and SCA48. *Genet Med*. 2022;24(1):29-40. doi:10.1016/j.gim.2021.08.003

10. Fujigasaki H, Martin JJ, De Deyn PP, et al. CAG repeat expansion in the TATA box-binding protein gene causes autosomal dominant cerebellar ataxia. *Brain*. 2001;124(Pt 10):1939-1947. doi:10.1093/brain/124.10.1939

11. Shatunov A, Fridman EA, Pagan FI, et al. Small de novo duplication in the repeat region of the TATA-box-binding protein gene manifest with a phenotype similar to variant Creutzfeldt-Jakob disease. *Clin Genet*. 2004;66(6):496-501. doi:10.1111/j.1399-0004.2004.00356.x

12. Gao R, Matsuura T, Coolbaugh M, et al. Instability of expanded CAG/CAA repeats in spinocerebellar ataxia type 17. *Eur J Hum Genet*. 2008;16(2):215-222. doi:10.1038/sj.ejhg.5201954

13. Desvignes JP, Bartoli M, Delague V, et al. VarAFT: a variant annotation and filtration system for human next generation sequencing data. *Nucleic Acids Res*. 2018;46(W1):W545-W553. doi:10.1093/nar/gky471

14. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189-198. doi:10.1016/0022-3956(75)90026-6

15. Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB: a Frontal Assessment Battery at bedside. *Neurology*. 2000;55(11):1621-1626. doi:10.1212/wnl.55.11.1621

16. Golden CJ. Identification of brain disorders by the Stroop Color and Word Test. *J Clin Psychol*. 1976;32(3):654-658. doi:10.1002/1097-4679(197607)32:3<654::aid-jclp2270320336>3.0.co;2-z

17. Nelson HE. A modified card sorting test sensitive to frontal lobe defects. *Cortex*. 1976;12(4):313-324. doi:10.1016/s0010-9452(76)80035-4

18. Buschke H. Cued recall in amnesia. *J Clin Neuropsychol*. 1984;6(4):433-440. doi:10.1080/01688638408401233

19. Grober E, Buschke H, Crystal H, Bang S, Dresner R. Screening for dementia by memory testing. *Neurology*. 1988;38(6):900-903. doi:10.1212/wnl.38.6.900

20. Schmahmann JD, Sherman JC. The cerebellar cognitive affective syndrome. *Brain*. 1998;121 (Pt 4):561-579. doi:10.1093/brain/121.4.561

21. Hayer SN, Deconinck T, Bender B, et al. STUB1/CHIP mutations cause Gordon Holmes syndrome as part of a widespread multisystemic neurodegeneration: evidence from four novel mutations. *Orphanet J Rare Dis*. 2017;12(1):31. doi:10.1186/s13023-017-0580-x

22. Pakdaman Y, Berland S, Bustad HJ, et al. Genetic Dominant Variants in STUB1, Segregating in Families with SCA48, Display In Vitro Functional Impairments Indistinctive from Recessive Variants Associated with SCAR16. *Int J Mol Sci*. 2021;22(11):5870. doi:10.3390/ijms22115870

23. Rolfs A, Koeppen AH, Bauer I, et al. Clinical features and neuropathology of autosomal dominant spinocerebellar ataxia (SCA17). *Ann Neurol*. 2003;54(3):367-375. doi:10.1002/ana.10676

24. Stevanin G, Fujigasaki H, Lebre AS, et al. Huntington's disease-like phenotype due to trinucleotide repeat expansions in the TBP and JPH3 genes. *Brain*. 2003;126(Pt 7):1599-1603. doi:10.1093/brain/awg155

25. Zühlke C, Gehlken U, Hellenbroich Y, Schwinger E, Bürk K. Phenotypical variability of expanded alleles in the TATA-binding protein gene. Reduced penetrance in SCA17? *J Neurol*. 2003;250(2):161-163. doi:10.1007/s00415-003-0958-7

26. Zühlke C, Dalski A, Schwinger E, Finckh U. Spinocerebellar ataxia type 17: report of a family with reduced penetrance of an unstable Gln49 TBP allele, haplotype analysis supporting a founder effect for unstable alleles and comparative analysis of SCA17 genotypes. *BMC Med Genet*. 2005;6:27. doi:10.1186/1471-2350-6-27

27. Oda M, Maruyama H, Komure O, et al. Possible reduced penetrance of expansion of 44 to 47 CAG/CAA repeats in the TATA-binding protein gene in spinocerebellar ataxia type 17. *Arch Neurol*. 2004;61(2):209-212. doi:10.1001/archneur.61.2.209

28. Tremolizzo L, Curtò NA, Marzorati L, et al. Early-onset SCA17 with 43 TBP repeats: expanding the phenotype? *Neurol Sci*. 2011;32(5):941-943. doi:10.1007/s10072-011-0662-9

29. Heimdal K, Sanchez-Guixé M, Aukrust I, et al. STUB1 mutations in autosomal recessive ataxias - evidence for mutation-specific clinical heterogeneity. *Orphanet J Rare Dis*. 2014;9:146. doi:10.1186/s13023-014-0146-0

30. Elden AC, Kim HJ, Hart MP, et al. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature*. 2010;466(7310):1069-1075. doi:10.1038/nature09320

31. Lattante S, Millecamps S, Stevanin G, et al. Contribution of ATXN2 intermediary polyQ expansions in a spectrum of neurodegenerative disorders. *Neurology*. 2014;83(11):990-995. doi:10.1212/WNL.0000000000000778

32. Rosas I, Martínez C, Clarimón J, et al. Role for ATXN1, ATXN2, and HTT intermediate repeats in frontotemporal dementia and Alzheimer’s disease. *Neurobiol Aging*. 2020;87:139.e1-139.e7. doi:10.1016/j.neurobiolaging.2019.10.017

33. Reis MC, Patrun J, Ackl N, et al. A Severe Dementia Syndrome Caused by Intron Retention and Cryptic Splice Site Activation in STUB1 and Exacerbated by TBP Repeat Expansions. *Front Mol Neurosci*. 2022;15:878236. doi:10.3389/fnmol.2022.878236

34. Gostout B, Liu Q, Sommer SS. “Cryptic” repeating triplets of purines and pyrimidines (cRRY(i)) are frequent and polymorphic: analysis of coding cRRY(i) in the proopiomelanocortin (POMC) and TATA-binding protein (TBP) genes. *Am J Hum Genet*. 1993;52(6):1182-1190.

35. Lescure A, Lutz Y, Eberhard D, et al. The N-terminal domain of the human TATA-binding protein plays a role in transcription from TATA-containing RNA polymerase II and III promoters. *EMBO J*. 1994;13(5):1166-1175.

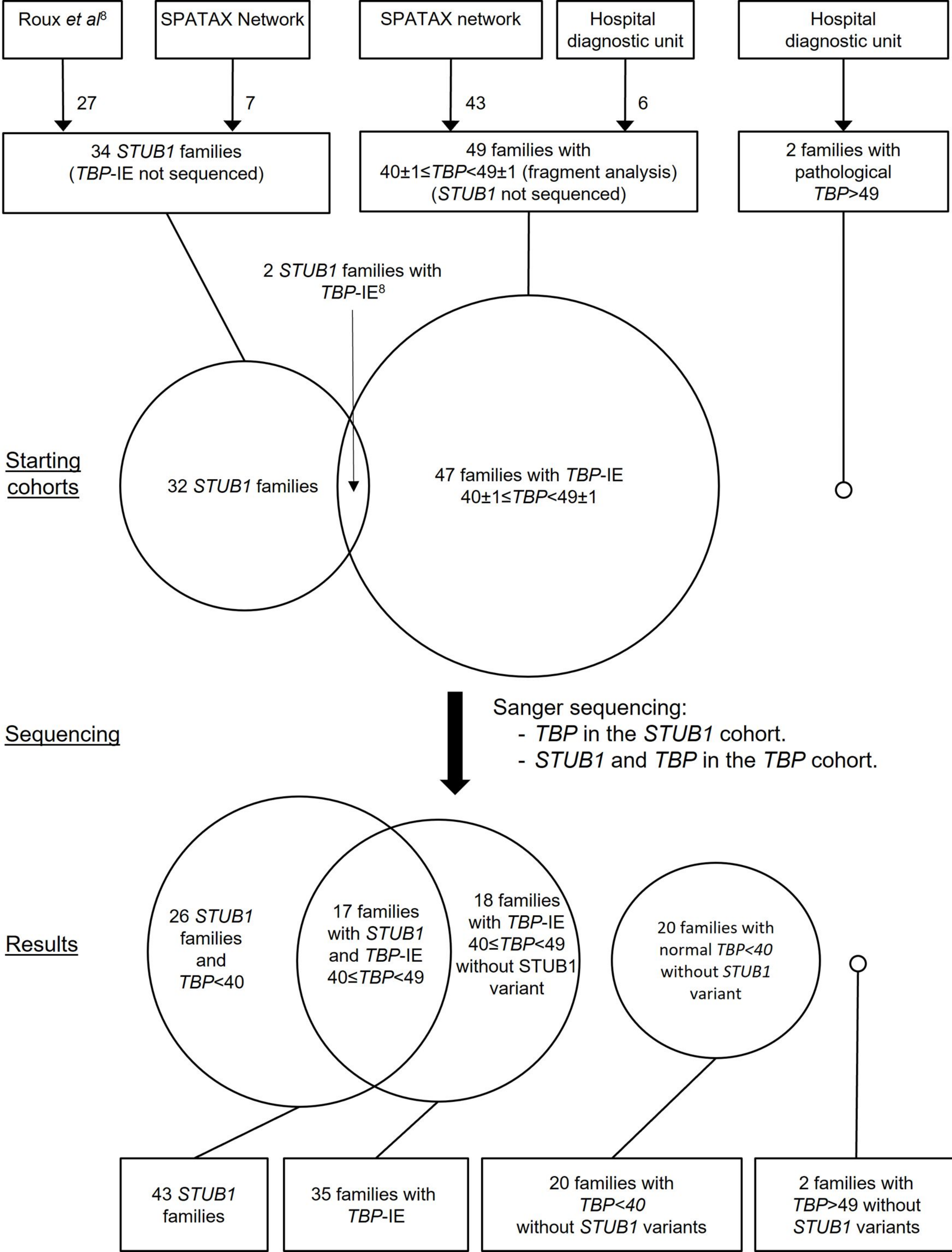
36. Kuddus R, Schmidt MC. Effect of the non-conserved N-terminus on the DNA binding activity of the yeast TATA binding protein. *Nucleic Acids Res*. 1993;21(8):1789-1796. doi:10.1093/nar/21.8.1789

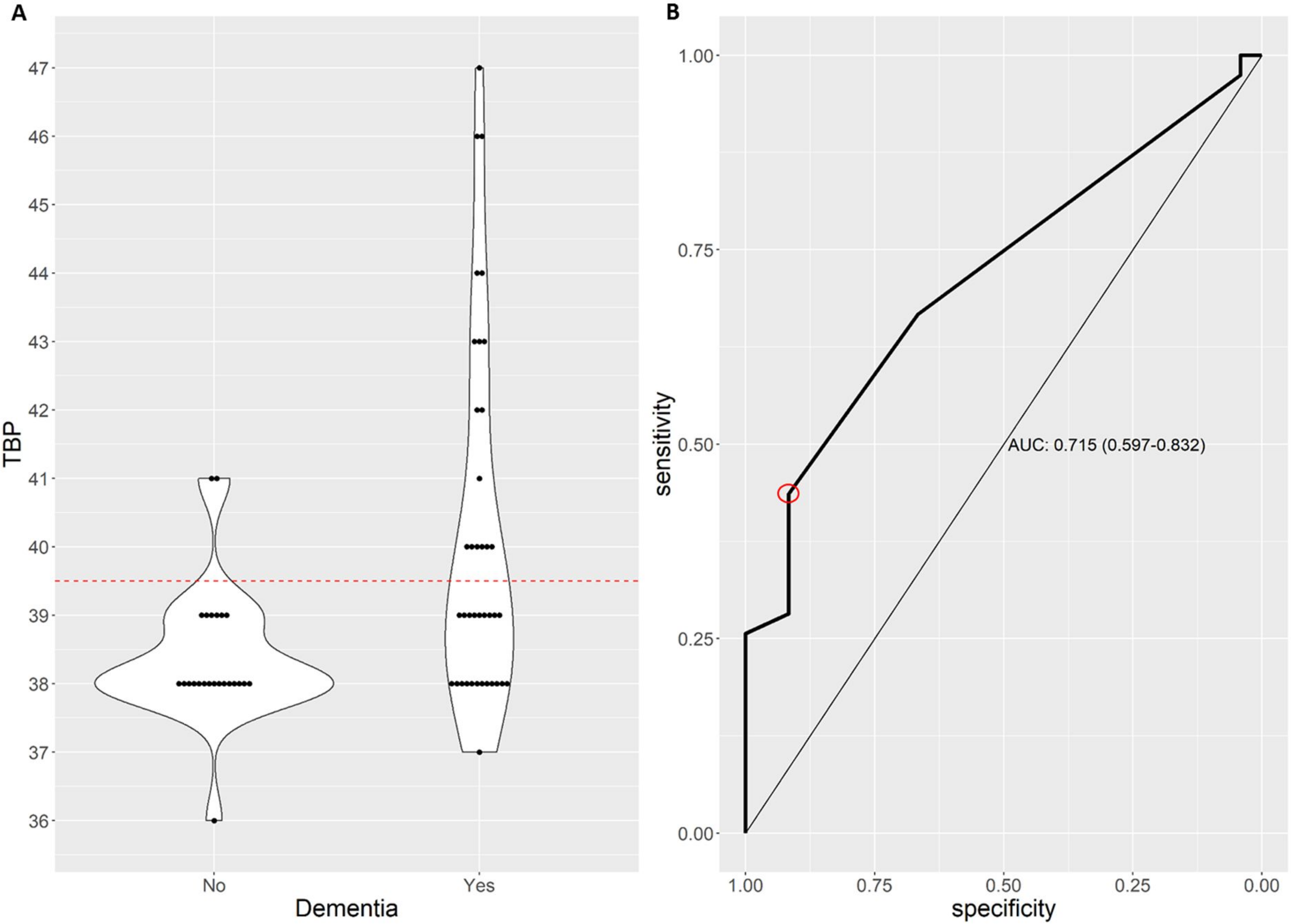
37. Bettencourt C, de Yébenes JG, López-Sendón JL, et al. Clinical and Neuropathological Features of Spastic Ataxia in a Spanish Family with Novel Compound Heterozygous Mutations in STUB1. *Cerebellum*. 2015;14(3):378-381. doi:10.1007/s12311-014-0643-7

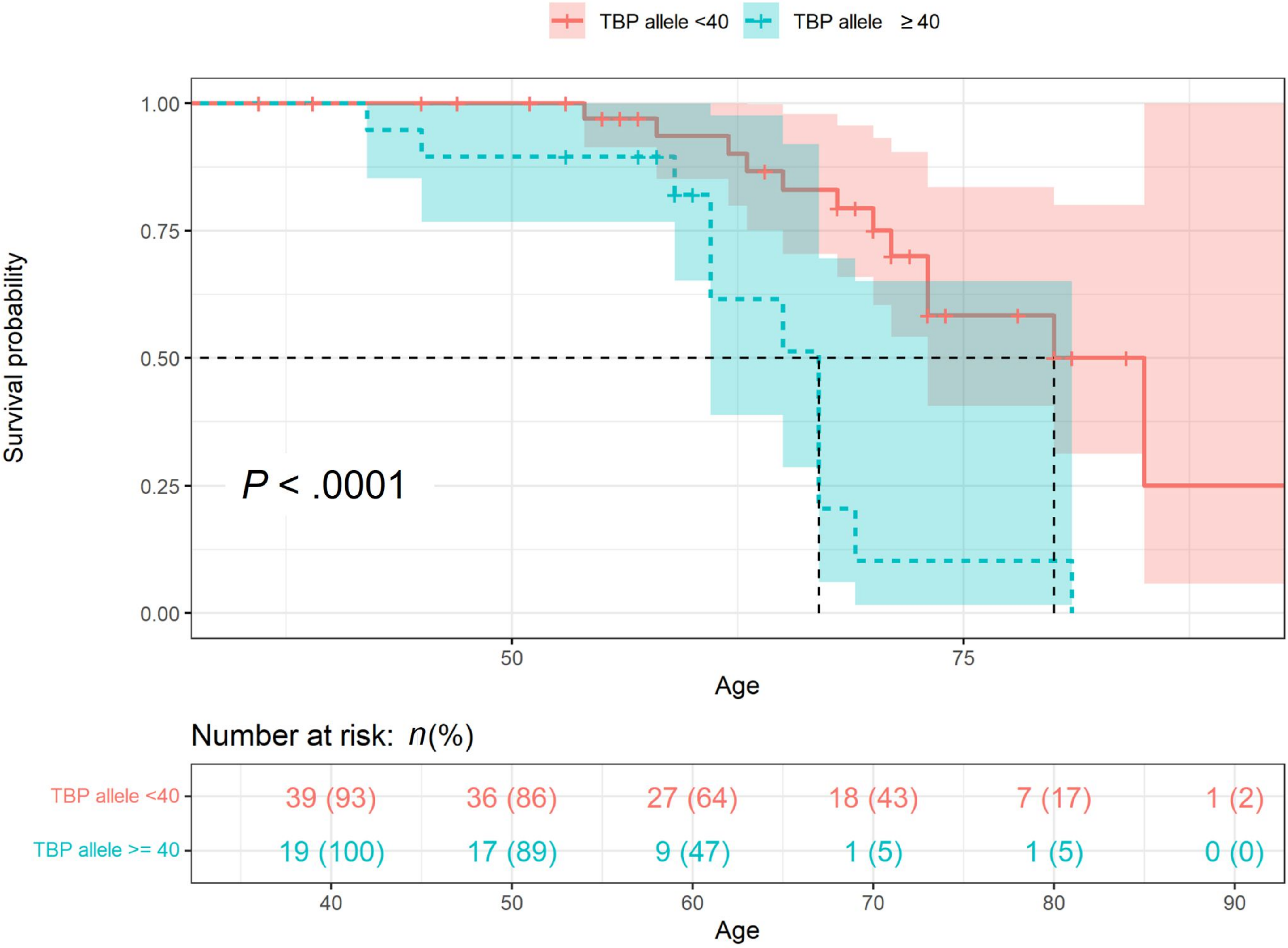
38. Chen DH, Latimer C, Yagi M, et al. Heterozygous STUB1 missense variants cause ataxia, cognitive decline, and STUB1 mislocalization. *Neurol Genet*. 2020;6(2):1-13. doi:10.1212/NXG.0000000000000397

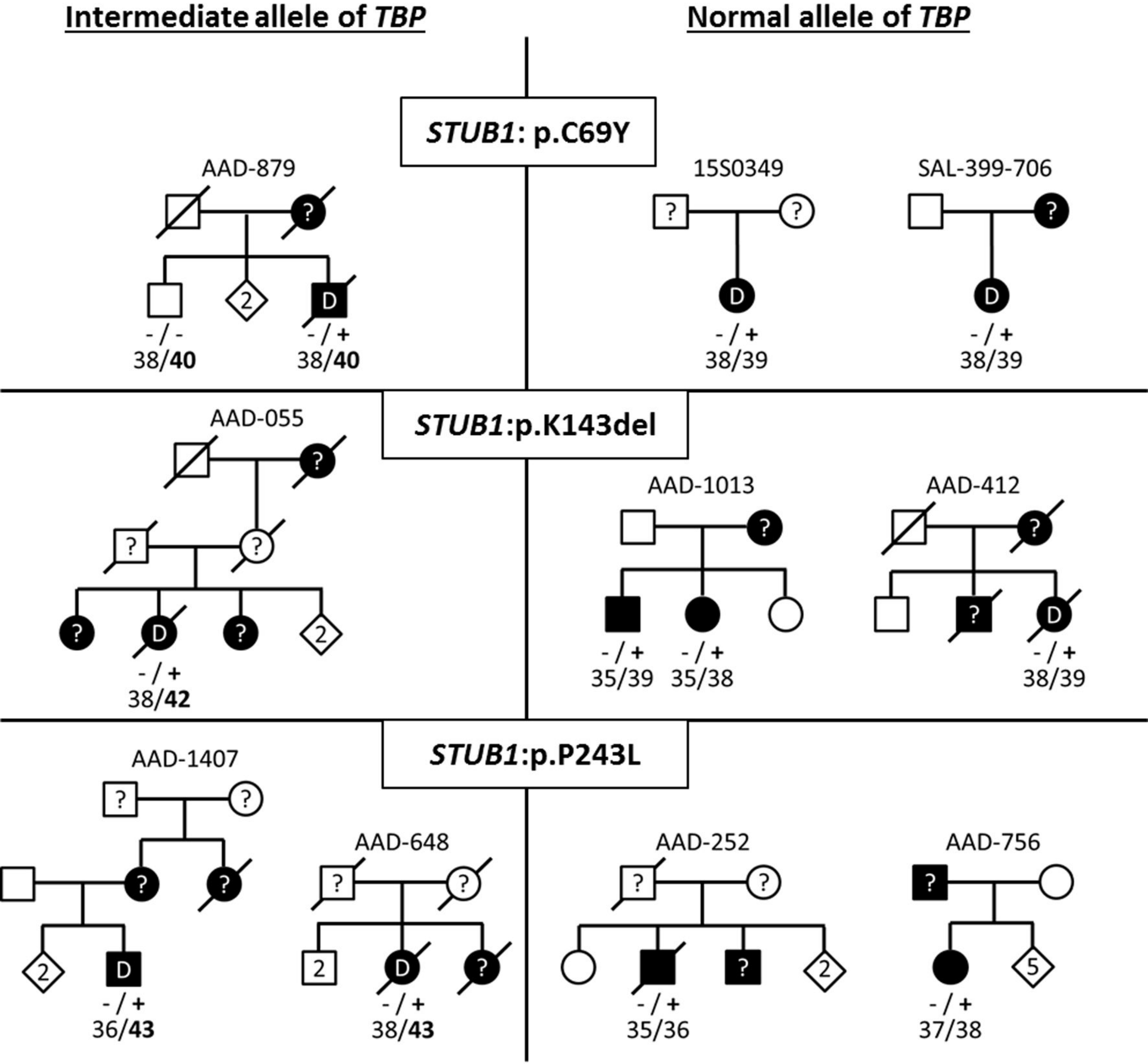
39. Mol MO, van Rooij JGJ, Brusse E, et al. Clinical and pathologic phenotype of a large family with heterozygous STUB1 mutation. *Neurol Genet*. 2020;6(3):e417. doi:10.1212/NXG.0000000000000417

40. Barman P, Sen R, Kaja A, et al. Genome-Wide Regulations of the Preinitiation Complex Formation and Elongating RNA Polymerase II by an E3 Ubiquitin Ligase, San1. *Mol Cell Biol*. 2022;42(1):e0036821. doi:10.1128/MCB.00368-21


Roux *et al*[8]
SPATAX Network
SPATAX network
Hospital diagnostic unit
Hospital diagnostic unit
27
7
43
6
34 *STUB1* families (*TBP*-IE not sequenced)
49 families with 40±1≤*TBP*<49±1 (fragment analysis) (*STUB1* not sequenced)
2 families with pathological *TBP*>49
2 *STUB1* families with *TBP*-IE[8]
Starting cohorts
32 *STUB1* families
47 families with *TBP*-IE 40±1≤*TBP*<49±1
Sequencing
Sanger sequencing:
- *TBP* in the *STUB1* cohort.
- *STUB1* and *TBP* in the *TBP* cohort.
Results
26 *STUB1* families and *TBP*<40
17 families with *STUB1* and *TBP*-IE 40≤*TBP*<49
18 families with *TBP*-IE 40≤*TBP*<49 without STUB1 variant
20 families with normal *TBP<40* without *STUB1* variant
43 *STUB1* families
35 families with *TBP*-IE
20 families with *TBP<40* without *STUB1* variants
2 families with *TBP*>49 without *STUB1* variants


A
TBP
47
46
45
44
43
42
41
40
39
38
37
36
No
Yes
Dementia
B
sensitivity
1.00
0.75
0.50
0.25
0.00
AUC: 0.715 (0.597-0.832)
1.00
0.75
0.50
0.25
0.00
specificity


TBP allele <40
TBP allele ≥ 40
Survival probability
1.00
0.75
0.50
0.25
0.00
P < .0001
50
75
Age
Number at risk: n(%)
TBP allele <40
TBP allele >= 40
39 (93)
36 (86)
27 (64)
18 (43)
7 (17)
1 (2)
19 (100)
17 (89)
9 (47)
1 (5)
1 (5)
0 (0)
40
50
60
70
80
90
Age


Intermediate allele of TBP
Normal allele of TBP
STUB1: p.C69Y
AAD-879
?
2
D
- / -
38/40
- / +
38/40
15S0349
?
?
D
- / +
38/39
SAL-399-706
?
D
- / +
38/39
STUB1:p.K143del
AAD-055
?
?
?
?
D
?
2
- / +
38/42
AAD-1013
?
- / +
35/39
- / +
35/38
AAD-412
?
?
D
- / +
38/39
STUB1:p.P243L
AAD-1407
?
?
?
?
2
D
- / +
36/43
AAD-648
?
?
2
D
?
- / +
38/43
AAD-252
?
?
?
2
- / +
35/36
AAD-756
?
5
- / +
37/38