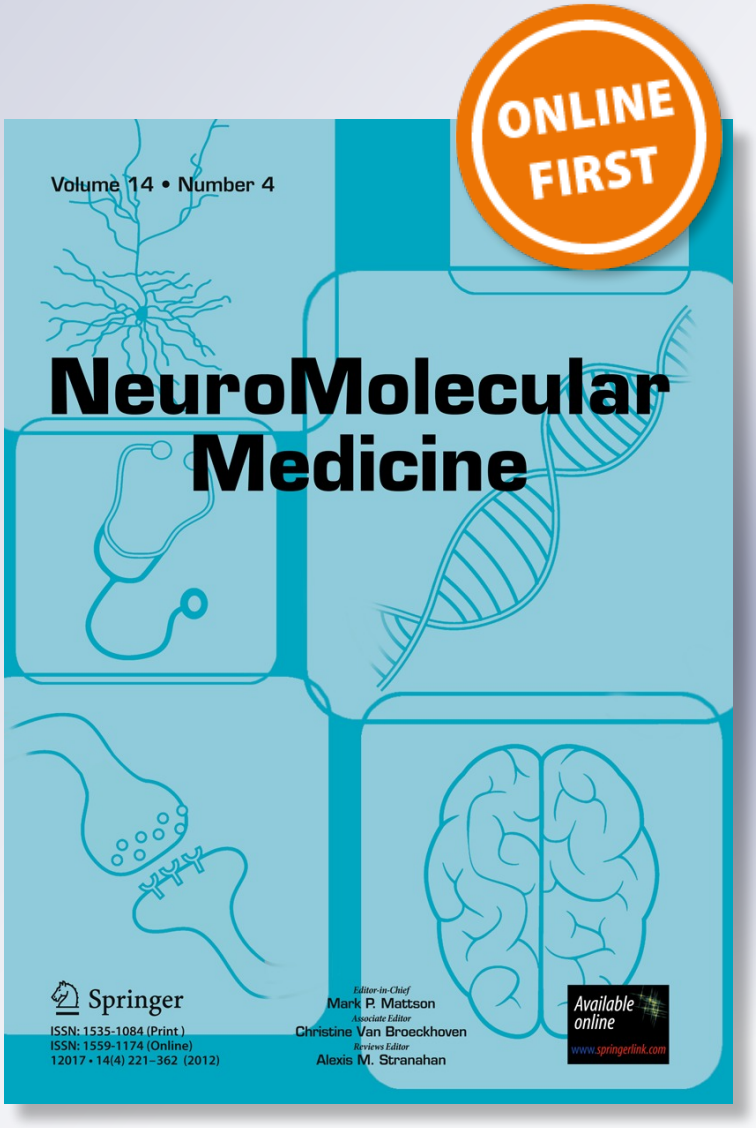
Novel Mutations in Cyclin-Dependent Kinase-Like 5 (CDKL5) Gene in Indian Cases of Rett Syndrome

Dhanjit Kumar Das, Bhakti Mehta, Shyla R. Menon, Sarbani Raha & Vrajesh Udani

NeuroMolecular Medicine

ISSN 1535-1084

Neuromol Med
DOI 10.1007/s12017-012-8212-z



Springer

Neuromol Med
DOI 10.1007/s12017-012-8212-z

ORIGINAL PAPER

Novel Mutations in Cyclin-Dependent Kinase-Like 5 (*CDKL5*) Gene in Indian Cases of Rett Syndrome

Dhanjit Kumar Das · Bhakti Mehta · Shyla R. Menon · Sarbani Raha · Vrajesh Udani

Received: 27 June 2012 / Accepted: 3 December 2012

Abstract Rett syndrome is a severe neurodevelopmental disorder, almost exclusively affecting females and characterized by a wide spectrum of clinical manifestations. Both the classic and atypical forms of Rett syndrome are primarily due to mutations in the methyl-CpG-binding protein 2 (*MECP2*) gene. Mutations in the X-linked cyclin-dependent kinase-like 5 (*CDKL5*) gene have been identified in patients with atypical Rett syndrome, X-linked infantile spasms sharing common features of generally early-onset seizures and mental retardation. *CDKL5* is known as serine/threonine protein kinase 9 (STK9) and is mapped to the Xp22 region. It has a conserved serine/threonine kinase domain within its amino terminus and a large C-terminal region. Disease-causing mutations are distributed in both the amino terminal domain and in the large C-terminal domain. We have screened the *CDKL5* gene in 44 patients with atypical Rett syndrome who had tested negative for *MECP2* gene mutations and have identified 6 sequence variants, out of which three were novel and three known mutations. Two of these novel mutations p.V966I and p.A1011V were missense and p.H589H a silent mutation. Other known mutations identified were p.V999M, p.Q791P and p.T734A. Sequence homology for all the mutations revealed that the two mutations (p.Q791P and p.T734A) were conserved across species. This indicated the importance of these residues in structure and function of the protein. The damaging effects of these mutations were analysed *in silico* using PolyPhen-2 online software. The PolyPhen-2 scores of p.Q791P and p.T734A were 0.998 and 0.48, revealing that these mutations could be deleterious and might have potential functional effect. All other mutations had a low score suggesting that they might not alter the activity of CDKL5. We have also analysed the position of the mutations in the CDKL5 protein and found that all the mutations were present in the C-terminal domain of the protein. The C-terminal domain is required for cellular localization through protein–protein interaction; any mutations in this domain might alter this function of the protein. This is the first report from India showing the mutation in *CDKL5* gene in Indian cases of Rett syndrome. Our study emphasizes the role of *CDKL5* mutation screening in cases of atypical Rett syndrome with congenital seizure variant.

Keywords *CDKL5* · Rett syndrome · RTT · *STK9* · Epilepsy

Introduction

Rett syndrome (RTT) is an X-linked dominant severe neurodevelopmental disorder, affecting almost exclusively girls. In the classic form, after a period of normal development, patients show growth retardation and regression of speech, along with loss of purposeful hand use and appearance of stereotyped hand movements. RTT variants have been described, including the "preserved speech variant (PSV)," characterized by the recovery of some degree of

Electronic supplementary material The online version of this article (doi:10.1007/s12017-012-8212-z) contains supplementary material, which is available to authorized users.

D. K. Das (✉) · B. Mehta · S. R. Menon
Genetic Research Centre, National Institute for Research in Reproductive Health (ICMR), Jahangir Merwanji Street, Parel, Mumbai 400 012, India
e-mail: dasd@nirrh.res.in

S. Raha · V. Udani
Department of Pediatric Neurology, Hinduja National Hospital and Research Centre, Mahim, Mumbai 400 016, India

Published online: 15 December 2012

speech; the "congenital variant" (recognized from birth); the "early seizure variant" (seizure onset before regression); and the "forme fruste," with a milder, incomplete clinical course (regression between 1 and 3 years) (Hagberg and Skjeldal 1994).

A very few cases of familial RTT made it initially difficult to determine the mode of inheritance of this disorder, but the virtual absence of affected males suggested an X-linked dominant inheritance pattern. Amir et al. (1999) showed that mutations in the methyl-CpG-binding protein 2 gene (*MECP2*) located in Xq28 are the primary cause of classic form of RTT. Mutations in *MECP2* account for more than 95 % of patients with classic RTT but have been found in only 20–40 % of atypical RTT patients (Cheadle et al. 2000). In our earlier study, we have identified a total of 19 different *MECP2* sequence variants in 27 patients. Of the 19 mutations, 4 mutations were found in atypical RTT (unpublished data). *CDKL5* defects have been found in some patients with the Hanefeld variant (Hanefeld 1985), a congenital form of RTT characterized by the presence of intractable seizures during the first months of life (Scala et al. 2005; Tao et al. 2004; Weaving et al. 2004). Moreover, mutations or chromosomal translocations involving *CDKL5* have also been identified in patients with infantile spasms associated with mental retardation and in West syndrome patients (Archer et al. 2006; Van Esch et al. 2007).

The human *CDKL5* gene encodes a protein of 1030 amino acids with an N-terminal catalytic domain highly homologous to members of the mitogen-activated protein (MAP) kinase and cyclin-dependent (CDK) kinase families (Montini et al. 1998). The protein is rather uncharacterized but its involvement in RTT has been explained by the fact that this kinase seems to function in a molecular pathway common to that of MeCP2 (Bertani et al. 2006 and Mari et al. 2005). In fact, the two proteins form a protein complex in vivo and the catalytic activity of CDKL5 mediates the phosphorylation of MeCP2 in vitro. Moreover, the two genes share temporal and spatial expression patterns in the brain and are simultaneously activated during neuronal maturation (Mari et al. 2005). According to the current model, CDKL5 works upstream of MeCP2 influencing directly or indirectly its phosphorylation state and thereby specific functions of MeCP2 (Rusconi et al. 2008). In the absence of CDKL5, these phosphorylation-dependent activities of MeCP2 would be altered causing a subset of Rett symptoms. In addition, other as yet non-discovered targets of the kinase would also be deregulated, leading to the specific phenotype associated with *CDKL5* mutations. In agreement with this model, so far, RTT-causing mutations in *CDKL5* have been found only in patients with the Hanefeld variant and never associated with classic RTT. The fact that the autonomous nervous system, whose malfunction in patients with *MECP2* mutations causes severe respiratory problems appears to be better preserved in patients with CDKL5 deficits. This reinforces the separate functions of the two proteins (Rosas-Vargas et al. 2008).

Mutations in *CDKL5* gene are located within the protein kinase domain and affect highly conserved amino acids; this strongly suggests that impaired CDKL5 catalytic activity plays an important role in the pathogenesis of Rett syndrome (Tao et al. 2004). More than 70 different point mutations have been described including missense mutations within the catalytic domain, nonsense mutations causing the premature termination of the protein distributed in the entire open reading frame, splice variants, and frameshift mutations (Bahi-Buisson and Bienvenu 2012). But no mutations were reported from Indian patients with Rett syndrome. Recently, Bahi-Buisson et al. (2012) showed the genotype-phenotype correlation of eight recurrent *CDKL5* mutations. They found that mutations in the kinase domain (such as p.Arg59X, p.Arg134X, p.Arg178Trp/Pro/Gln, or c.145+2T>C) and frameshift mutations in the C-terminal region (such as c.2635_2636delCT) had a more severe phenotype with infantile spasms, refractory epileptic encephalopathy, absolute microcephaly, and inability to walk. Hadzsiev et al. (2011) identified two novel mutations in *CDKL5*, c.607G>T resulting in a termination codon at amino acid 203, disrupting the catalytic domain, and c.1708G>T leading to a stop at amino acid 570 position of the C-terminus. Therefore, the position of mutations in CDKL5 protein is important for clinicians to correlate with the clinical severity of the disorders.

The aim of this study was to determine *CDKL5* mutations and its relevance in the atypical cases of RTT where *MECP2* mutations were not found. In this paper, we report the identification of novel and known *CDKL5* mutation, reinforcing the link between the gene and RTT.

Patients and Methods

Patients

All the patients screened in this study were sporadic cases of Rett syndrome attending the Paediatric Neurology OPD of Hinduja Hospital, Mumbai. All the children were clinically assessed according to the criteria of the *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition) (DSM IV) and found to be atypical Rett disorder. Initially, all the patients were referred for *MECP2* mutations. We have sequenced the entire coding region of *MECP2* gene, and those found to be negative for mutations in *MECP2* were included in this study (n=44) to screen for

mutations in *CDKL5* gene. The study was approved by Institutional Ethics Committee for Clinical Studies, and samples were collected after obtaining written informed consent (Institutional Ethics approval No 168/2009).

Genomic DNA Isolation and PCR Amplification

Whole blood samples from above patients were collected in EDTA tubes. Genomic DNA was isolated from 2.0 ml of blood collected from the above patients using QIAmp DNA extraction kit (Qiagen, GmBH, Germany). After isolation, the integrity was checked by running on 0.8 % agarose gel electrophoresis.

PCR amplification was performed in 50 μl of 10 mM Tris HCl (pH 8.3) containing 50 mM KCl, 1.5 mM MgCl2, 200 mM of each dNTPs (dATP, dCTP, dGTP and dTTP), 20 pmol primers, 250 ng template DNA and 1.5 units of Taq DNA polymerase (MBI-Fermantas, MD). All the 21 exons were amplified separately using specific primers designed from the wild-type *CDKL5* sequence. PCR was cycled 35 times; each cycle consisted of denaturation for 1 min at 94 °C, annealing for 45 s at 56–59 °C and extension for 1 min at 72 °C. After amplification, 3 μl of PCR product was subjected to electrophoresis on 1 % agarose gel for 45 min at 100 V in TAE buffer and bands were stained with ethidium bromide (0.5 mg/ml).

Sequencing and Sequence Analysis

The PCR products were gel purified using QIAquick Gel extraction kit (QIAGEN, GmBH, Germany) according to the manufacturer's protocol. The gel-purified PCR products were sequenced using gene-specific primers on ABI PRISM 3130xl version 3.1 DNA Sequencer (Applied Biosystems, Foster City, CA, USA). The sequences were analysed for the presence of mutations using Lasergen program (DNASTAR, Inc, Madison, USA). The mutations were screened based on reference sequence of *CDKL5* gene (Accession No. NM_003159). Multiple sequence alignment was carried out using MegAlign program of DNASTAR, and WebLogo was created using online software (http://weblogo.berkeley.edu/ WebLogo: A Sequence Logo Generator). The sequences used for alignment were retrieved from NCBI database and those are *Homo sapiens* (Accession No NP_001032420), *Pan troglodytes* (XP_003317430), *Pan paniscus* (XP_003819587), *Danio rerio* (NP_001124243), *Macaca mulatta* (XP_002808564), Callithrix jacchus (XP_002762761), *Saimiri boliviensis boliviensis* (XP_003924179), *Bos Taurus* (XP_002700425), *Equus caballus* (XP_001491126), *Sus scrofa* (XP_003360287), *Canis lupus familiaris* (XP_548881), *Gallus gallus* (XP_425571), *Oreochromis niloticus* (XP_003439841), *Oryctolagus cuniculus* (XP_002719977), *Rattus norvegicus* (ACY73182) and *Mus musculus* (NP_001019795). Ab initio structure modelling of MeCP2 protein was carried out using two popular algorithms freely available online, namely I-Tasser (Roy et al. 2010) and Bhageerath (Jayaram et al. 2012).

X Chromosome Inactivation (XCI)

X inactivation ratios were determined as described by Allen et al. (1992). Briefly, separate aliquots of DNA were digested using the methylation-sensitive restriction enzymes HpaII and HhaI (New England Biolabs). The polymorphic repeat at the androgen receptor locus was then amplified using fluorescent labelled PCR primers, and the allele peak areas were analysed using an ABI PRISM 3130xl version 3.1 DNA Sequencer and Genotyper software (Applied Biosystems).

Results

In this study, we have screened 44 cases with features of atypical Rett syndrome for presence of mutations in the *CDKL5* gene and were negative for *MECP2* gene mutation. Patients were classified based on clinical symptoms according to DSM-IV criteria.

Sequencing and Sequence Analysis

A total of 6 sequence *CDKL5* variants in five patients were identified, out of which three were novel and three known mutations. Novel mutation p.V966I (Fig. 1a) located at exon 20 of *CDKL5* gene was identified in patient 1. Multiple sequence alignment revealed that this mutation is conserved with the *CDKL5* sequence of *Pan troglodytes*, *Pan paniscus* and *Macaca mulatta.* The other novel mutation p.A1011V (Fig. 1b) located at exon 21 found to be present in patient 2 and is located at C-terminal domain of the protein. Upon multiple sequence alignment, this mutation was also found to be conserved with *Pan troglodytes*, *Pan paniscus* and *Danio rerio.*

The patient 3 had two mutations: one known (p.V999M) (Fig. 1c) and the other novel mutation (p.H589H) (Fig. 1d). The mutation p.V999M is located at exon 21, and the silent mutation p.H589H was found to be present at exon 12 of *CDKL5* gene. Upon sequence analysis, it was evident that the mutation p.H589H was found to be conserved with its homologues available across all species. Analysis of p.V999M mutation showed the sequence conservation among *Pan troglodytes*, *Pan paniscus* and *Macaca mulatta.*

In patient 4, we have identified p.Q791P mutation located at exon 16 of *CDKL5* gene (Fig. 1e). Another mutation p.T734A located at exon 15 has been identified in

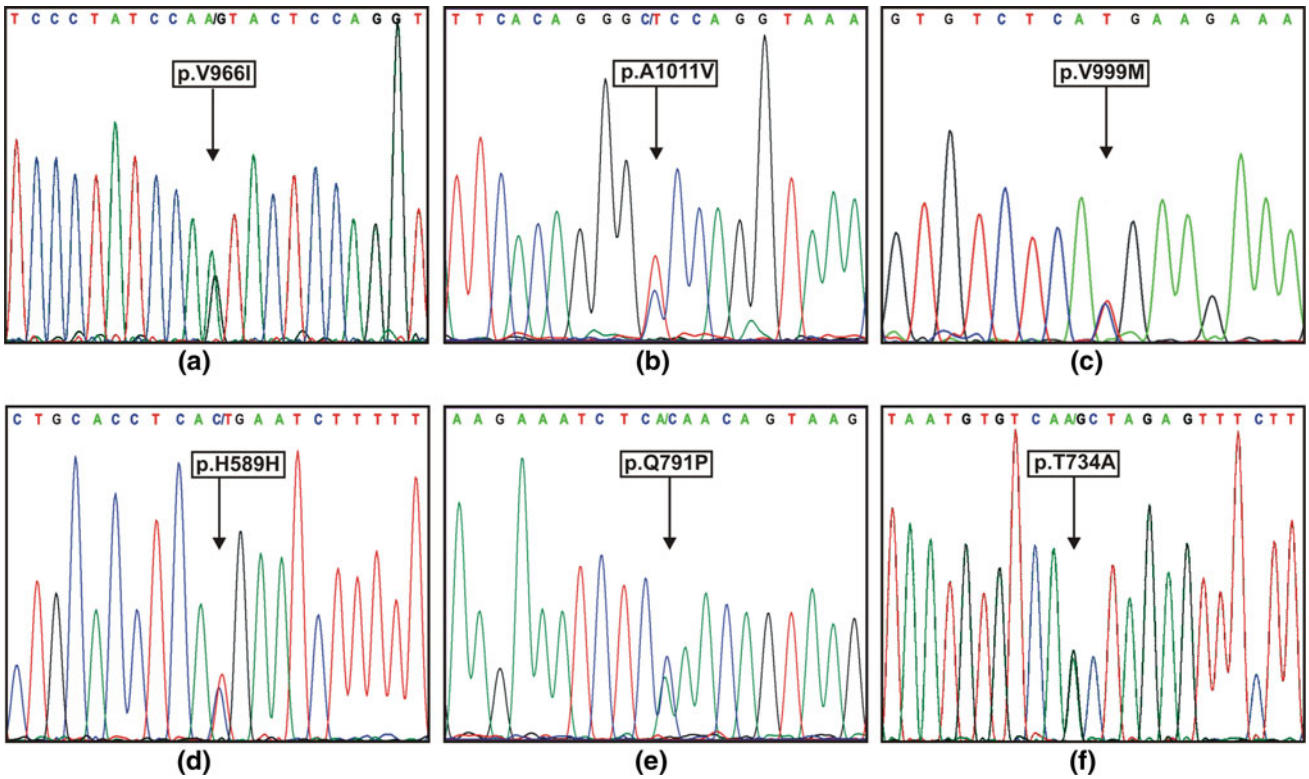


Fig. 1 DNA sequence chromatogram showing the presence of mutations. **a** p.V966I mutation in exon 20 of *CDKL5* gene. Since the mutation is heterozygous, two peaks at location indicated by an *arrow* corresponding to two nucleotides, which are G (wild type) and A (mutant). **b** p.A1011V mutation in exon 21 of patient 2. This is also a heterozygous mutation (C for wild changed to T for mutant allele). **c** p.V999M mutation in exon 21 of patient 3 (C for wild changed to T for mutant allele). **d** p.H589H mutation in exon 12 of patient 3 (C for wild changed to T for mutant allele). **e** p.Q791P mutation in exon 16 of patient 4 (A for wild changed to C for mutant allele). **f** p.T734A mutation in exon 15 of patient 5 (A for wild changed to G for mutant allele)

patient 5 (Fig. 1f). Multiple sequence alignment of both these two mutations revealed that these were well conserved across species (Fig. 2). The WebLogo was also created using online software (http://weblogo.berkeley.edu/ WebLogo: A Sequence Logo Generator). It was observed that the p.Q791P and p.T734A were found to be well conserved among sequences of all the species (Fig. 3).

Parent samples of all the mutation-positive patients were analysed, but none were found to have mutation in *CDKL5* gene. We also sequenced 100 control samples for presence of three novel mutations to rule out polymorphism. But we have not been identified these three novel mutations in the control population.

Structure Analysis of CDKL5 Protein

We also tried to analyse the functional effect of novel mutation using *in-silico* approach. However, the entire structure of CDKL5 protein has not yet been experimentally elucidated; only the structure of kinase domain which spans N-terminus of the protein is available. We tried to generate a theoretical model for the entire protein. There were no templates identified for the C-terminal region, hence comparative homology modelling could not be carried out. De novo modelling was also attempted using two popular algorithms, namely I-Tasser (Roy et al. 2010) and Bhageerath (Jayaram et al. 2012). Due to the absence of suitable templates, de novo modelling failed to predict a suitable structure (supplementary Fig).

Clinical Description of Mutation-Positive Patients

Patient 1 (p.V966I)

She was born to non-consanguineous parents following a normal pregnancy. She was presented with loss of acquired speech and motor hand skill with infrequent seizure. After treatment with antiepileptic drug, seizures were under controlled. Developmental history revealed the delay in head holding and sitting. On clinical examination, the head circumference (44.5 cm) fell below 3rd percentile at her 2½ years of age. Hand stereotypies started at the age of

1 year. She also had breathing difficulties like infrequent hyperventilation. She had showed lack of attachment with family members, and also cry or smile without any apparent reason. Neuroimaging was normal, and I.Q. assessment revealed presence of severe degree of mental retardation.

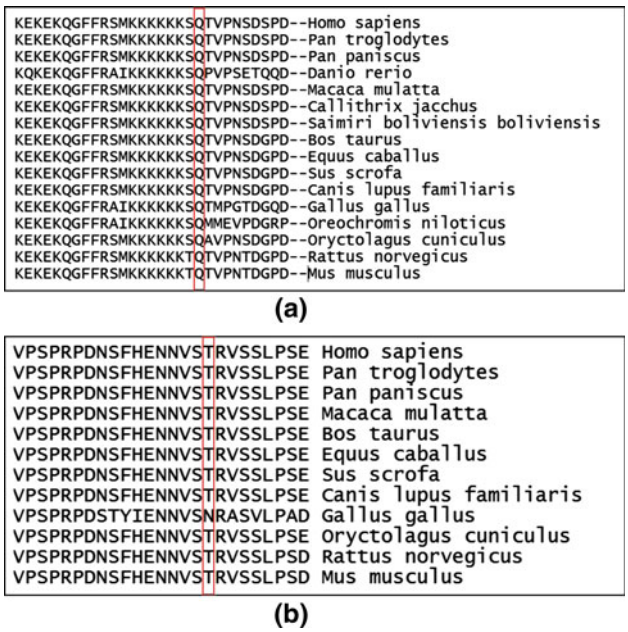


Fig. 2 Multiple sequence alignment with homologous sequence from different species showing the extent of conservation of the mutation site. **a** Multiple sequence alignment of p.Q791P mutation showing the conserved glutamine (Q) residue indicated by *box* across species. **b** Multiple sequence alignment of p.T734A mutation showing the conserved threonine (T) residue indicated by *box* across species

Patient 2 (p.A1011V)

She was born at term following a normal pregnancy with a birth weight of 3.5 kg and head circumference was 35.5 cm (75th centile). She was a placid and sleepy baby and showed milder phenotype. There were no problems during neonatal period, but there were global developmental delays. Stereotypic hand movement was started at the age of 1½ years. But the patient had no epilepsy, which was characteristic of *CDKL5* mutation reported by Castren et al. (2010).

Patient 3 (p.V999M and p.H589H)

Clinically, she was placid, sleepy baby and head circumference was 33 cm (2nd percentile) at the time of presentation. Her milestones were delayed and hands always closed as an infant. She was noticed to have become quite lethargic with diminished social and emotional reciprocity to her mother. Seizures were noted at the age of 4 weeks, and these were not infantile spasm. Clinically, she fulfilled the criteria of atypical RTT (early seizure variant).

Patient 4 (p.Q791P)

She was born following a normal pregnancy and delivery at 37 weeks. There were no problems during the neonatal period, but she was very placid and sleepy. There were no concerns about her social interaction in the first 6 months, however; subsequently, autistic features became apparent. Hand stereotypies, hand wringing, clapping and mouthing appeared at 1½ years. She developed a mixed seizure

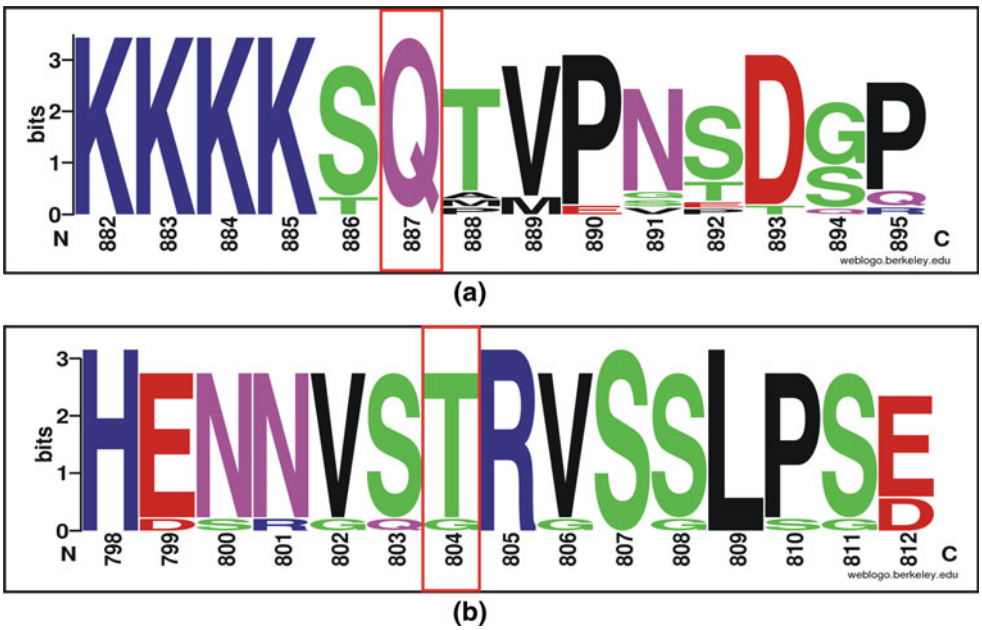


Fig. 3 WebLogo showing the conservation of the mutation. **a** p.Q791P mutation showing the conserved glutamine (Q) residue indicated by *box*. **b** p.T734A mutation showing the conserved threonine (T) residue indicated by *box*

disorder with generalized tonic–clonic and myoclonic jerks at the age of 8 months. Seizure was under control with medication. No dysmorphic features were noted. Clinically, she also fulfilled the criteria for atypical RTT.

Patient 5 (p.T734A)

The child was born to consanguineous parents. She was born following a normal pregnancy, birth weight was 2.5 kg and head circumference was 32 cm (less than 5th centile). She was normal till 2 months of age. She had recurrent episodes of facial grimacing at 2 months of age. She was then hospitalized for 4 days, and under antiepileptic drug, the frequency of seizures was minimized. Her milestone was delayed; she was able to sit at the age of 9 months. No stereotypic hand movement was noticed but she uses her hand to a minimum. She had a good eye contact during 6 months of age but later on she developed poor eye contact. At the age of 1 year, she had severe gastro-oesophageal reflux and was predominantly fed by gastrostomy tube. Her EEG showed consistent hypsarrhythmia and MRI brain revealed atrophy of basal ganglia. She fulfil the clinical characteristics of atypical Rett syndrome.

X Chromosome Inactivation (XCI) Status

X chromosome inactivation testing was completed in 35 of the 43 patients. After excluding the 8 who were homozygous, 11 % (4 out of 35) were shown to have skewed XCI. We did not find any significant relation between mutation type and XCI, using logistic transformation of the percentage of the smaller allele present. The only XCI result of importance was associated with patient 3 who had two sequence variants of *CDKL5* gene. This patient had more skewing (where skewing was defined as greater than 80 % of one X allele present), and clinical presentation did not have any significance on the type of mutation (Fig. 4). There were no significant relations found between any other mutations and phenotypic presentation.

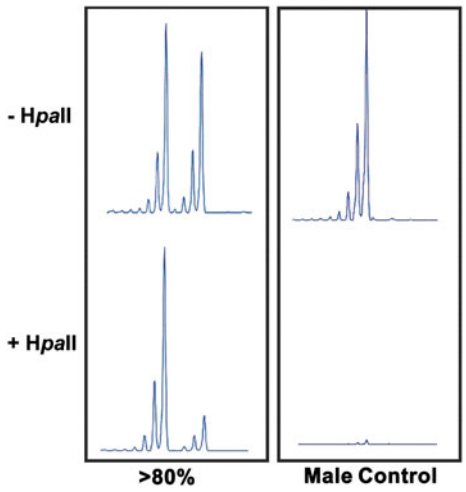


Fig. 4 Pattern of X chromosome inactivation (XCI) in blood sample of RTT patients. The *top* figure in each *box* represents the undigested DNA (−H*pa*II) from female RTT patients and from a male control; the *bottom* figure of each *box* represents DNA digested with H*pa*II prior to PCR (+HpaII). The relative intensity of the two alleles after digestion represents the XCI pattern for each individual (expressed as ratio and normalized to the undigested samples). Skewed inactivation is considered as at least 80 % inactivation of one allele. The *bottom* figure in male disappears, representing complete digestion of the unmethylated allele on the active X chromosome

Discussion

The X-linked cyclin-dependent kinase (CDK)-like 5 gene (CDKL5, OMIM 300203), formerly known as the STK9 gene (Montini et al. 1998), encodes a serine/threonine kinase that has been associated with the X-linked infantile spasm syndrome (ISSX, OMIM 308350) and the early-onset seizure variant of Rett syndrome (Evans et al. 2005). The phenotypes associated with *CDKL5* mutations range from X-linked infantile spasms (ISSX) and infantile epileptic encephalopathy to atypical RTT.

CDKL5 is a large protein of 1030 amino acids with an estimated molecular mass of 116 kDa containing a conserved serine/threonine kinase domain within its NH2 terminus and a large COOH-terminal region. The N-terminal domain (1–297 amino acid) contains ATP-binding domain (14–47 amino acid) and serine/threonine kinase domain (127–144 amino acid). The C-terminal domain (298–1030 amino acid) spans two-third of the protein. It has a signal peptidase serine active site, GTSMCPTL (amino acid 971–978). C-terminal regulates cellular localization, catalytic activity and stability of the protein (Bertani et al. 2006).

All patients in this study had initially been referred for *MECP2* mutation testing. Some *CDKL5* mutation-positive patients have been reported to have atypical RTT but none presented with classic RTT. However, there are some similarities between the most severe *CDKL5* phenotype and the male encephalopathic presentation of RTT (Villard et al. 2000).

We have identified a novel mutation p.V966I, which is present in the C-terminal domain of CDKL5 protein. We have used PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/) for predicting damaging effects of this missense mutation. This p.V966I mutation predicted to be begin with a score of 0.008; therefore, this mutation might not affect the activity of the protein. The other novel mutation p.A1011V was present in exon 21 of patient 2, and it was found to be present in C-terminal domain of the protein. PolyPhen-2 score was

found to be very low for this mutation. Therefore, this mutation might not affect the activity of the protein. Absence of these two novel mutations (p.V966I and p.A1011V) in control population (n=100) ruled out the possibility of polymorphism. Moreover, sequence analysis also showed that these two mutations were conserved among the homologues from ape family. This indicated the importance of these two residues on its structure and function of the protein.

The mutation p.V999M present in exon 21 of patient 3 changed valine to methionine. As both the amino acids are non-polar, this mutation may or may not affect the activity of CDKL5. PolyPhen score was found to be 0.275 that indicates that this mutation may affect the structure of the protein. The mutation was found to be present in C-terminal domain of CDKL5; therefore, this may affect the normal cellular localization, catalytic activity and stability of the C-terminus of the protein. In the same patient, another silent mutation (p.H589H) was found in exon 12. Since the silent mutation does not change the amino acid residue, it may not have any functional implication.

The mutation p.Q791P in exon 16 present in patient 4 was found to be present in the C-terminal domain. In this mutation, the glutamine is changed to proline. Glutamine is a polar amino acid changed to non-polar amino acid proline. Multiple sequence analysis also revealed that this mutation is well conserved across the species that indicates the importance of this residue for structure or function of the protein. We have predicted the damaging effect of this mutation using PolyPhen-2 that showed the presence of high score (0.998), means that the mutation might probably affect the function of the protein. This might be due to the change in polarity of the protein and it might affect the secondary structure of the protein. As the mutation is present in C-terminal domain, it might affect the cellular localization, catalytic activity and its stability.

Another missense mutation p.T734A located at C-terminal domain was found to be present in patient 5. In this mutation, threonine, which is polar amino acid, changed to non-polar neutral amino acid. Therefore, it may affect the structure of the protein that is important for its function. Sequence analysis showed that this mutation is well conserved across the species indicating the importance of this residue. The damaging effect of this mutation was determined by PolyPhen, and the score was found to be 0.48. This indicated that the mutation might have an important role in structure and function of the protein.

Effect of X Chromosome Inactivation

The inactivation of one of the X chromosome occurs randomly in differentiating embryonal cells in females, resulting in cells that are mosaic with respect to which chromosome is active. The purpose is to equalize X-linked gene products between XX females and XY males. Preferential inactivation of the X chromosome with mutated *MECP2* gene protects against the deleterious effects of mutations. Females carriers of RTT-causing mutations but asymptomatic or who suffer from only mild learning disability had a non-random XCI (Zoghbi et al. 1990). We have demonstrated the skewing XCI associated with patient 3 but there was no difference in clinical presentation. In this case, there might be a favourable XCI where the mutant allele might be present in the inactive X chromosome. In all other mutations, positive patients had a random X inactivation. Therefore, no significant relation between XCI and clinical severity was found.

In conclusion, we have identified 6 sequence variations in *CDKL5* gene. No *CDKL5* mutations have so far been found in classic RTT cases. Mutations in *CDKL5* are relatively common in patients whose seizures begin before the age of 6 months. The phenotypes of Rett syndrome have many differentials based on X chromosome inactivation pattern. However, a genetically based confirmed diagnosis would help in management and counselling. Although clinical severity of the affected individual may differ on the type of mutation present, they could contribute to a genotype-phenotype correlation. The molecular diagnosis approach described in this study can be a basis for further development in clinical use.

Acknowledgments We thank all the patients and their families for contribution to this study. The authors are also thankful to Director, NIRRH, for providing necessary facilities and Indian Council of Medical Research (ICMR) and Department of Science & Technology (SR/FT/LS-87/2010), Govt of India, for providing financial grant for the study. The help in computational analysis provided by Dr Susan Thomas; Saravanan & Nanda in sequencing are also acknowledged.

References

Allen, R. C., Zoghbi, H. Y., Moseley, A. B., Rosenblatt, H. M., & Belmont, J. W. (1992). Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *American Journal of Human Genetics, 51*(6), 1229–1239.

Amir, R. D., Van den Veyver, I. B., Wan, M., Tra, C. Q., Francke, U., & Zoghbi, H. Y. (1999). Rett syndrome is caused by mutations in X-linked MeCP2, encoding methyl CpG binding protein 2. *Nature Genetics, 23*(2), 185–188.

Archer, H. L., Evans, J., Edwards, S., Colley, J., Newbury-Ecob, R., O'Callaghan, F., et al. (2006). *CDKL5* mutations cause infantile spasms, early onset seizures, and severe mental retardation in female patients. *Journal of Medical Genetics, 43*(9), 729–734.

Bahi-Buisson, N., & Bienvenu, T. (2012). CDKL5-related disorders: From clinical description to molecular genetics. *Molecular Syndromology, 2*(3–5), 137–152.

Bahi-Buisson, N., Villeneuve, N., Caietta, E., Jacquette, A., Maurey, H., Matthijs, G., et al. (2012). Recurrent mutations in the CDKL5 gene: Genotype-phenotype relationships. *American Journal of Medical Genetics Part A, 158A*(7), 1612–1619.

Bertani, I., Rusconi, L., Bolognese, F., Forlani, G., Conca, B., De Monte, L., et al. (2006). Functional consequences of mutations in CDKL5, an X-linked gene involved in infantile spasms and mental retardation. *Journal of Biological Chemistry, 281*(42), 32048–32056.

Castren, M., Gaily, E., Tengström, C., Lahdetie, J., Archer, H., & Ala-Mello, S. (2010). Epilepsy caused by CDKL5 mutations. *European Journal of Paediatric Neurology, 15*(1), 65–69.

Cheadle, J. P., Gill, H., Fleming, N., Maynard, J., Kerr, A., Leonard, H., et al. (2000). Long-read sequence analysis of the *MECP2* gene in Rett syndrome patients: Correlation of disease severity with mutation type and location. *Human Molecular Genetics, 9*(7), 1119–1129.

Evans, J. C., Archer, H. L., Colley, J. P., Ravn, K., Nielsen, J. B., Kerr, A., et al. (2005). Early onset seizures and Rett-like features associated with mutations in CDKL5. *European Journal of Human Genetics, 13*(10), 1113–1120.

Hadzsiev, K., Polgar, N., Bene, J., Komlosi, K., Karteszi, J., Hollody, K., et al. (2011). Analysis of Hungarian patients with Rett syndrome phenotype for MECP2, CDKL5 and FOXG1 gene mutations. *Journal of Human Genetics, 56*(3), 183–187.

Hagberg, B. A., & Skjeldal, O. H. (1994). Rett variants: A suggested model for inclusion criteria. *Pediatric Neurology, 11*(1), 5–11.

Hanefeld, F. (1985). The clinical pattern of the Rett syndrome. *Brain Development, 7*(3), 320–325.

Jayaram, B., Dhingra, P., Lakhani, B., & Shekhar, S. (2012). Bhageerath—Targeting the near impossible: Pushing the frontiers of atomic models for protein tertiary structure prediction. *Journal of Chemical Sciences, 124*(1), 83–91.

Mari, F., Azimonti, S., Bertani, I., Bolognese, F., Colombo, E., Caselli, R., et al. (2005). CDKL5 belongs to the same molecular pathway of MeCP2 and it is responsible for the early-onset seizure variant of Rett syndrome. *Human Molecular Genetics, 14*(14), 1935–1946.

Montini, E., Andolfi, G., Caruso, A., Buchner, G., Walpole, S. M., Mariani, M., et al. (1998). Identification and characterization of a novel serine-threonine kinase gene from the Xp22 region. *Genomics, 51*(3), 427–433.

Rosas-Vargas, H., Bahi-Buisson, N., Philippe, C., Nectoux, J., Girard, B., N'Guyen Morel, M. A., et al. (2008). Impairment of CDKL5 nuclear localisation as a cause for severe infantile encephalopathy. *Journal of Medical Genetics, 45*(3), 172–178.

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: A unified platform for automated protein structure and function prediction. *Nature Protocol, 5*(4), 725–738.

Rusconi, L., Salvatoni, L., Giudici, L., Bertani, I., Kilstrup-Nielsen, C., Broccoli, V., et al. (2008). CDKL5 expression is modulated during neuronal development and its subcellular distribution is tightly regulated by the C-terminal tail. *Journal of Biological Chemistry, 283*(44), 30101–30111.

Scala, E., Ariani, F., Mari, F., Caselli, R., Pescucci, C., Longo, I., et al. (2005). CDKL5/STK9 is mutated in Rett syndrome variant with infantile spasms. *Journal of Medical Genetics, 42*(2), 103–107.

Tao, J., Van Esch, H., Hagedorn-Greiwe, M., Hoffmann, K., Moser, B., Raynaud, M., et al. (2004). Mutations in the X-linked cyclin-dependent kinase-like 5 (CDKL5/STK9) gene are associated with severe neurodevelopmental retardation. *American Journal of Human Genetics, 75*(6), 1149–1154.

Van Esch, H., Jansen, A., Bauters, M., Froyen, G., & Fryns, J. P. (2007). Encephalopathy and bilateral cataract in a boy with an interstitial deletion of Xp22 comprising the CDKL5 and NHS genes. *American Journal of Medical Genetics Part A, 143*(4), 364–369.

Villard, L., Kpebe, A., Cardoso, C., Chelly, B., Tardieu, P., & Fontes, M. (2000). Two affected boys in a Rett syndrome family. *Neurology, 55*(8), 1188–1193.

Weaving, L. S., Christodoulou, J., Williamson, S. L., Friend, K. L., McKenzie, O. L. D., Archer, H., et al. (2004). Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. *American Journal of Human Genetics, 75*(6), 1079–1093.

Zoghbi, H. Y., Percy, A. K., Schultz, R. J., & Fill, C. (1990). Patterns of X chromosome inactivation in the Rett syndrome. *Brain Development, 12*(1), 131–135.