

Intellectual ability in Tuberous Sclerosis Complex correlates with predicted effects of mutations on TSC1 and TSC2 proteins

Ho Tin Wong,¹ Deborah L McCartney,² Julia C Lewis,² Julian R Sampson,²
Christopher J Howe,³ Petrus J de Vries⁴

¹ St James's University Hospital, Beckett Street, Leeds, West Yorkshire, LS9 7TF, UK, ²Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff, CF14 4XN, UK, ³Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UK, ⁴Division of Child & Adolescent Psychiatry, University of Cape Town, 46 Sawkins Road, Rondebosch, 7700, South Africa

Correspondence to: P J de Vries (petrus.devries@uct.ac.za)

ABSTRACT

Background

Tuberous sclerosis complex is a multisystem genetic disease, caused by mutation in the *TSC1* or *TSC2* genes, associated with many features, including intellectual disability (ID). We examined psychometric profiles of patients with *TSC1* or *TSC2* mutations and tested if different mutation types were associated with different degrees of intellectual ability.

Methods

One hundred subjects with known *TSC1/TSC2* mutations were assessed using a range of IQ (intellectual quotient) or DQ (developmental quotient) measures. Effects of mutations on *TSC1/TSC2* proteins were inferred from sequence data and published biochemical studies.

Results

Most individuals with *TSC1* mutations fell on a normal distribution identical to the general population, with ~10% showing profound ID. Of individuals with *TSC2* mutations, 34% showed profound ID, and the remainder a pattern of IQ/DQ more variable and shifted to the left than in *TSC1* or the general population. Truncating *TSC1* mutations were all predicted to be subject to nonsense-mediated mRNA decay. Mutations predicted to result in unstable protein were associated with less severe effects on IQ/DQ. There was a statistically significant negative correlation between length of predicted aberrant C-terminal tails arising from frameshift mutations in *TSC1* and IQ/DQ; for *TSC2* a positive but not statistically significant correlation was observed.

Conclusion

We propose a model where (i) IQ/DQ correlates inversely with predicted levels and/or deleterious biochemical effects of mutant TSC1 or TSC2 protein, and (ii) longer aberrant C-terminal tails arising from frameshift mutations are more detrimental for TSC1 and less for TSC2. Predictions of the model require replication and biochemical testing.

Key words

genotype-phenotype / tuberous sclerosis complex / tuberin / hamartin / nonsense-mediated decay

INTRODUCTION

Tuberous sclerosis complex (TSC) is a multisystem genetic disease with widespread manifestations including neuropsychiatric features such as intellectual disability (ID)[1,2]. It is an autosomal dominant disorder caused by mutations in the *TSC1* [3] or *TSC2* [4] genes. Roles attributed to TSC1 include stabilisation and correct localisation of TSC2 [5,6]. The TSC1-TSC2-TBC1D7 complex [7-9] modulates multiple intracellular processes, notably through TSC2's GTPase activating protein (GAP) domain [10]. The GAP activity of TSC2 counteracts stimulation by RAS-homologue enriched in brain (Rheb) of mammalian (or mechanistic) target of rapamycin complex 1 (mTORC1), which phosphorylates proteins including p70 S6 kinase (S6K1) [9,11]. TSC1 and TSC2 receive many inputs with activating or inhibitory effects on signalling [1], and loss of TSC1/TSC2 disturbs downstream signalling.

Intellectual disability (ID) is defined as an Intelligence Quotient (IQ) below 2 standard deviations of the mean (IQ/DQ <70), and associated functional impairment in daily life [12]. Although prevalence of ID in the general population is approximately 2.2%, epidemiological analysis of TSC in south-west England showed ID prevalence in TSC of 44.4% [13]. The pattern of ID was not a typical unimodal Gaussian distribution, but bimodal, with 70% of individuals on a normal distribution shifted slightly down (mean (μ) = 93.6, standard deviation (σ) =13.8) and 30% with 'profound' intellectual disability ($\mu < 20$) [13,14].

Until discovery of the roles of the TSC1-TSC2-TBC1D7 complex, it was assumed that the predominant causal contributions to intellectual and neurocognitive manifestations of TSC were from structural abnormalities (cortical tubers), seizure phenomena (epilepsy) or interaction between the two [15,16]. However, de Vries and Howe suggested that seizures and tubers were neither necessary nor sufficient to explain the neuropsychiatric variability of TSC [17]. For example, less than 50% of IQ variance was explained by seizure and tuber factors in the study by O'Callaghan et al [16]. Instead de Vries and Howe proposed TSC-associated neuropsychiatric disorders to be direct molecular consequences of disturbance of cell signalling/development [17 and refs therein]. Supporting this, animal models show a direct route from mutation to disruption of neurocognitive functions through mTOR overactivation [17-19]. There is also evidence for reversal of neurocognitive manifestations by mTOR inhibitors in animal models [19,20] and possibly in humans [21].

Consequences of mutations in *TSC2* are typically more severe than for *TSC1* [22-27]. De Vries and Howe proposed that, given that not all mutations were necessarily null, different mutations within *TSC1* or *TSC2* would have different consequences according to biochemical effects on the proteins [17]. This raised the possibility of using the nature of a given mutation to predict intellectual ability in TSC individuals, which could be scientifically and clinically valuable.

To date only three studies of genotype-phenotype associations in TSC utilised validated and standardised measures of intellectual ability/developmental level [24,27,28]. Lewis et al. [24] measured IQ of 92 individuals with TSC and known

mutations, performing a categorical genotype-phenotype analysis, dividing subjects into 'impaired' and 'non-impaired' based on IQ score below or above 70. They suggested *TSC2* mutations were more likely to be associated with ID than *TSC1* mutations. Jansen and colleagues performed comprehensive neuropsychological assessment of 59 TSC subjects [27]. The battery included measures of IQ, memory, attentional and executive skills and motor reaction times, and a 'cognition index' was calculated by summing scaled scores across neuropsychological domains. They concluded that *TSC1* and *TSC2* 'groups overlapped almost completely' and prediction of cognitive phenotype should not be based on whether an individual had a *TSC1* or *TSC2* mutation.

In the most recent genotype-phenotype study of IQ in TSC, van Eeghen et al. examined 137 subjects from a clinical case series [28], utilizing a range of standardised IQ/DQ measures. Rates of ID were 23% in those with *TSC1* and 59% in those with *TSC2* mutations, and several inter-mutational correlations were observed, some contrasting with earlier findings.

These studies had methodological weaknesses, particularly in the psychometric approach to IQ/DQ data. Intelligence is not a categorical phenomenon, being normally distributed across a population with $\mu = 100$ and $\sigma = 15$. Given that IQ/DQ distribution in TSC is bimodal, de Vries and Prather suggested that individuals in the profoundly impaired range (IQ<20) and those in the normal distribution be considered as separate cognitive phenotypes, perhaps representing different underlying mechanisms [14]. Statistically, it is inappropriate to perform parametric analysis with the TSC distribution pattern. Earlier studies [24,28] did not examine the actual shape

of IQ distribution across *TSC1* and *TSC2*, represented by mean and standard deviation data. *TSC2* may, for instance, have been reported as more severe than *TSC1* because most individuals with profound ID had a *TSC2* mutation. There are two main criticisms of the study by Jansen et al. [27]. In spite of the comprehensive neuropsychological battery, scores were collapsed into 4 categories. Given that a score of 1 represented performance < 70, more than 50% of the study data would have been in this category, preventing separation of mild, moderate, severe and profound intellectual disability. Secondly, Jansen's 'cognition index' is of unknown biological meaning, since neurobiological substrates of skills such as memory, attention, executive skills and motor function are highly different, and mutational or molecular correlates may be domain specific.

In van Eeghen et al's study only 44% of the clinic series had neuropsychological evaluations, possibly skewing the sample towards milder phenotypes as shown by the low proportion with profound ID [28]. No mean and standard deviation data were presented, limiting interpretation of the 'shape' of IQ/DQ distributions. Also the study did not consider familial/sporadic status (the latter being associated with a more severe phenotype [22]) resulting in a possible ascertainment bias.

Given the earlier proposal that different mutation types have different effects [17], we tested if different kinds of *TSC1* or *TSC2* mutations correlated with effects on IQ/DQ. We were interested in predicted effects of mutations on levels of mutant proteins through nonsense-mediated decay (NMD) or protein degradation (PD). NMD is a degradation-mediated reduction in levels of transcripts carrying premature termination codons (PTC), occurring if the PTC is shortly before the last exon-exon junction [29].

NMD has been demonstrated for *TSC1* but was shown to be incomplete in RNA isolated from blood, and from lymphoid and fibroblast cell lines [30], raising the possibility that mutant protein might be present in individuals with *TSC1* or *TSC2* mutations. We therefore tested for association between the predicted nature of any residual mutant protein and phenotype. PD is known to cause significant reduction in levels of TSC1 with mutations close to the N terminus [31-33].

We examined the association between genotype and intellectual phenotype, extending and reanalysing the sample used by Lewis et al. [24], and interpret the results in terms of a model for stability and effects of mutant proteins.

METHODS

Subjects

Ninety-two unrelated subjects were recruited and assessed as described previously (the ‘Cardiff’ study) [24]. An additional 21 unrelated subjects were volunteers from the Tuberous Sclerosis Association, UK, recruited for a separate (‘Cambridge’) study [34]. After confirmation of clinical criteria for TSC [35], subjects consented for their information to be used for research. Protocols received ethical approval. Only subjects with complete psychological and genetic profiles were included, excluding individuals with no mutation identified, no IQ/DQ score, or undetermined *TSC1/TSC2* status.

IQ/DQ assessments

In the Cardiff study, JCL assessed IQ using Wechsler Adult Intelligence Scale Revised (WAIS-R) [36], Wechsler Intelligence Scale for Children III (WISC-III) [37], Raven's Coloured Progressive Matrices (RCPM) [38] or the Vineland Adaptive Behaviour Scale (VABS) [39]. WAIS-R and WISC-III are standard measures of IQ suitable for individuals able to participate in direct testing. The floor of IQ measurement using these tools is ~ 45. Although WAIS-R and WISC-III are more reliable, they are unsuitable to determine IQs in individuals with lower abilities. In these cases the RCPM or VABS may be useful. RCPM is recognised as a measure of 'fluid intelligence' whereas VABS is an indirect measure of adaptive behaviours in daily life, used to estimate a developmental quotient (DQ) as proxy for IQ. In the Cambridge study, DLM used the Wechsler Abbreviated Scale of Intelligence (WASI), WISC-III or WAIS-R [40] to calculate full-scale IQ. Given the challenges of measurement across age and ability, we calculated a 'best estimate IQ/DQ' using a hierarchical approach (WAIS-R/WISC-III > WASI > RCPM > VABS). As per ICD-10 convention, ID was defined as IQ/DQ < 70 and profound ID as IQ/DQ < 20 [41].

Mutation detection and categorisation

Mutations were analysed by screening coding sequences and intronic splice site regions of *TSC1* and *TSC2*, using single strand conformation polymorphism analysis, denaturing high performance liquid chromatography, high resolution melt analysis and Sanger sequencing, and assessing both loci for large rearrangements using multiplex probe ligation analysis [42,43]. Mutations were categorised as missense,

nonsense, frameshift, in-frame deletions, large deletions, splice-site mutations and undetermined. Missense mutations were classified as non-conservative if they resulted in substitution between biochemical groups as defined by Clustal X [44]. Regions of TSC1 and TSC2 considered in mutation mapping are shown in Supplementary Figure 1. NMD was predicted if the PTC either generated directly by mutation, or resulting from a frameshift, was >18 codons upstream of the last exon-exon junction [29]. Junctions were as in the Leiden Open Variation Database (LOVD) *TSC1* and *TSC2* reference sequences [45]. Splice-site mutations, splice variants and possible effects on splicing of mutations away from splice sites were not considered, given the lack of detailed information on splicing. For frameshifts, the position of the next termination codon and the length of the ‘aberrant tail’ between it and the mutation site were determined using the LOVD sequences. To predict mutation distributions we assumed that each codon in *TSC1* and *TSC2* was equally likely to mutate to a PTC. Thus 190/1164 (16.3%) of possible *TSC1* truncation or frameshift mutations were predicted not to result in NMD, as the final exon-exon boundary occurs 172 codons prior to the 3’ end of the coding sequence, and 974/1164 (83.7%) were predicted to result in NMD. Similarly, 72/1807 (4.0%) of possible *TSC2* mutations were predicted not to result in NMD and 1735/1807 (96.0%) were predicted to result in NMD. For *TSC1* we carried out additional calculations using the reference sequence to estimate the actual number of positions at which a single insertion, deletion or substitution could lead to escape from NMD. We assumed insertions and deletions were equally likely, and calculated under models in which (i) only transitions occurred or (ii) transitions and transversions were equally likely. For *TSC1* truncating mutations, PD was predicted to occur if the PTC was within the first 350 codons [32], although it is recognized that some non-pathogenic mutations in this region may not affect the

structure [46] and some mutations outside this region may destabilise the protein. Thus 350/1164 (30.1%) of possible truncating mutations were predicted to exhibit PD. Predicted and observed numbers of mutations showing PD and/or NMD were evaluated from the population sample and from the LOVD database [45]. For *TSC2* truncating mutations, the GAP status was assumed to depend on whether the domain was fully intact (present) or truncated (absent).

Statistical Analysis

Data analysis was performed using SPSS version 21. Differences in IQ/DQ scores were analysed using the non-parametric Mann-Whitney U test. The Kolmogorov-Smirnov test was used to assess normality of IQ/DQ distributions above 19. 2x2 contingency tables were assessed using Fisher's exact test. Spearman's rank correlation test was used to correlate two independent variables. A threshold of $p < 0.05$ was considered statistically significant in all tests and no corrections were made for multiple testing.

RESULTS

Subject characteristics

Of the 113 participants (92 from the Cardiff study [24] and 21 from the Cambridge study [34]), 13 were excluded because intellectual or genetic data were incomplete. Of the remaining 100 subjects, 26 had *TSC1* and 74 *TSC2* mutations. There were 47 males and 53 females, 82 sporadic and 15 familial cases, with 3 of undetermined familial/sporadic origin. Details of mutations are given in Supplementary Table 1.

IQ/DQ in *TSC1* and *TSC2*

IQ/DQ scores (Supplementary Figure 2) were stratified and compared to the expected distribution in the general population (Table 1). Proportions with profound impairment were 28%, 11% and 34% (all, *TSC1* and *TSC2* respectively). Excluding subjects with profound ID, Kolmogorov-Smirnov tests showed normal distributions for remaining IQ scores (All $p = 0.574$; *TSC1* $p = 0.915$; *TSC2* $p = 0.551$).

Table 1 Stratification of IQ/DQ in the normal population and of all, *TSC1* and *TSC2* subjects in the study (n= 100). The lower panel in grey shows subdivisions of the category with IQ<70.

IQ	Description	Normal pop ⁿ (%)	All subjects (%) (n=100)	<i>TSC1</i> subjects (%) (n=26)	<i>TSC2</i> subjects (%) (n=74)
>130	Very superior	2.2	0.0	0.0	0.0
120-129	Superior	6.7	3.0	11.5	0.0
110-119	High average	16.1	6.0	7.7	5.4
90-109	Average	50.0	20.0	42.3	12.2
80-89	Low average	16.1	12.0	15.4	10.8
70-79	Borderline	6.7	10.0	3.8	12.2
<70	Extremely low	2.2	49.0	19.2	59.5
50-69	Mild ID		9%	4%	11%
35-49	Moderate ID		4%	4%	4%
20-34	Severe ID		8%	0%	11%
<20	Profound ID		28%	11%	34%

TSC2 was associated with significantly lower IQ/DQ than *TSC1* (Table 2) and the association held after exclusion of cases with profound ID (IQ<20), and after removal of familial cases (reducing possible ascertainment bias). Of those with IQ/DQ>20 (i.e. the ‘normal distribution phenotype’ [14]), those with *TSC1* mutations were psychometrically very similar to the general population ($\mu=97.5$, $\sigma=19.1$; compared to $\mu=100.0$, $\sigma=15.0$). Those with *TSC2* mutations showed significantly lower mean IQ

and very different psychometric distribution ($\mu=71.6$, $\sigma=27.1$). No significant difference between males and females was seen across any comparison (Supplementary Table 2), with no significant gender differences in the proportion of those with ID (males 51.1% vs females 47.1%) or profound ID (males 27.7% vs females 28.3%).

Table 2 Means (μ) and standard deviations (σ) of IQ/DQ with different TSC mutation status in the total sample, in those with IQ>20 and in the sporadic cases with IQ>20

TSC status	Number	μ	σ	Z value	P
Total sample					
All individuals	100	62.8	36.0		
<i>TSC1</i>	26	88.4	31.2		
<i>TSC2</i>	74	53.8	33.3		
<i>TSC1</i> vs <i>TSC2</i>				-4.252	<0.001*
IQ >20 only					
All individuals	72	79.9	27.5		
<i>TSC1</i>	23	97.5	19.1		
<i>TSC2</i>	49	71.6	27.1		
<i>TSC1</i> vs <i>TSC2</i>				-3.878	<0.001*
IQ>20 and sporadic only					
All individuals	54	78.8	26.9		
<i>TSC1</i>	11	94.8	22.9		
<i>TSC2</i>	43	74.7	26.6		
<i>TSC1</i> vs <i>TSC2</i>				-2.353	0.019*

Statistical comparisons were made with the Mann-Whitney U test. * denotes a statistically significant result.

Mutation Characteristics for TSC1 – NMD, Protein Degradation and TSC2

interaction

TSC1 frameshift mutations (Table 3 and Supplementary Fig 3A) were associated with significantly lower IQ/DQ than nonsense mutations (difference between mean IQ/DQ = 34.9; Z = -2.233; p=0.026). No *TSC1* missense mutations were found. All 25 *TSC1* truncating (i.e. nonsense or frameshift) mutations were predicted to result in NMD.

Strikingly, all truncating mutations in *TSC1* in the LOVD database [45] (2 May 2014) were also predicted to result in NMD, which was statistically highly significant compared to the null hypothesis that they were evenly distributed throughout the gene ($p < 0.0001$, Table 4). Predictions based on the more detailed calculation of sites where a single mutation could result in escape from NMD gave similar conclusions (not shown).

Table 3 Means (μ) and standard deviations (σ) of IQ/DQ with different mutational categories in *TSC1* and *TSC2*

Mutation category	Number	μ	σ	Z value	P
<i>TSC1</i>					
Nonsense	14	103.0	16.5		
Frameshift	11	68.1	36.0		
Nonsense vs frameshift				-2.233	0.026*
<i>TSC2</i>					
Nonsense	17	50.1	30.6		
Frameshift	13	56.7	34.2		
Nonsense vs frameshift				-0.558	0.577

Statistical comparisons were made with the Mann-Whitney U test. The individual with a *TSC1* splice site mutation is omitted. * denotes a statistically significant result.

Table 4 Fisher's exact tests of the distribution of mutations predicted to be associated with nonsense-mediated decay (NMD) and protein degradation (PD) status in the sample and of all mutations listed in the *TSC1* and *TSC2* Leiden Open Variation Database (accessed on 2 May 2014). Individuals with splice site mutations are omitted.

Type of analysis	Expected ratio	Observed Ratio	Two-tailed P value
This study (TSC1 = 25; TSC2 = 30)			
<i>TSC1</i> (no NMD:NMD)	4:21	0:25	0.111
<i>TSC2</i> (no NMD:NMD)	1:29	3:27	0.612
<i>TSC1</i> (no PD:PD)	17:8	15:10	0.551
LOVD Database (TSC1 = 588; TSC2 = 1020)			
<i>TSC1</i> (no NMD:NMD)	96:492	0:588	<0.0001*
<i>TSC2</i> (no NMD:NMD)	41:979	45:975	0.741
<i>TSC1</i> (no PD:PD)	411:177	417:171	0.749

* denotes a statistically significant result.

The numbers of *TSC1* mutations predicted to cause truncation in the region leading to protein degradation ('PD mutations') rather than the non-PD region ('no-PD mutations') in our dataset and the LOVD database were not significantly different from expectation. However, mutations predicted to result in PD were associated with higher IQ/DQ than non-PD mutations (Table 5). When only sporadic cases were considered, to reduce ascertainment bias, the difference was even more marked and significant (mean difference in IQ/DQ = 45.8; $Z = 2.275$; $p=0.027$). Hoogeveen-Westerveld et al. reported that substitutions in the first 224 amino acids could lead to protein instability [33]. Mutations in our dataset causing truncations within this more restricted region were also associated with a higher IQ/DQ than mutations leading to distal truncations, but sample numbers were too small for statistical analysis.

Individuals with mutations within the tuberin-interaction domain (TID) of TSC1 showed higher mean IQ/DQ ($n=6$, $\mu=101.3$, $\sigma= 13.6$) than others ($n=19$, $\mu=83.3$, $\sigma= 34.6$) and higher mean than those with mutations elsewhere in the first half of the protein ($n=7$, $\mu=90.4$, $\sigma= 16.9$). The differences were not significant, although this may reflect small sample sizes.

Table 5 Means (μ) and standard deviations (σ) of IQ/DQ with different expression categories in *TSC1* and *TSC2*. Individuals with splice site mutations are omitted.

Expression category	Number	μ	σ	Z value	P
<i>TSC1</i>					
No PD	15	79.7	37.6		
PD	10	99.5	14.6		
No PD vs PD				1.083	0.279
<i>TSC1</i> sporadic cases only					
No PD	10	65.5	36.8		
PD	4	111.3	14.2		
No PD vs PD				2.275	0.027*
<i>TSC2</i>					
No NMD	3	31.0	20.8		
NMD	27	55.4	32.2		
No NMD vs NMD				1.383	0.167
<i>TSC2</i> sporadic cases only					
No NMD	3	31.0	20.8		
NMD	26	55.0	32.7		
No NMD vs NMD				1.326	0.185

Statistical comparisons were made with the Mann-Whitney U test. * denotes a statistically significant result.

Correlation of predicted length of aberrant tail in *TSC1* and IQ/DQ

We tested for a correlation between the predicted lengths of aberrant tails generated by the frameshift or nonsense mutations and IQ/DQ (Figure 1A and Supplementary Table 3). Mutant proteins (frameshift or nonsense, 15 individuals) not predicted to be subject to PD showed significant negative correlation between tail length and IQ/DQ ($\rho = -0.706$; $p = 0.003$).

Mutation Characteristics for *TSC2* - NMD, *TSC1* interaction and GAP status

Unlike *TSC1*, there were several missense *TSC2* mutations (supplementary Fig 3B). Most were non-conservative, suggesting that many conservative substitutions would be non-pathogenic. The single individual in the sample with a conservative substitution (L1594M) showed an IQ/DQ of 19. Van Eeghen et al. found that individuals with missense mutations in exons 22-33 were less seriously affected than

individuals with mutations elsewhere [28]. In our study there was only one participant with a missense mutation in exons 22-33, with IQ/DQ=19. Of the 12 with missense mutations elsewhere, five also had IQ/DQ <20. Missense mutations were associated with lower IQ/DQ than truncating mutations, but the difference was not significant (n=13, $\mu=37.8$, $\sigma=29.8$ and n=30, $\mu=52.9$, $\sigma=31.8$ respectively, p=0.224). Among truncating mutations, there was no significant difference between frameshift and nonsense mutations (Table 3). In-frame deletions were associated with a higher IQ/DQ (n=4, $\mu=71.0$, $\sigma=16.8$) than truncations, but the sample was too small for statistical testing. There were 9 splice site mutations ($\mu=68.6$, $\sigma=29.8$), and 18 large deletions ($\mu=55.8$, $\sigma=39.5$).

For *TSC2* PTC mutations the numbers of NMD and no-NMD mutations were not significantly different from expected, either in our dataset or in the LOVD database (Table 4). No-NMD mutations were associated with a lower IQ/DQ than NMD mutations (n = 3; $\mu=31$; $\sigma=20.8$; n = 27; $\mu=55.4$; $\sigma=32.2$), although results were not statistically significant (Table 5).

There was one individual with a missense mutation in the TSC1-interaction domain of TSC2, defined as extending to the end of exon 11 [47], with IQ/DQ =21. When the region considered was extended (HID, dotted line in Supplementary Figure 1) to the N-terminal half of TSC2, including exon 22 (and including missense mutations potentially affecting the interaction domain as defined by van Eeghen et al.) there was no significant difference compared to individuals with missense mutations elsewhere. Van Eeghen et al. [28] by contrast found lower IQ/DQ for individuals with mutations in the interaction domain. Three of the 30 *TSC2* truncating mutations were predicted

to result in retention of the GAP domain, whereas 27 were predicted to result in its loss. IQ/DQ was not significantly different between the two groups ($n = 3$; $\mu=31.0$, $\sigma=20.8$, and $n = 27$; $\mu=55.4$, $\sigma= 32.2$, $Z =-1.383$; $p=0.167$).

Correlation of length of aberrant tail in predicted mutant TSC2 proteins and IQ/DQ

After discounting the no-NMD mutations which may result in lower IQ/DQ for other reasons (see below) there was a positive, but not statistically significant, correlation ($\rho = 0.308$; $p = 0.119$) between length of predicted aberrant tail and IQ/DQ (Figure 1B and Supplementary Table 3). However, the nonsense mutations within this group showed a wide spread of IQ/DQ values.

DISCUSSION

IQ/DQ distribution

The IQ/DQ of our patients (Supplementary Figure 2) was comparable to those in another epidemiological study in the UK from 2003 [13]. Both studies showed bimodal IQ distribution, with similar proportions of ID and profound impairment (49.0% vs 44.4% and 28.0% vs 30.5% respectively), and suggesting our sample was representative of the TSC population at the time patient phenotypes were assessed.

Even excluding individuals with profound ID, the *TSC1* and *TSC2* ‘normal distribution’ phenotypes [14] remained psychometrically and statistically significantly different, with *TSC1* individuals psychometrically very similar to the general population, and *TSC2* individuals associated with significantly lower IQ (Table 2).

Results confirmed that the higher proportion of ID in *TSC2* was not simply attributable to more individuals with profound ID.

The lack of significant difference between males and females in mean IQ (Supplementary Table 2) or the proportion of ID (males 51.1% vs females 47.1%) or profound ID (males 27.7% vs females 28.3%) refutes earlier, non-standardised reports suggesting more ID in males with TSC [25,26], and replicates the observations made by van Eeghen et al. [28].

Mutation characteristics for *TSC1*

The lack of *TSC1* missense mutations in our sample of 26 individuals is consistent with previous reports that such mutations are rare [48]. It was remarkable that all 25 *TSC1* truncating mutations in our sample and all 588 *TSC1* truncating mutations in the LOVD database were predicted to show NMD. There are two possible interpretations. Truncation of the *TSC1* protein within the C-terminal 190 residues may have such a slight effect that mutations do not cause TSC, and are not seen in a sample of individuals chosen because they have the disorder. Alternatively, *TSC1* proteins truncated within the C-terminal 190 residues may be lethal if present at wild-type levels, because of a dominant effect. We favour the latter explanation, for two reasons. The TSC LOVD includes 157 *TSC1* variants classified as having ‘no known pathogenicity’ [45] (Sept 2014) and 25 of these are located in the no-NMD region, consistent with expectation if equally distributed. However, none is a truncating mutation, although 56% of mutations overall result in truncation. Thus, if truncating mutations in the no-NMD region were indeed non-pathogenic, we would expect to see examples of them in LOVD. In addition, most truncating mutations in the no-NMD

region would be expected to affect the ERM domain (believed to have an important role in mediating cell adhesion [49]) and a region involved in interaction with TSC2 [32]. Although the suggestion that truncating mutants of TSC1 not subject to NMD are embryonic-lethal because of dominant effects contrasts with the general view that *TSC1* mutations have a mild phenotype, it is consistent with the demonstration [31] that TSC1 protein with substitutions at positions 50 or 117 can have a dominant negative effect, and the suggestion by these authors that protein with mutations elsewhere might also have a dominant negative effect [31].

A deleterious effect of mutant TSC1 would also explain our observation that individuals with mutations predicted to result in PD showed higher IQ/DQ than those with mutations outside the PD region. Because NMD does not lead to complete loss of transcripts [30], production of some truncated protein would be expected. We propose that this is not immediately degraded, but that lower levels persist for protein truncated within the PD region, resulting in less marked effects on IQ/DQ. In fact, these individuals showed a mean and standard deviation in IQ/DQ scores essentially identical to the general population. Furthermore, although no patients in this study had deletions of the entire *TSC1* gene, we have observed 13 patients with such deletions in a larger diagnostic series and none was reported to have ID. Given the location of the TSC2-interaction domain towards the N terminus of TSC1, it is not surprising that individuals with mutations in the interaction domain also showed higher IQ/DQ than others (although this contrasts with the observation of van Eeghen et al).

Length of aberrant tail on TSC1

For individuals with frameshift *TSC1* mutations outside the PD region, effects on IQ/DQ were greater the longer the aberrant C-terminal tail on the predicted mutant protein, and frameshift mutations were associated with lower IQ than nonsense mutations. This could be explained if the aberrant tail increases the deleterious biochemical effects of the protein (or reduces its rate of degradation), again consistent with the notion that mutant protein is detrimental.

Mutation characteristics for *TSC2*

Results for *TSC2* were quite different from those for *TSC1*, with no significant clustering of mutations in the NMD region in our population or in the *TSC2* LOVD (Table 4). Although truncating mutations predicted to give rise to NMD were associated with a higher IQ/DQ than no-NMD mutations, this was not statistically significant (Table 5). These observations suggest that mutant *TSC2* proteins truncated near the C-terminus are only slightly deleterious in terms of IQ/DQ. The calmodulin-binding domain, the oestrogen receptor alpha binding site and a target site for the ribosomal protein S6 kinase are all within the no-NMD region [47], although the functional relevance of these domains has not yet been fully established.

Van Eeghen et al. [28] reported that individuals with missense mutations in *TSC2* had a higher IQ/DQ than those with premature truncation mutations, whereas we observed a tendency for the reverse. This may reflect the smaller sample size in our study.

Similarly, van Eeghen et al. observed a lower IQ/DQ for individuals with a mutation in the first half of the protein compared to those with a mutation in the second half, whereas we observed no difference (and a slightly higher IQ/DQ for individuals with

a mutation in the TSC1-interaction domain). Like van Eeghen et al., we observed little difference between mutations affecting the GAP domain and those that did not.

Length of aberrant tail on TSC2

In individuals with frameshift mutations, there was a positive but not statistically significant correlation between the length of the aberrant tail and IQ (Figure 1B). This might be explained if longer aberrant tails led to more effective clearing of mutant protein (for which degradation systems have been described [50]) or reduced the deleterious effects of mutant protein.

Model for the effects of mutations

We propose the following model for the different effects of mutant TSC1 and TSC2 proteins. Mutant TSC1 protein is highly deleterious, although NMD leads to partial reduction of the expression of truncated protein. Mutant protein is degraded slowly, and has an adverse effect on cognitive functions, including IQ/DQ. Proteins with longer aberrant C-terminal tails are more deleterious in this respect. TSC1 protein altered or truncated in the N-terminal 200-300 residues is degraded more rapidly, causing a less severe phenotype. The model is consistent with observations reported by van Eeghen et al. For TSC1, most individuals in their study with truncating mutations up to and including exon 5 showed an IQ/DQ of 70 or more, consistent with our proposal that more rapid clearance of mutant protein in those individuals, leaves them relatively mildly affected [32]. The destabilizing effect of mutations close to the N-terminus of TSC1 is consistent with the existence of a conserved, tightly-folded domain in this region [46].

We suggest that levels of truncated TSC2 protein are also reduced through NMD, with mutations in the no-NMD region leading to more severe, but not lethal, consequences. Mutant protein is degraded more effectively than TSC1, and proteins destabilized by having a long aberrant C-terminal tail are less deleterious.

However, it will be important to test biochemically the predictions of the model for levels and function of mutant TSC1 and TSC2 proteins. If the predictions are correct, it will also be important to understand why longer aberrant C-terminal tails have different effects on TSC1 and TSC2.

Our analysis solely considered genotypes, and other factors will also affect outcomes, illustrated by some individuals with the same mutation having very different IQ/DQ (e.g. TSC2 mutations R751X: IQ/DQ 19, 85, 94; Q1192X: IQ/DQ 19, 79 and R1459X: IQ/DQ 19, 19, 68). These factors include somatic mosaicism [51], modifier genes (perhaps affecting NMD efficiency) [52] and environmental effects such as seizures [28] (although tubers and seizures combined accounted for only 47% of the variance in IQ in the study of O'Callaghan et al [16]). Nevertheless, it is clear that there are differing effects of distinct mutations in *TSC1* and *TSC2* (and specific mutations associated with a mild phenotype have been recognized for some time, e.g., [53]). , Larger-scale exploration of the relationship between the molecular consequences of TSC mutations, including direct measurements of levels and function of mutant proteins *in vivo*, and intellectual phenotype is required to confirm our findings and may lead to discoveries of direct relevance in the clinical setting.

ACKNOWLEDGEMENTS

We thank the Tuberous Sclerosis Association, the Wales Gene Park, the National Research Foundation of South Africa, and the Struengmann Fund for financial support. We thank Prof Chris Smith for helpful comments on the manuscript.

Contributors All authors contributed to design, data collection, analysis and interpretation of results. All authors approved the final manuscript submitted.

Competing interests None declared.

Ethics approval Cardiff and Cambridge MRECs.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No new data were generated in this study—all data used were contributed to the publically accessible LOVD database (<http://chromium.liacs.nl/LOVD2/TSC/home.php>). All data referred to in the manuscript are also listed in online supplementary table S1.

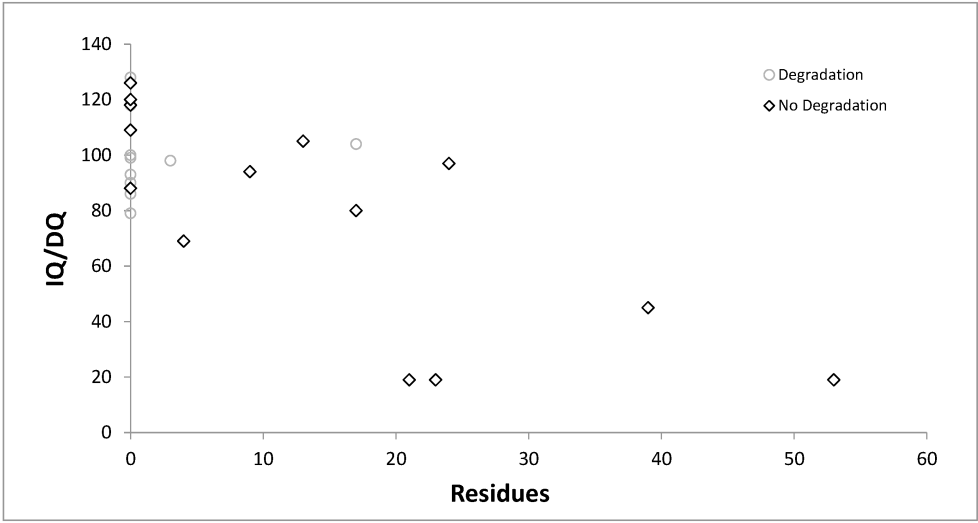
REFERENCES

- 1 Curatolo P, Moavero R, de Vries PJ. Neurological and neuropsychiatric aspects of tuberous sclerosis complex. *Lancet Neurol* 2015;14:733-745.
- 2 Northrup H, Krueger DA, International Tuberous Sclerosis Complex Consensus Group. Tuberous sclerosis complex diagnostic criteria update: recommendations of the 2012 international tuberous sclerosis complex consensus conference. *Pediatr Neurol* 2013;49:243-54.
- 3 van Slegtenhorst M, de Hoogt R, Hermans C et al. Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34. *Science* 1997; 277:805-808.
- 4 European Chromosome 16 Tuberous Sclerosis Consortium. Identification and characterization of the tuberous sclerosis gene on chromosome 16. *Cell* 1993;75:1305-15.
- 5 Chong-Kopera H, Inoki K, Li Y, Zhu T, Garcia-Gonzalo FR, Rosa JL, Guan KL. TSC1 stabilizes TSC2 by inhibiting the interaction between TSC2 and the HERC1 ubiquitin ligase. *J Biol Chem* 2006;281:8313-8316.
- 6 Cai SL, Tee AR, Short JD, Bergeron JM, Kim J, Shen J, Guo R, Johnson CL, Kiguchi K, Walker CL. Activity of TSC2 is inhibited by AKT-mediated phosphorylation and membrane partitioning. *J Cell Biol* 2006;173:279-289.
- 7 Plank TL, Yeung RS, Henske EP. (1998). Hamartin, the product of the tuberous sclerosis 1 (TSC1) gene, interacts with tuberlin and appears to be localized to cytoplasmic vesicles. *Cancer Res* 1998;58:4766-4770.
- 8 Hoogeveen-Westerveld M, van Unen L, van den Ouweland A, Hally D, Hoogeveen A, Nellist M. The TSC1-TSC2 complex consists of multiple TSC1 and TSC2 subunits. *BMC Biochem* 2012;13:18.
- 9 Dibble CC, Elis W, Menon S, Qin W, Klekota J, Asara JM, Finan PM, Kwiatkowski DJ, Murphy LO, Manning BD TBC1D7 is a third subunit of the TSC1-TSC2 complex upstream of mTORC1. *Mol Cell* 2012;47:535-546.
- 10 Tee AR, Manning BD, Roux PP, Cantley LC, Blenis J. Tuberous sclerosis complex gene products, Tuberlin and Hamartin, control mTOR signaling by acting as a GTPase-activating protein complex toward Rheb. *Curr Biol* 2003;13:1259-1268.
- 11 Inoki K, Li Y, Xu T, Guan KL. Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling. *Genes Dev* 2003;17:1829-1834.
- 12 World Health Organization. 1993. The ICD-10 Classification of Mental and Behavioural Disorders. Diagnostic Criteria for Research. World Health, Organization, Geneva.
- 13 Joinson C, O'Callaghan FJ, Osborne JP, Martyn C, Harris T, Bolton PF Learning disability and epilepsy in an epidemiological sample of individuals with tuberous sclerosis complex. *Psychol Med* 2003;33:335-344.
- 14 de Vries PJ, Prather PA. The tuberous sclerosis complex. *N Engl J Med* 2007;356:92.
- 15 Goodman M, Lamm SH, Engel A, Shepherd CW, Houser OW, Gomez MR. Cortical tuber count: a biomarker indicating neurologic severity of tuberous sclerosis complex. *J Child Neurol* 1997;12:85-90.
- 16 O'Callaghan FJ, Harris T, Joinson C, Bolton P, Noakes M, Presdee D, Renowden S, Shiell A, Martyn CN, Osborne JP. The relation of infantile spasms, tubers, and intelligence in tuberous sclerosis complex. *Arch Dis Child* 2004;89:530-3.
- 17 de Vries PJ, Howe CJ. The tuberous sclerosis complex proteins - a GRIPP on cognition and neurodevelopment. *Trends Mol Med* 2007;13:319-326.

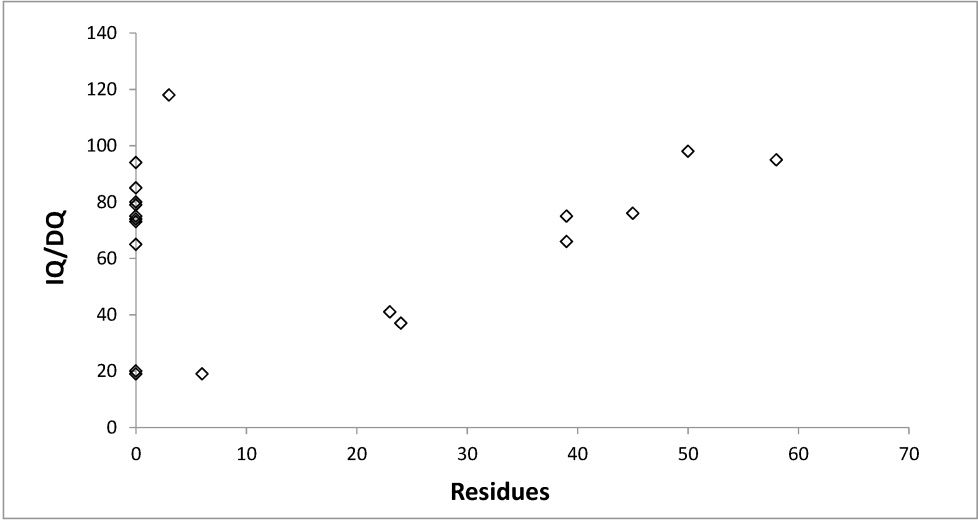
- 18 Goorden SM, van Woerden GM, van der Weerd L, Cheadle JP, Elgersma Y. Cognitive deficits in Tsc1^{+/-} mice in the absence of cerebral lesions and seizures. *Ann Neurol* 2007;62:648-655.
- 19 Ehninger D, Han S, Shilyansky C, Zhou Y, Li W, Kwiatkowski DJ, Ramesh V, Silva AJ. Reversal of learning deficits in a Tsc2^{+/-} mouse model of tuberous sclerosis. *Nat Med* 2008;14:843-848.
- 20 Tsai PT, Hull C, Chu Y, Greene-Colozzi E, Sadowski AR, Leech JM, Steinberg J, Crawley JN, Regehr WG, Sahin M. Autistic-like behaviour and cerebellar dysfunction in Purkinje cell Tsc1 mutant mice. *Nature* 2012;488:647-651.
- 21 Davies DM, de Vries PJ, Johnson SR, McCartney DL, Cox JA, Serra AL, Watson PC, Howe CJ, Doyle T, Pointon K, Cross JJ, Tattersfield AE, Kingswood JC, Sampson JR. Sirolimus therapy for angiomyolipoma in tuberous sclerosis and sporadic lymphangiomyomatosis: a phase 2 trial. *Clin Cancer Res* 2011;17:4071-4081.
- 22 Jones AC, Daniells CE, Snell RG, Tachataki M, Idziaszczyk SA, Krawczak M, Sampson JR, Cheadle JP. Molecular genetic and phenotypic analysis reveals differences between TSC1 and TSC2 associated familial and sporadic tuberous sclerosis. *Hum Mol Genet* 1997;6:2155-2161
- 23 Dabora SL, Jozwiak S, Franz DN, Roberts PS, Nieto A, Chung J, Choy YS, Reeve MP, Thiele E, Egelhoff JC, Kasprzyk-Obara J, Domanska-Pakiela D, Kwiatkowski DJ. Mutational analysis in a cohort of 224 tuberous sclerosis patients indicates increased severity of TSC2, compared with TSC1, disease in multiple organs. *Am J Hum Genet* 2001;68:64-80.
- 24 Lewis JC, Thomas HV, Murphy KC. Genotype and psychological phenotype in tuberous sclerosis. *J Med Genet* 2004;41:203-207.
- 25 Sancak O, Nellist M, Goedbloed M, Elfferich P, Wouters C, Maat-Kievit A, Zonnenberg B, Verhoef S, Halley D, van den Ouweland A. Mutational analysis of the TSC1 and TSC2 genes in a diagnostic setting: genotype - phenotype correlations and comparison of diagnostic DNA techniques in Tuberous Sclerosis Complex. *Eur J Hum Genet* 2005;13:731-741.
- 26 Au KS, Williams AT, Roach ES, Batchelor L, Sparagana SP, Delgado MR, Wheless JW, Baumgartner JE, Roa BB, Wilson CM, Smith-Knuppel TK, Cheung MY, Whittemore VH, King TM, Northrup H. Genotype/phenotype correlation in 325 individuals referred for a diagnosis of tuberous sclerosis complex in the United States. *Genet Med* 2007;9:88-100.
- 27 Jansen FE, Braams O, Vincken KL, Algra A, Anbeek P, Jennekens-Schinkel A, Halley D, Zonnenberg BA, van den Ouweland A, van Huffelen AC, van Nieuwenhuizen O, Nellist M. Overlapping neurologic and cognitive phenotypes in patients with TSC1 or TSC2 mutations. *Neurology* 2008;70:908-915.
- 28 van Eeghen AM, Black ME, Pulsifer MB, Kwiatkowski DJ, Thiele EA. Genotype and cognitive phenotype of patients with tuberous sclerosis complex. *Eur J Hum Genet* 2012;20:510-515.
- 29 Nagy E, Maquat LE. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem Sci* 1998;23:198-199.
- 30 Jeganathan D, Fox MF, Young JM, Yates JRW, Osborne JP, Povey S. Nonsense-mediated RNA decay in the TSC1 gene suggests a useful tool pre- and post-positional cloning. *Hum Genet* 2002;111:555-565.
- 31 Mozaffari M, Hoogeveen-Westerveld M, Kwiatkowski D, Sampson J, Ekong R, Povey S, den Dunnen JT, van den Ouweland A, Halley D, Nellist M.

- Identification of a region required for TSC1 stability by functional analysis of TSC1 missense mutations found in individuals with tuberous sclerosis complex. *BMC Med Genet* 2009;10:88.
- 32 Hoogeveen-Westerveld M, Exalto C, Maat-Kievit A, van den Ouweland A, Halley D, Nellist M. Analysis of TSC1 truncations defines regions involved in TSC1 stability, aggregation and interaction. *Biochim Biophys Acta* 2010;1802:774-781.
 - 33 Hoogeveen-Westerveld M, Wentink M, van den Heuvel D, Mozaffari M, Ekong R, Povey S, den Dunnen JT, Metcalfe K, Vallee S, Krueger S, Bergoffen J, Shashi V, Elmslie F, Kwiatkowski D, Sampson J, Vidales C, Dzarir J, Garcia-Planells J, Dies K, Maat-Kievit A, van den Ouweland A, Halley D, Nellist M. Functional assessment of variants in the TSC1 and TSC2 genes identified in individuals with Tuberous Sclerosis Complex. *Hum Mutat* 2011;32:424-435.
 - 34 Tierney KM, McCartney DL, Serfontein JR, de Vries PJ. Neuropsychological attention skills and related behaviours in adults with tuberous sclerosis complex. *Behav Genet* 2011;41:437-444.
 - 35 Roach ES, Gomez MR, Northrup H. Tuberous sclerosis complex consensus conference: revised clinical diagnostic criteria. *J Child Neurol* 1998;13:624-628.
 - 36 Wechsler D. 1981. Wechsler Adult Intelligence Scale - Revised. Psychological Corporation, London.
 - 37 Wechsler D. 1992. Wechsler Intelligence Scale for Children - III. Psychological Corporation, London.
 - 38 Raven J, Raven JC, Court JH. 1998. Coloured Progressive Matrices. Oxford Psychologists Press, Oxford.
 - 39 Sparrow SS, Balla DA, Cicchetti DV. 1984. Vineland Adaptive Behavior Scales. Circle Press, Minnesota.
 - 40 Wechsler D. 1999. Wechsler Abbreviated Scale of Intelligence. Psychological Corporation, London.
 - 41 World Health Organization. 1994. The ICD-10 Classification of Mental and Behavioural Disorders. Clinical Descriptions and Diagnostic Guidelines. World Health Organization, Geneva.
 - 42 Jones AC, Shyamsundar MM, Thomas MW, Maynard J, Idziaszczyk S, Tomkins S, Sampson JR, Cheadle JP. Comprehensive mutation analysis of TSC1 and TSC2-and phenotypic correlations in 150 families with tuberous sclerosis. *Am J Hum Genet* 1999;64:1305-1315.
 - 43 Jones AC, Sampson JR, Hoogendoorn B, Cohen D, Cheadle JP. Application and evaluation of denaturing HPLC for molecular genetic analysis in tuberous sclerosis. *Hum Genet* 2000;106:663-668.
 - 44 Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and Clustal X version 2.0. *Bioinformatics* 2007; 23:2947-2948.
 - 45 <http://chromium.liacs.nl/LOVD2/TSC/home.php>
 - 46 Sun W, Zhu YJ, Wang Z, Zhong Q, Gao F, Lou J, Gong W, Xu W. Crystal structure of the yeast TSC1 core domain and implications for tuberous sclerosis pathological mutations. *Nature Comms* 2013;4:2135
 - 47 Serfontein J, Nisbet RER, Howe CJ, de Vries PJ. Conservation of structural and functional elements of TSC1 and TSC2: a bioinformatic comparison across animal models. *Behav Genet* 2011;41:349-356.

- 48 Kwiatkowski DJ. Genetics of Tuberous Sclerosis Complex. In Kwiatkowski DJ, Whittemore VH, Thiele EA eds. *Tuberous Sclerosis Complex: Genes, Clinical Features and Therapeutics* pp29-60. Wiley, Wienheim 2010.
- 49 Lamb RF, Roy C, Diefenbach TJ, Vinters HV, Johnson MW, Jay DG, Hall A. The TSC1 tumour suppressor hamartin regulates cell adhesion through ERM proteins and the GTPase Rho. *Nat Cell Biol* 2000;2:281-287.
- 50 Han S, Witt RM, Santos TM, Polizzano C, Sabatini BL, Ramesh V. Pam (Protein associated with Myc) functions as an E3 ubiquitin ligase and regulates TSC/mTOR signaling. *Cell Signal* 2008;20:1084-1091.
- 51 Emmerson P, Maynard J, Jones S, Butler R, Sampson JR, Cheadle JP. Characterizing mutations in samples with low-level mosaicism by collection and analysis of DHPLC fractionated heteroduplexes. *Hum Mutat* 2003;21:112-115.
- 52 Seoighe C, Gehring C. Heritability in the efficiency of nonsense-mediated mRNA decay in humans. *PLoS One* 2010;5:e11657
- 53 Jansen AC, Sancak O, D'Agostino MD, Badhwar A, Roberts P, Gobbi G, Wilkinson R, Melanson D, Tampieri D, Koenekoop R, Gans M, Maat-Kievit A, Goedbloed M, van den Ouweland AMW, Nellist M, Pandolfo M, McQueen M, Sims K, Thiele EA, Dubeau FA, Andermann F, Kwiatkowski DJ, Halley DJJ, Andermann E. Unusually mild tuberous sclerosis phenotype is associated with TSC2 R905Q mutation. *Ann Neurol* 2006; 60:528-539.



1A



IB

Figure Legend

Figure 1 Correlations between IQ/DQ and the length of predicted aberrant C-terminal tail of nonsense and frameshift mutations. A) *TSC1* mutations with and without predicted protein degradation, B) *TSC2* mutations not predicted to be subject to NMD. For both panels the y-axis indicates IQ/DQ and the x-axis length of the predicted aberrant C-terminal tail arising from frameshift mutations (0 for nonsense mutations). For *TSC1* cases where mutant protein was not expected to be degraded, there was a strong negative correlation between the length of the C-terminal aberrant tail and IQ/DQ ($\rho = -0.706$; $p = 0.003$).

Supplementary Table 1. List of individual patient mutations, familial or sporadic status of mutation and best estimate IQ/DQ by *TSC1* and *TSC2*.

Mutation	Familial/Sporadic	IQ	Mutation	Familial/Sporadic	IQ
<i>TSC1</i>			<i>TSC2 continued</i>		
R37SP	Familial	108	S946SP	Sporadic	65
S91X	Unknown	86	T993SP	Sporadic	99
S91X	Familial	79	R1032X	Sporadic	19
W103X	Sporadic	118	G1133FSX	Sporadic	95
S108FSX	Unknown	104	Q1192X	Sporadic	19
R245X	Sporadic	128	Q1192X	Sporadic	79
Q328X	Familial	93	Q1419X	Sporadic	73
Q328X	Familial	90	S1433X	Sporadic	65
Q328FSX	Familial	98	R1459X	Sporadic	73
S334X	Sporadic	99	R1459X	Sporadic	19
T339X	Sporadic	100	R1459X	Sporadic	19
H371FSX	Sporadic	69	P1497R	Sporadic	19
T417FSX	Familial	97	F1499FSX	Sporadic	37
Q527FSX	Familial	80	V1531FSX	Sporadic	76
N634FSX	Sporadic	19	L1594M	Sporadic	19
Q701FSX	Sporadic	19	N1643K	Sporadic	117
R715FSX	Sporadic	94	Q1665X	Sporadic	20
W750X	Sporadic	109	P1675K	Unknown	38
R786X	Sporadic	88	N1681K	Sporadic	19
R786X	Sporadic	88	M1691SP	Sporadic	72
E788X	Familial	126	M1721SP	Sporadic	54
R804FSX	Sporadic	19	P1737FS	Sporadic	55
E890FSX	Sporadic	105	R1743P	Sporadic	19
R892FSX	Sporadic	45	H1746del	Sporadic	65
Q898X	Familial	118	H1746del	Sporadic	50
Q898X	Familial	120	H1746del	Sporadic	84
<i>TSC2</i>			T1780FS	Sporadic	19
L259SP	Sporadic	34	F1803FS	Sporadic	19
G294E	Familial	21	Large Deletion	Sporadic	24
Y336X	Sporadic	74	Large Deletion	Familial	81
I365DEL	Sporadic	85	Large Deletion	Sporadic	19
P419SP	Sporadic	87	Large Deletion	Sporadic	19
R458X	Sporadic	19	Large Deletion	Sporadic	19
E482SP	Sporadic	75	Large Deletion	Sporadic	19
I496FSX	Familial	66	Large Deletion	Sporadic	25
R505X	Sporadic	80	Large Deletion	Sporadic	105
Y598X	Sporadic	75	Large Deletion	Familial	41
R611Q	Sporadic	80	Large Deletion	Sporadic	82
R611Q	Sporadic	33	Large Deletion	Sporadic	19
R611W	Familial	50	Large Deletion	Sporadic	19
R611W	Sporadic	19	Large Deletion	Sporadic	105
R611W	Sporadic	19	Large Deletion	Sporadic	100
A614D	Sporadic	38	Large Deletion	Sporadic	114
T659FSX	Sporadic	75	Large Deletion	Sporadic	100
A675FSX	Sporadic	41	Large Deletion	Sporadic	94
V692FSX	Sporadic	19			
R751X	Sporadic	85			
R751X	Sporadic	94			
R751X	Sporadic	19			
A765FSX	Sporadic	19			
Q785SP	Sporadic	19			
Q785SP	Sporadic	112			
Y790X	Sporadic	19			
L844FSX	Sporadic	98			
S879FSX	Sporadic	118			

Supplementary Table 2. Means (μ) and standard deviations (σ) of IQ/DQ with different gender in different TSC status.

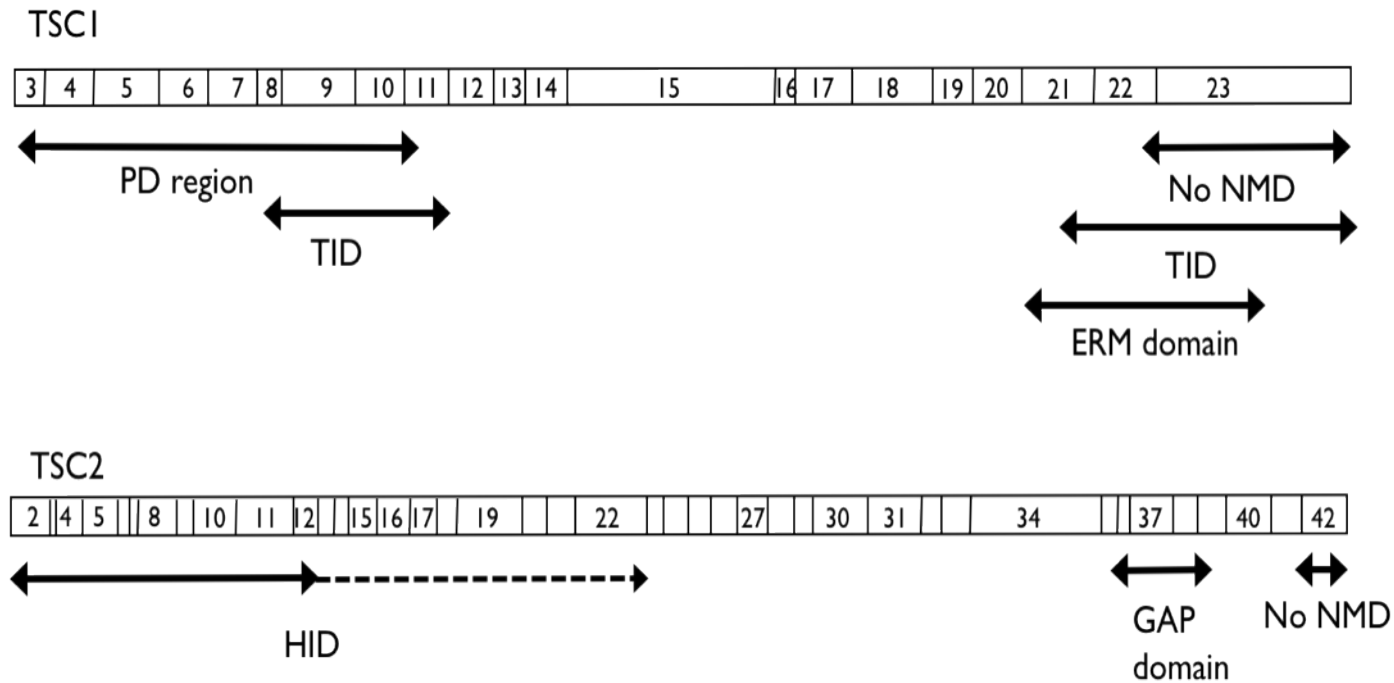
Gender status	Number	μ	σ	Z value	P
All subjects					
Males	47	62.0	36.5		
Females	53	63.5	35.9		
Males vs Females				-0.164	0.870
TSC1					
Males	10	89.3	33.9		
Females	16	87.9	30.6		
Males vs Females				0.343	0.732
TSC2					
Males	37	54.6	33.9		
Females	37	53.0	33.2		
Males vs Females				0.182	0.856

Statistical comparisons were made with the Mann-Whitney U test. * denotes a statistically significant result.

Supplementary Table 3. Spearman's rank correlations for *TSC1* and *TSC2* aberrant tail length and IQ/DQ. Only individuals with nonsense and frameshift mutations are included. Individuals with missense (*TSC2*) or splice site (*TSC1* and *TSC2*) mutations are excluded.

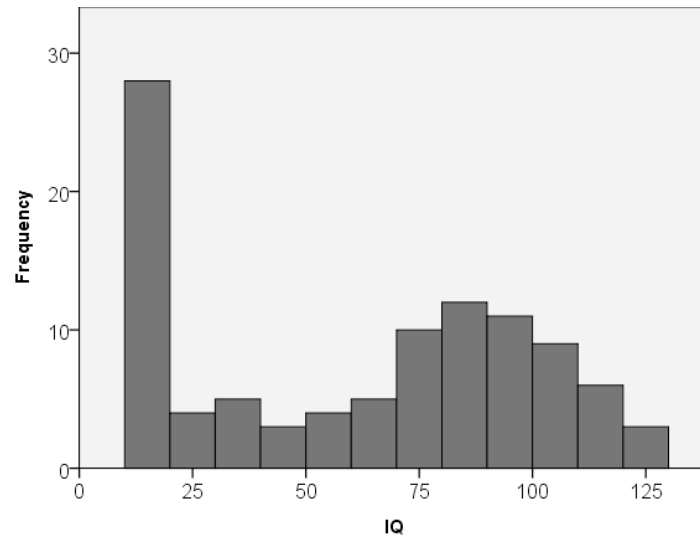
	N	ρ	P
TSC1 all subjects	25	-0.539	0.005*
TSC1 no degradation	15	-0.706	0.003*
TSC1 degradation	10	0.199	0.582
TSC1 frameshift	11	-0.552	0.078
TSC2 all subjects	30	0.092	0.628
TSC2 no NMD	3	0.866	0.333
TSC2 NMD	27	0.308	0.119
TSC2 frameshift	13	-0.031	0.921
TSC2 frameshift NMD	10	0.422	0.224

* denotes a statistically significant result

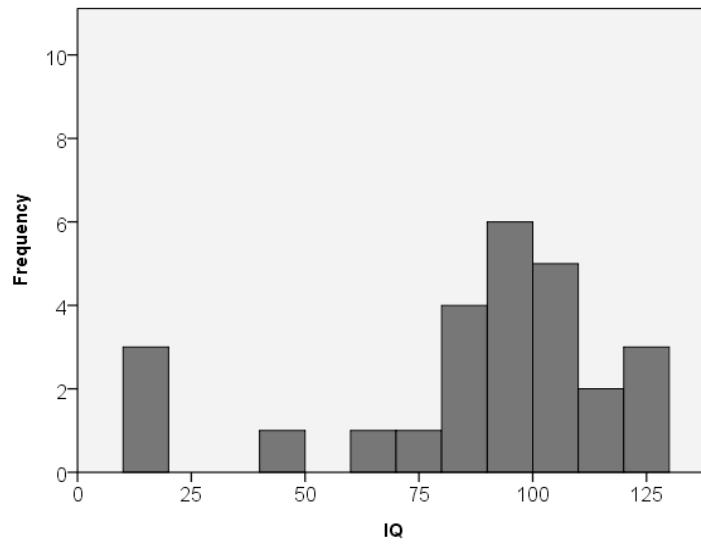


Supplementary Figure 1. Schematic representation of the regions of the TSC1 and TSC2 proteins considered here. Exon sizes are represented approximately to scale. Exon numbers correspond to those in the reference sequences used in the analysis. Numbers of some exons for *TSC2* have been omitted for clarity. NMD = nonsense mediated decay; PD = protein degradation; TID = tuberlin interaction domain; HID = hamartin interaction domain (see text for discussion of extent of this); GAP = GTPase activating protein; ERM = (Ezrin–Radixin–Moesin)-binding

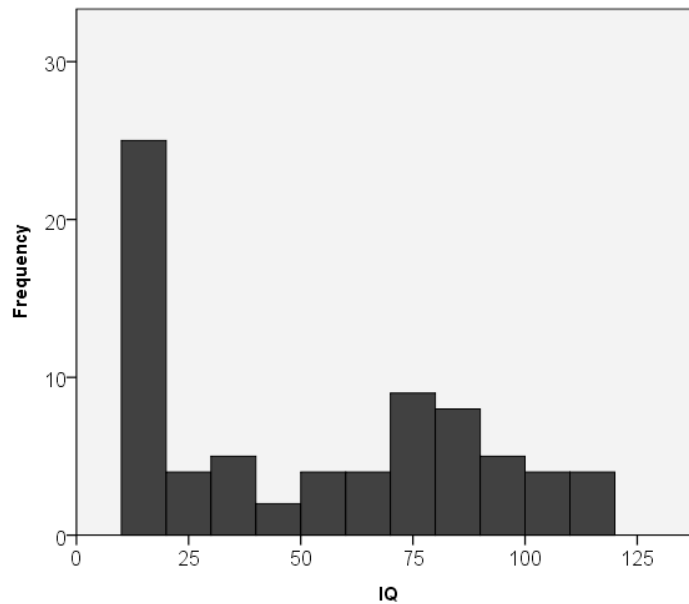
A (all subjects)



B (TSC1)

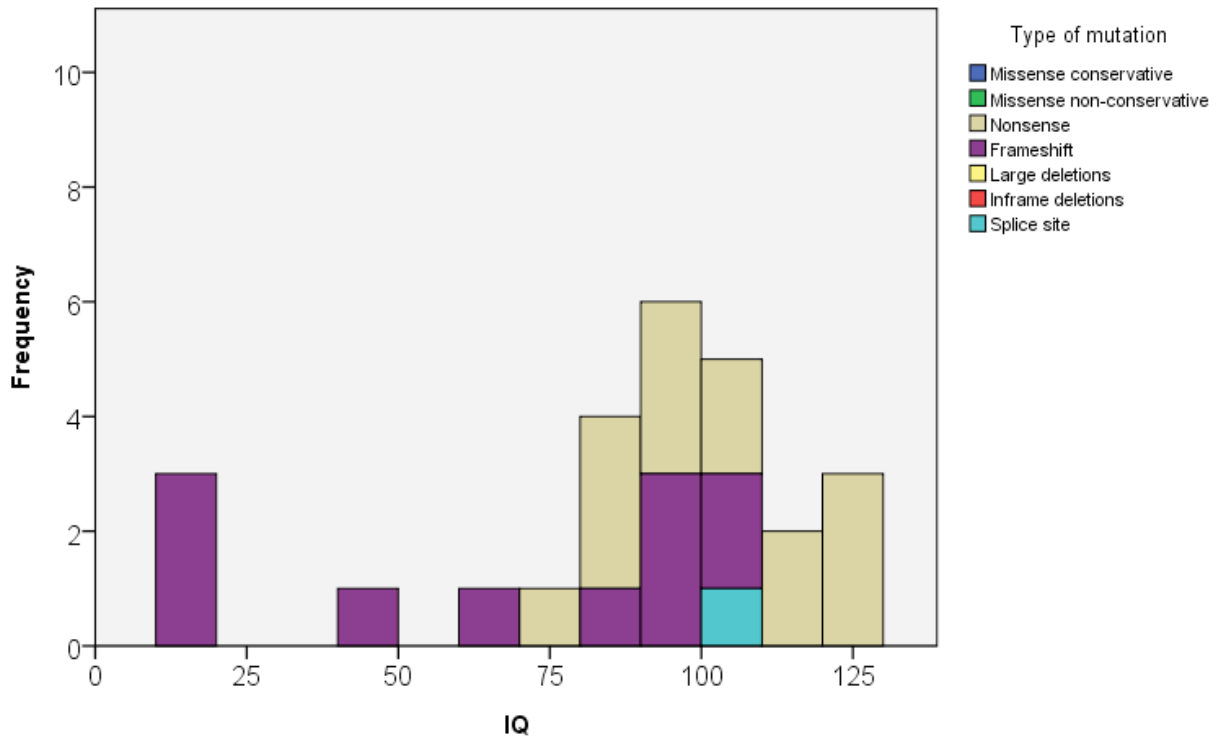


C (TSC2)

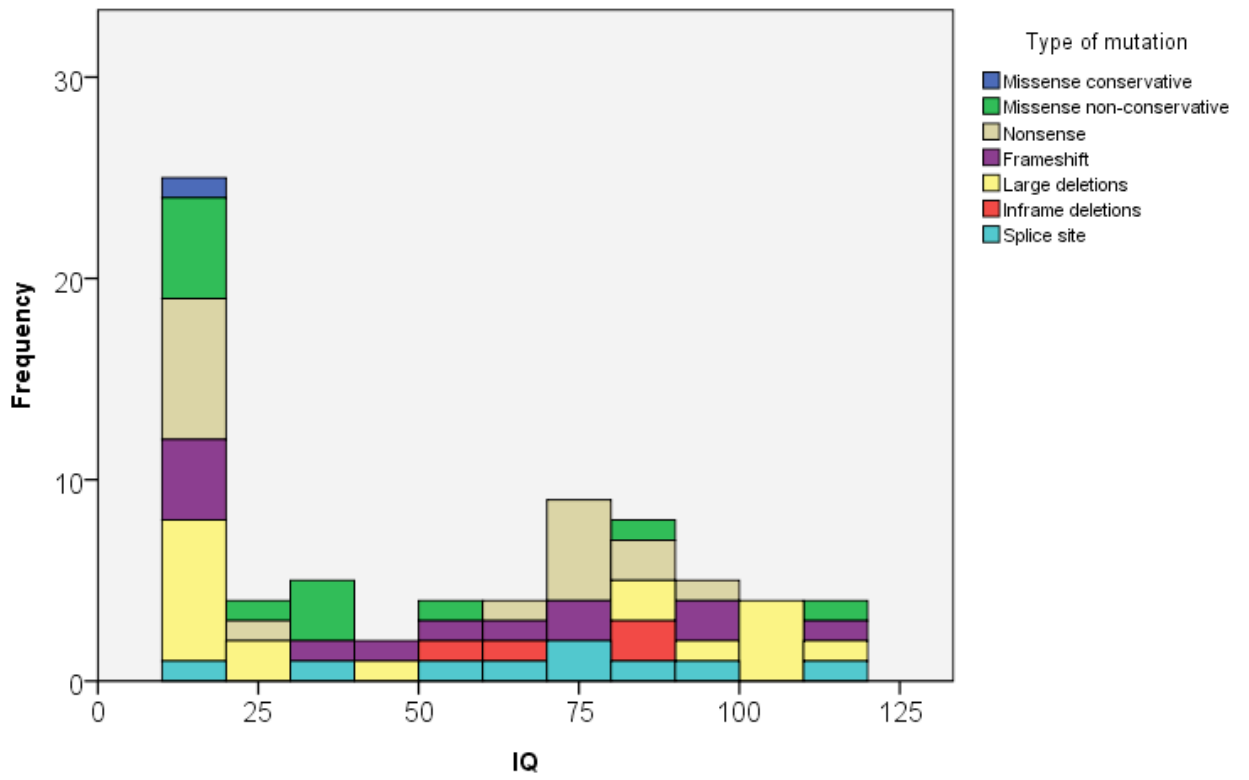


Supplementary Figure 2. Histograms showing the bimodal distribution of IQ/DQ in the study. A) The total sample (n = 100), B) *TSC1* group (n = 26), C) *TSC2* group (n = 74).

A
(TSC1)



B
(TSC2)



Supplementary Figure 3. Stacked histogram of IQ/DQ in relation to type of mutation. A) *TSC1* mutations, B) *TSC2* mutations.