

RESEARCH

Open Access



Functional characterization of rare FOXP2 variants in neurodevelopmental disorder

Sara B. Estruch^{1†}, Sarah A. Graham^{1†}, Swathi M. Chinnappa¹, Pelagia Deriziotis¹ and Simon E. Fisher^{1,2*}

Abstract

Background: Heterozygous disruption of *FOXP2* causes a rare form of speech and language impairment. Screens of the *FOXP2* sequence in individuals with speech/language-related disorders have identified several rare protein-altering variants, but their phenotypic relevance is often unclear. *FOXP2* encodes a transcription factor with a forkhead box DNA-binding domain, but little is known about the functions of protein regions outside this domain.

Methods: We performed detailed functional analyses of seven rare FOXP2 variants found in affected cases, including three which have not been previously characterized, testing intracellular localization, transcriptional regulation, dimerization, and interaction with other proteins. To shed further light on molecular functions of FOXP2, we characterized the interaction between this transcription factor and co-repressor proteins of the C-terminal binding protein (CTBP) family. Finally, we analysed the functional significance of the polyglutamine tracts in FOXP2, since tract length variations have been reported in cases of neurodevelopmental disorder.

Results: We confirmed etiological roles of multiple FOXP2 variants. Of three variants that have been suggested to cause speech/language disorder, but never before been characterized, only one showed functional effects. For the other two, we found no effects on protein function in any assays, suggesting that they are incidental to the phenotype. We identified a CTBP-binding region within the N-terminal portion of FOXP2. This region includes two amino acid substitutions that occurred on the human lineage following the split from chimpanzees. However, we did not observe any effects of these amino acid changes on CTBP binding or other core aspects of FOXP2 function. Finally, we found that FOXP2 variants with reduced polyglutamine tracts did not exhibit altered behaviour in cellular assays, indicating that such tracts are non-essential for core aspects of FOXP2 function, and that tract variation is unlikely to be a highly penetrant cause of speech/language disorder.

Conclusions: Our findings highlight the importance of functional characterization of novel rare variants in FOXP2 in assessing the contribution of such variants to speech/language disorder and provide further insights into the molecular function of the FOXP2 protein.

Keywords: Transcription factor, Speech, Language, Functional genetics, Neuroscience

Background

FOXP2 is a member of the forkhead box (FOX) family of transcription factors and has crucial roles in the development of the brain and other organs [1, 2]. Heterozygous disruptions of the *FOXP2* gene cause a rare and severe speech and language disorder (OMIM 602081) [3]. This disorder was first reported in a three-generation

pedigree (the KE family), in which approximately half of the individuals have difficulties with learning to make the co-ordinated orofacial movements required for speech (childhood apraxia of speech, CAS), together with wide-ranging impairments in comprehension and production of spoken and written language, but without major deficits in other aspects of cognitive functioning [4]. All affected members of the family were found to carry a missense variant in *FOXP2* that alters a critical residue within the DNA-recognition helix of the FOX domain and thus prevents DNA binding and regulation of transcription [4–6]. A number of individuals have since been reported to

* Correspondence: Simon.Fisher@mpi.nl

†Equal contributors

¹Language and Genetics Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, the Netherlands

²Donders Institute for Brain, Cognition and Behaviour, Radboud University, 6525 EN Nijmegen, the Netherlands

present with severe speech/language difficulties together with heterozygous whole gene deletions or chromosomal translocations disrupting *FOXP2*, confirming the necessity of two functional copies of this gene for typical speech and language development [3].

Screening of the *FOXP2* coding region for protein-altering variants has been performed in a few small cohorts of children with speech articulation disorders similar to those reported in the KE family [7–9]. In addition, to address the possibility that *FOXP2* disruption might also be a factor in other disorders characterized by speech/language problems, similar screens have been performed in individuals with specific language impairment, speech sound disorder, autism, schizophrenia, and epilepsy of the speech cortex [10–18]. There are no common non-synonymous variants in *FOXP2* in the general population, and relatively little coding sequence variation has been observed in individuals with speech/language-related disorders, indicating that *FOXP2* disruptions are a rare cause of such disorders, which likely have a highly heterogeneous genetic basis. Nonetheless, screening for *FOXP2* variants in individuals with neurodevelopmental phenotypes has identified a small number of rare protein-altering variants, including five missense variants, one stop-gain variant, one 2-bp deletion resulting in a frame-shift, and several in-frame insertions or deletions of glutamine residues within polyglutamine tracts. However, the contribution of individual rare variants to disorder often remains unclear because the genetic evidence in isolation is insufficient to confirm a causal or contributory role, and the effect of the variant on protein function is unknown.

To clarify the etiological contribution of the rare *FOXP2* variants reported to date in individuals with neurodevelopmental disorders, we performed functional characterization of these variants by assaying their effects on a range of molecular properties. In addition, we characterize the interaction between *FOXP2* and the co-repressors of the C-terminal binding protein (CTBP) family, which may have a central role in *FOXP2*-mediated transcriptional repression. Finally, we provide the first detailed examination of the role of the polyglutamine tract in *FOXP2* function, in order to shed light on the contribution of tract length variation to disorder.

Methods

DNA constructs

The cloning of human *FOXP2* (NM_014491), *FOXP1* (NM_032682), and *CTBP1* (NM_001328) and mouse *Foxp2* (NM_053242) has been described previously [6, 19]. The coding sequence of *CTBP2* (NM_001329) was amplified from human foetal brain cDNA using the primers listed in Additional file 1. The *FOXP2* p.Q17L, p.M406T, p.P416T, p.R553H, p.N597H, p.N303T, p.S325N, and p.N303T/p.S325N variants were generated by site-

directed mutagenesis using the Quick-Change Site-Directed Mutagenesis kit (Stratagene) following the manufacturer's protocol. (Note that the numbering of *FOXP2* variants throughout this manuscript is given with respect to the coding sequence of the predominant isoform, NM_014491, which normally encodes a 715 amino-acid protein.) Primers used in site-directed mutagenesis are listed in Additional file 2. The *FOXP2* p.R328* variant and synthetic truncated forms of *FOXP2* were generated using the primers listed in Additional file 1. Synthetic *FOXP2* variants with reduced polyglutamine tracts were generated using a PCR-based strategy [20]. All *FOXP2* variants were initially generated in an intermediary plasmid, pCR2.1-TOPO (Life Technologies), and the entire coding sequence of *FOXP2* was verified by Sanger sequencing before being subcloned into the final expression vector. For expression of fusion proteins with *Renilla* luciferase, yellow fluorescent protein (YFP), mCherry, and three tandem N-terminal Myc tags, cDNAs were subcloned into the pLuc, pYFP, pmCherry, and pMyc expression vectors, respectively, which have been described previously [19–21]. The SRPX2 luciferase reporter plasmid was generated by subcloning a 1146-bp region of the SRPX2 promoter into the promoterless firefly luciferase vector pGL4.23 (Promega), as described previously [21]. All constructs were verified by Sanger sequencing. Plasmid sequences are available upon request.

Cell culture and transfection

HEK293 were obtained from ECACC (cat. no. 85120602) and cultured in DMEM supplemented with 10% foetal bovine serum. Transfections were performed using GeneJuice (Merck-Millipore) according to the manufacturer's instructions.

Western blotting

HEK293 cells were transfected in 6-well plates and cultured for 48 h. Cells were lysed for 10 min at 4 °C with 100 mM Tris pH 7.5, 150 mM NaCl, 10 mM EDTA, 0.2% Triton X-100, 1% PMSE, and protease inhibitor cocktail. Cell lysates were cleared by centrifugation at 10,000×g for 3 min at 4 °C. Proteins were resolved on 10% SDS-polyacrylamide gels and transferred to PVDF membranes using a TransBlot Turbo Blotting apparatus (Bio-Rad). Membranes were blocked in PBS containing 5% milk and 0.1% Tween-20 and incubated overnight at 4 °C with primary antibody. The following antibodies were used: anti-GFP (Clontech cat. no. 632380, 1:8000, for YFP constructs); anti-Myc tag (Abcam cat. no. ab9106, 1:1000); and anti-β-actin (Sigma cat. no. A5441, 1:10,000). After washing, membranes were incubated with horseradish peroxidase-conjugated goat anti-mouse or anti-rabbit IgG for 45 min at room temperature. Proteins were visualized

using Novex ECL Chemiluminescent Substrate Reagent Kit (Invitrogen) and a ChemiDoc XRS+ imaging system (Bio-Rad).

BRET assay

Bioluminescence resonance energy transfer (BRET) assays were performed as described [19]. Briefly, HEK293 cells were transfected in 96-well plates with DNA plasmids encoding YFP- and luciferase-fusion proteins. After 36–48 h, Enduren Live Cell Luciferase Substrate (Promega) was added at a final concentration of 60 μ M. Cells were cultured for a further 4 h, and emission readings (integrated over 10 s) were taken using a TECAN F200PRO microplate reader using the Blue1 and Green1 filter sets. Expression levels of the YFP-fusion proteins were measured by taking fluorescent readings using the filter set and dichroic mirror suitable for green fluorescent protein (excitation 480 nm, emission 535 nm). The corrected BRET ratio was calculated with the following formula: $[\text{Green1}_{(\text{experimental condition})}/\text{Blue1}_{(\text{experimental condition})}] - [\text{Green1}_{(\text{control condition})}/\text{Blue1}_{(\text{control condition})}]$. The control conditions used luciferase or YFP fused to a C-terminal nuclear localization signal.

Fluorescence microscopy

HEK293 cells were seeded on coverslips coated with poly-L-lysine. Cells were cultured for 30 h post-transfection and then fixed with methanol. Nuclei were stained with Hoechst 33342. Fluorescence images were acquired using an Axiovert A-1 fluorescent microscope with ZEN Image software (Zeiss).

Fluorescence-based quantitation of protein expression levels

HEK293 cells were transfected in triplicate with YFP-tagged FOXP2 variants and mCherry, in clear-bottomed black 96-well plates. Cells were cultured at 37 °C with 5% CO₂ in a TECAN M200PRO microplate reader equipped with a Gas Control Module for live-cell kinetic assays. Fluorescence intensity was measured 48 h post-transfection. For each well and time point, the background-subtracted YFP intensity was divided by the background-subtracted mCherry intensity. Triplicate conditions were averaged.

Luciferase reporter assays

HEK293 cells were seeded in clear-bottomed white 96-well plates and transfected in triplicate. For the SV40 assay, cells were transfected with 12 ng of pGL3-promoter firefly luciferase reporter construct containing the SV40 promoter (Promega), 5 ng of pRL-TK *Renilla* luciferase normalization control (Promega), and 16 ng of YFP-FOXP2 (wild-type or variant) or YFP control construct. For the SRPX2 assay, cells were transfected with 4.3 ng of SRPX2 luciferase reporter construct, 5 ng of pGL4.74

Renilla luciferase normalization control (Promega), and 45 ng of YFP-FOXP2 (wild-type or variant) or YFP control construct. After 48 h, luciferase activity was measured in a TECAN F200PRO microplate reader using the Dual-Luciferase Reporter Assay system (Promega).

Statistical analysis

The statistical significance of the luciferase reporter assays and BRET assays was analysed using a one-way analysis of variance (ANOVA) followed by Bonferroni's post hoc correction.

Results

Rare FOXP2 variants implicated in neurodevelopmental disorder

We examined seven rare FOXP2 variants that have been observed in individuals with neurodevelopmental disorders, including five missense variants, one stop-gain variant, and one frameshift variant (Table 1, Fig. 1a). For three of the variants examined (p.N597H, p.P416T, p.Q390Vfs*7), this is the first report of any functional characterization. The remaining four variants have been studied previously to varying extents and are included here for comparison with the uncharacterized variants. All of the variants were additionally characterized using novel assays that have not been used to study FOXP2 in prior literature. Two of the variants are of particular interest because of their uncertain significance with regard to the phenotype of the affected cases. The p.N597H variant was found by targeted sequencing in a proband with CAS and was described as a likely pathogenic variant, but it was not ascertained if this variant occurred de novo, and no functional characterization was performed [7]. The p.M406T variant was identified in a proband with an epilepsy-aphasia spectrum disorder (focal epilepsy with continuous spike-and-waves during sleep), cognitive and language deficits, and polymicrogyria of the left rolandic operculum [15]. The variant was also carried by two siblings who were not known to have any neurological abnormality and was inherited from the father, who did not display any neurological or MRI abnormalities. It was suggested that this variant plays an etiological role in the epilepsy-aphasia spectrum disorder observed in the proband, acting as a risk factor with incomplete penetrance [15]. However, the contribution of this variant to the phenotype is tentative because of the lack of segregation with the disorder, the atypical phenotypic presentation in comparison to other cases of FOXP2 disruption, and the limited functional characterization performed to date.

The seven FOXP2 variants were expressed as fusion polypeptides with YFP in HEK293 cells. The expression of these YFP-tagged variants (and of all other YFP-tagged forms of FOXP2 used in this study) was verified by western blotting, as shown in Additional file 3. Direct measurement of fluorescence intensity in live cells indicated that the

Table 1 Rare FOXP2 variants in individuals with neurodevelopmental disorders

Variant description ^a	Frequency in ExAC ^b	Phenotype and inheritance pattern	Role in disorder ^c	Reference
p.Q17L (missense) c.50A > T chr7:114426561A > T rs201649896	56/120570 (1 homozygote)	Found in a proband with CAS but not in an affected sibling. Parental genotypes not determined	Probably incidental	MacDermot et al. [8]
p.M406T (missense) c.1217 T > C chr7:114653960 T > C no rs ID	Not observed	Found in a proband with rolandic epilepsy and polymicrogyria, in two unaffected siblings and in unaffected father	Uncertain significance	Roll et al. [15]
p.P416T (missense) c.1246C > A chr7:114653989C > A rs369313543	1/121328	Present in two siblings with severe stuttering. Absent in affected father. Inherited from mother, who does not stutter but has oral motor impairments	Probably incidental	Turner et al. [9]
p.R553H (missense) c.1658G > A chr7:114662075G > A rs121908377	Not observed	Segregates with CAS in three generations of the KE family	Causal	Lai et al. [4]
p.N597H (missense) c.1789A > C chr7:114663469A > C no rs ID	1/120986	Found in a proband with CAS. Parental genotypes not determined	Uncertain significance	Laffin et al. [7]
p.R328* (stop-gain) c.982C > T chr7:114642616C > T rs121908378	Not observed	Present in a proband with CAS and in affected sibling. Inherited from affected mother	Causal	MacDermot et al. [8]
p.Q390Vfs*7 (frameshift) c.1168_1169del chr7:114652276_114652277del no rs ID	Not observed	De novo variant in a proband with sporadic CAS, dysarthria, and fine motor apraxia	Causal	Turner et al. [9]

^aVariants are described in accordance with Human Genome Variation Society recommendations (www.hgvs.org/mutnomen, accessed June 2016) with reference to the major transcript NM_014491.3 (ENST00000350908). Genomic coordinates refer to the hg38 assembly. The rs ID number is provided for variants that are present in dbSNP

^bVariant allele frequency in the Exome Aggregation Consortium (ExAC) dataset (<http://exac.broadinstitute.org>, accessed June 2016)

^cVariants are described as causal if they segregate perfectly with childhood apraxia of speech (CAS) affection status, or if they occurred de novo in sporadic cases, and if they additionally have been demonstrated to cause loss of protein function, or are very likely to do so because of protein truncation. Variants are described as probably incidental if they do not segregate with CAS and are observed in the Exome Aggregation Consortium (ExAC) dataset. Other variants are described as of uncertain significance

Note: This table does not include newly described variants that were reported by Reuter et al. [58] after the completion of the present study

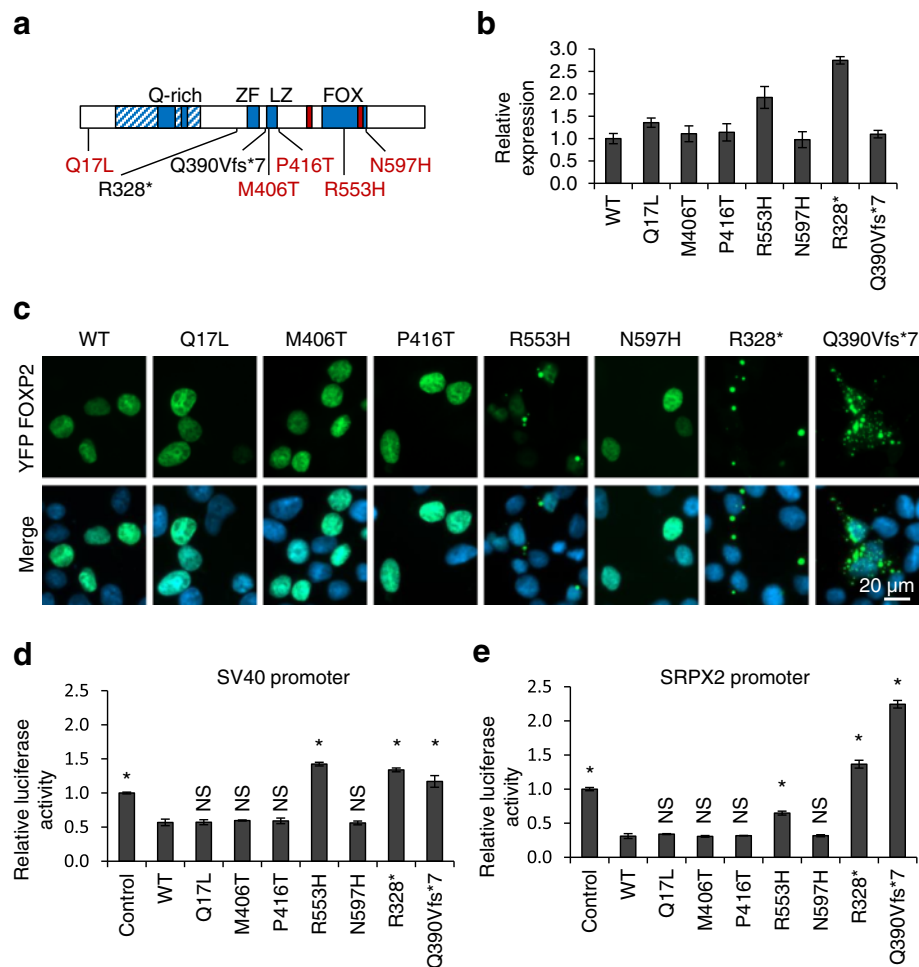


Fig. 1 Functional characterization of rare FOXP2 variants. **a** Schematic representation of the FOXP2 protein showing rare variants found in individuals with neurodevelopmental disorders. Stop-gain and frameshift variants are shown in black and missense variants in red. Known domains are labelled: glutamine-rich (*Q-rich*) region (hatched shading) including polyglutamine tracts (*solid shading*), zinc finger (ZF), leucine zipper (LZ) and forkhead domain (FOX). Nuclear localization signals are indicated with red bars. **b** Fluorescence-based measurement of FOXP2 expression level. HEK293 cells were transfected with YFP-FOXP2, together with mCherry for normalization. Fluorescence intensity was measured 48 h post-transfection. Values are mean YFP/mCherry fluorescence ratios \pm S.D. ($n = 3$), expressed relative to the value for wild-type (WT) FOXP2. **c** Fluorescence micrographs of HEK293 cells transfected with YFP-FOXP2. Nuclei were stained with Hoechst 33342. **d, e** Luciferase reporter assays for transcriptional regulatory activity of FOXP2. HEK293 cells were transfected with a firefly luciferase reporter vector containing the SV40 promoter (**d**) or the human *SRPX2* promoter (**e**), together with a *Renilla* luciferase normalization plasmid and YFP-FOXP2, or YFP alone (control). Values are mean relative luciferase activity \pm S.D. ($n = 3$), expressed relative to the control. Asterisks indicate significant differences compared to wild-type (WT) FOXP2 ($p < 0.05$, one-way ANOVA followed by Bonferroni post hoc correction). NS not significant. Exact p values for both the SV40 and the *SRPX2* assays are < 0.0001 for the control and the R553H, R328*, and Q390Vfs*7 variants, and > 0.9999 for all other variants

missense variants and the truncated p.Q390Vfs*7 variant are expressed at broadly similar levels to the wild-type protein, whereas the p.R553H and p.R328* variants may have slightly increased expression (Fig. 1b). In cells of people carrying the p.R328* and p.Q390Vfs*7 variants, the expression levels of these protein variants are likely to be very low due to nonsense-mediated decay of the aberrant transcripts, although this has not been formally tested due to the lack of FOXP2 expression in accessible tissue [8, 9]. The inclusion of these truncated variants in functional assays provides a useful comparison with missense variants because the

truncated variants lack the DNA-binding domain and hence cannot function normally in transcriptional regulation.

Effects of variants on nuclear localization and transcriptional regulation

The effect of the variants on FOXP2 localization was examined by direct imaging of the YFP-fusion proteins in transfected cells (Fig. 1c). Wild-type FOXP2 exhibited nuclear localization, with exclusion from nucleoli, as reported previously (Fig. 1c) [6, 22]. The two previously

characterized etiological variants, p.R553H and p.R328*, showed disrupted nuclear localization and formation of aggregates, also consistent with earlier reports (Fig. 1c) [6, 22]. The uncharacterized Q390Vfs*7 variant formed cytoplasmic aggregates similar to those observed for the p.R328* variant, consistent with the loss of endogenous nuclear localization signals in these truncated variants (Fig. 1a,c) [22]. The p.Q390Vfs*7 variant therefore has a similarly deleterious effect on protein localization as the two known etiological variants. None of the other missense variants showed any sign of abnormal localization, including the putatively pathogenic p.N597H and p.M406T variants (Fig. 1c). Altered localization has been reported previously for the p.M406T variant, but in multiple independent experiments, we did not observe any loss of nuclear localization for this variant [15]. Thus, with the exception of the p.R553H variant, all the missense variants retain the nuclear localization necessary for transcriptional regulatory activity.

To assess the ability of the FOXP2 variants to regulate transcription, we performed luciferase reporter assays using the SV40 promoter, a viral promoter which is repressed by FOXP2 [6, 23]. The etiological p.R553H and p.R328* variants exhibited the expected loss of repression activity in this assay [6] (Fig. 1d). The uncharacterized p.Q390Vfs*7 variant showed a comparable loss of repression to the other etiological variants (Fig. 1f). In contrast, the remaining four missense variants showed similar activity to the wild-type protein (Fig. 1d). To verify these results using an endogenous human FOXP2 target gene, we performed luciferase reporter assays using a region of the *SRPX2* promoter [15]. As observed for the SV40 promoter, the truncated variants and the p.R553H variant showed loss of repressive activity, whereas the other missense variants did not differ significantly in activity from the wild-type protein (Fig. 1e). The p.M406T variant has previously been reported to show reduced transcriptional repression activity in relation to the *SRPX2* promoter [15]. Small differences between the promoter regions and vector backbones used in this and the previous study might account for these conflicting results. Nonetheless, our data indicate that the p.M406T variant does not exhibit a generalized reduction in transcriptional regulatory activity.

Effects of variants on protein dimerization

FOXP2 and other proteins of the FOXP subfamily form dimers via their leucine zipper domains, a property which appears to be essential for transcriptional regulatory activity [24] (Fig. 1a). In FOXP3, loss of dimerization capacity as a result of an in-frame single amino acid deletion in the leucine zipper domain results in an immunological disorder known as IPEX syndrome (MIM 304790), with disease severity comparable to that resulting from loss of DNA-binding activity in FOXP3 [25, 26]. Variants

affecting the dimerization of FOXP2 might therefore be a cause of speech/language disorder. Additionally, FOXP2 variants that have lost DNA-binding capacity but retain dimerization ability might interfere with the function of wild-type protein in cells of affected individuals, and thus exert a dominant-negative effect. Notably, FOXP2 not only forms homodimers but can also heterodimerize with FOXP1 and FOXP4, proteins which have partially overlapping expression patterns in the developing brain [24, 27–29]. (Note that FOXP3 is expressed only in haematopoietic cells.) Heterodimerization among certain FOXP proteins may therefore be an important mechanism for differential regulation of target genes in different neuronal subtypes. Of note, heterozygous disruption of *FOXP1* results in a severe neurodevelopmental phenotype that includes language deficits, potentially reflecting dysregulation of some of the same target genes impacted by *FOXP2* disruption [30, 31].

Two of the five FOXP2 missense variants reported in individuals with neurodevelopmental disorders lie within or near the leucine zipper domain, which spans residues p.V388-L409 (Fig. 1a). These variants (p.M406T and p.P416T) could therefore affect FOXP2 dimerization. The interaction between FOXP2 variants and wild-type FOXP2 was assayed using a BRET assay, which enables protein-protein interactions to be monitored in live cells [19]. The interaction of the FOXP2 variants with wild-type FOXP1 and FOXP4 was also examined. As expected, wild-type FOXP2 was able to homodimerize and also to heterodimerize with FOXP1 and FOXP4 (Fig. 2a–c) [24]. The truncated p.R328* and p.Q390Vfs*7 variants were not able to dimerize with the wild-type proteins, consistent with the loss of the leucine zipper domain in these variants (Fig. 2a–c). Thus, even if these abnormal proteins are present in cells of affected individuals despite the activation of nonsense-mediated decay mechanisms, they could not interfere with the function of wild-type proteins via dimerization. In contrast, the p.R553H variant found in the KE family has a normal leucine zipper domain and showed only a modest reduction in interaction with wild-type FOXP2, FOXP1, and FOXP4 (Fig. 2a–c). Furthermore, co-transfection of wild-type FOXP2 with the p.R553H variant led to some wild-type protein being mislocalized to the cytoplasm (Fig. 2d). Mislocalization of wild-type protein was not observed upon co-transfection with the non-interacting variants p.R328* and p.Q390Vfs*7 (Fig. 2d). The p.R553H variant might therefore interfere with the function of wild-type FOXP2 in cells of people carrying this variant, and this effect could contribute to the phenotype observed in the affected members of the KE family. All four of the other missense variants generally demonstrated normal dimerization capacity (Fig. 2a–c). The p.P416T variant showed a statistically significant difference in interaction with FOXP1 (Fig. 2b),

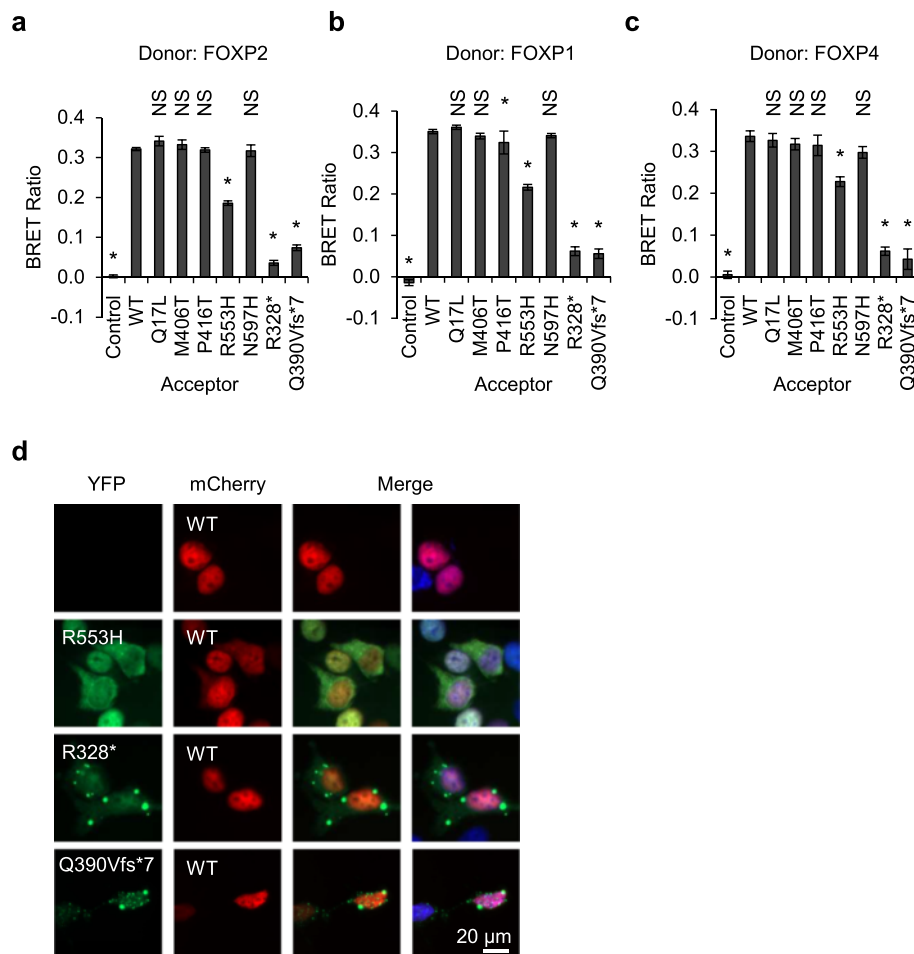


Fig. 2 Interaction of rare FOXP2 variants with wild-type FOXP proteins. **a–c** Bioluminescence resonance energy transfer (BRET) assays for interaction of FOXP2 variants with wild-type (WT) FOXP proteins. HEK293 cells were transfected with YFP-FOXP2 variants (acceptor) and *Renilla* luciferase (donor) fusions of FOXP2 (**a**), FOXP1 (**b**) or FOXP4 (**c**). The control acceptor protein is a nuclear-targeted YFP. Values are mean corrected BRET ratios \pm S.D. ($n = 3$). Asterisks indicate significant differences compared to wild-type (WT) FOXP2 ($p < 0.05$, one-way ANOVA followed by Bonferroni post hoc correction). NS not significant. Exact p values for **a** are <0.0001 for the control and the R553H, R328* and Q390Vfs*7 variants and >0.9999 for all other variants. Exact p values for **b** are <0.0001 for the control and the R553H, R328* and Q390Vfs*7 variants, 0.024 for the P416T variant, and >0.9999 for all other variants. Exact p values for **c** are <0.0001 for the control and the R553H, R328* and Q390Vfs*7 variants, 0.088 for the N597H variant, and >0.9999 for all other variants. **d** Fluorescence micrographs of HEK293 cells transfected with YFP-fusions of the FOXP2 variants p.R553H, p.R328* and p.Q390Vfs*7, together with wild-type FOXP2 fused to mCherry. Nuclei were stained with Hoechst 33342

but this difference was not consistently observed across independent experiments. Therefore, despite being located in or near the leucine zipper domain, the p.M406T and p.P416T variants do not appear to affect FOXP2 dimerization and are unlikely to exert a pathogenic effect via this mechanism.

Effects of variants on interaction with the co-repressors CTBP1 and CTBP2

A further mechanism by which variants in FOXP2 might lead to speech/language disorder is through disruption of protein-protein interactions between FOXP2 and crucial mediators of transcriptional regulation. Relatively, little is

known about the interaction partners of FOXP2; however, proteins of the CTBP family have been identified as candidate FOXP2 interactors in multiple independent yeast two-hybrid screens [24, 32–34]. The CTBP family consists of two proteins, CTBP1 and CTBP2, which function as co-repressors for multiple transcription factors [35, 36]. The interaction between FOXP2 and CTBP1 has previously been validated by us using the BRET assay, and by others via co-immunoprecipitation, but the interaction with CTBP2 has not yet been confirmed [19, 24]. Using the BRET assay, we validated the interaction between CTBP2 and FOXP2 and found that this interaction produces a notably higher BRET signal than the

interaction between CTBP1 and FOXP2 (Fig. 3a,b). The difference in the magnitude of the BRET signal may reflect subcellular localization differences between the CTBPs: CTBP2 is wholly nuclear and therefore shows a large degree of co-localization with wild-type FOXP2, whereas CTBP1 is found in both the cytoplasm and nucleus, and therefore has an overall lower degree of co-localization with FOXP2 (Fig. 3c) [37, 38]. Strikingly, we found that CTBP1 and CTBP2 are also able to interact with FOXP1, but not with FOXP4, in line with previously reported findings for CTBP1 (Fig. 3a,b) [24]. The lack of interaction between FOXP4 and CTBPs suggests that the FOXP proteins have partially divergent mechanisms for transcriptional regulation, which is particularly interesting given the ability of FOXP2 to dimerize with FOXP4 (Fig. 2c).

We next evaluated the effects of rare variants in FOXP2 on the interaction with CTBP1/2. The p.R553H variant found in the KE family showed slightly reduced interaction with CTBP2, but increased interaction with CTBP1, when compared with wild-type FOXP2 (Fig. 3d,e). Interestingly, the p.R553H variant displayed a mixed nuclear and cytoplasmic localization when transfected alone or with CTBP2, but became predominantly cytoplasmic in the presence of CTBP1 (Figs. 1e and 3f,g). In contrast, the nuclear localization of wild-type FOXP2 was not affected by co-transfection with CTBP1 (Fig. 3c). The cytoplasmic retention of the p.R553H variant, but not of wild-type FOXP2, upon co-transfection with CTBP1 may be a consequence of the loss of DNA-binding capacity in the p.R553H variant. These co-transfection experiments indicate that the p.R553H variant shows greater co-localization with CTBP1 but reduced co-localization with CTBP2, compared with wild-type FOXP2, consistent with the BRET data on interactions between these proteins (Fig. 3d,e). Although altered interaction with CTBP1/2 probably does not play a substantive role in the speech/language pathology in the KE family, the retained interaction between the p.R553H variant and CTBP1/2 is notable because it indicates that DNA-binding activity is not a prerequisite for interaction of FOXP2 with co-repressors of the CTBP family. The four other missense variants did not differ substantially from the wild-type protein in their interaction with CTBP1/2 in the BRET assay (Fig. 3d,e). These four variants thus have comparable properties to wild-type FOXP2 in assays of subcellular localization, transcriptional repression, protein dimerization, and interaction with co-repressor proteins, suggesting that all of these variants are benign rare polymorphisms (Table 2). In particular, we do not find any support for the etiological roles suggested previously for the p.M406T and p.N597H variants [7, 15].

Interestingly, the p.R328* and p.Q390Vfs*7 variants partially retained the ability to interact with CTBP1/2, despite lacking a large proportion of the normal FOXP2

polypeptide (Fig. 3d,e). The interaction with CTBP1 in particular was only mildly affected (Fig. 3d). Furthermore, co-transfection of CTBP1 with the p.R328* and p.Q390Vfs*7 variants resulted in clear co-localization of CTBP1 with the truncated variants within cytoplasmic aggregates (Fig. 3d,f). The interaction with CTBP2 was more severely affected, but co-localization of CTBP2 and the p.R328* variant within aggregates was still readily apparent (Fig. 3e,g). The interaction between these severely truncated FOXP2 variants and CTBP1/2 suggests that some key determinants of CTBP binding lie near the N-terminus of FOXP2.

Characterization of the interaction between FOXP proteins and CTBPs

The retention of CTBP binding by the severely truncated p.R328* and p.Q390Vfs*7 variants (Fig. 3d,e) prompted us to investigate further the interaction between FOXP2 and CTBPs. The CTBP binding site in FOXP2 was previously suggested to be a PLNLV motif at residues 422–426, due to the similarity of this motif to the consensus CTBP-binding motif PXDLS (Fig. 4a) [24, 36]. The PLNLV motif is conserved in FOXP1 but not FOXP4, consistent with the lack of interaction between CTBP1 and FOXP4 [24] (Fig. 3a,b). However, the residual interaction of CTBPs with the p.R328* and p.Q390Vfs*7 variants indicates that the PLNLV motif is not essential for the interaction and that the CTBP binding site is at least partially localized to a more N-terminal region of FOXP2 (Fig. 3d,e). Furthermore, one of the yeast two-hybrid screens that identified CTBP1 and CTBP2 as FOXP2 interactors employed a fragment of FOXP2 encompassing residues 122–382, indicating that this region, which lacks the PLNLV motif, is sufficient for the interaction [34].

To map the region of FOXP2 involved in binding to CTBP2, we performed a BRET assay using a series of synthetic truncated versions of FOXP2 (Fig. 4a) [21]. Truncated protein variants lacking the endogenous nuclear localization signals were targeted to the nucleus by addition of an artificial localization signal to the protein C-terminus (Fig. 4a,b). The efficacy of this truncation series in mapping interaction domains was demonstrated using a BRET assay for FOXP2 homodimerization (Fig. 4c). Truncation of FOXP2 after residue 423 did not interfere with homodimerization, whereas truncation after residue 329 dramatically reduced the interaction, reflecting the loss of the critical leucine zipper domain (Fig. 4c) [24].

In a BRET assay for interaction with CTBP2, a FOXP2 variant truncated after residue 487 showed a similar level of interaction to the full-length protein (Fig. 4d), indicating that the C-terminal region of FOXP2 including the FOX domain is not involved in the interaction with CTBPs. More severely truncated forms of FOXP2 showed a progressive reduction in interaction with CTBP2. The

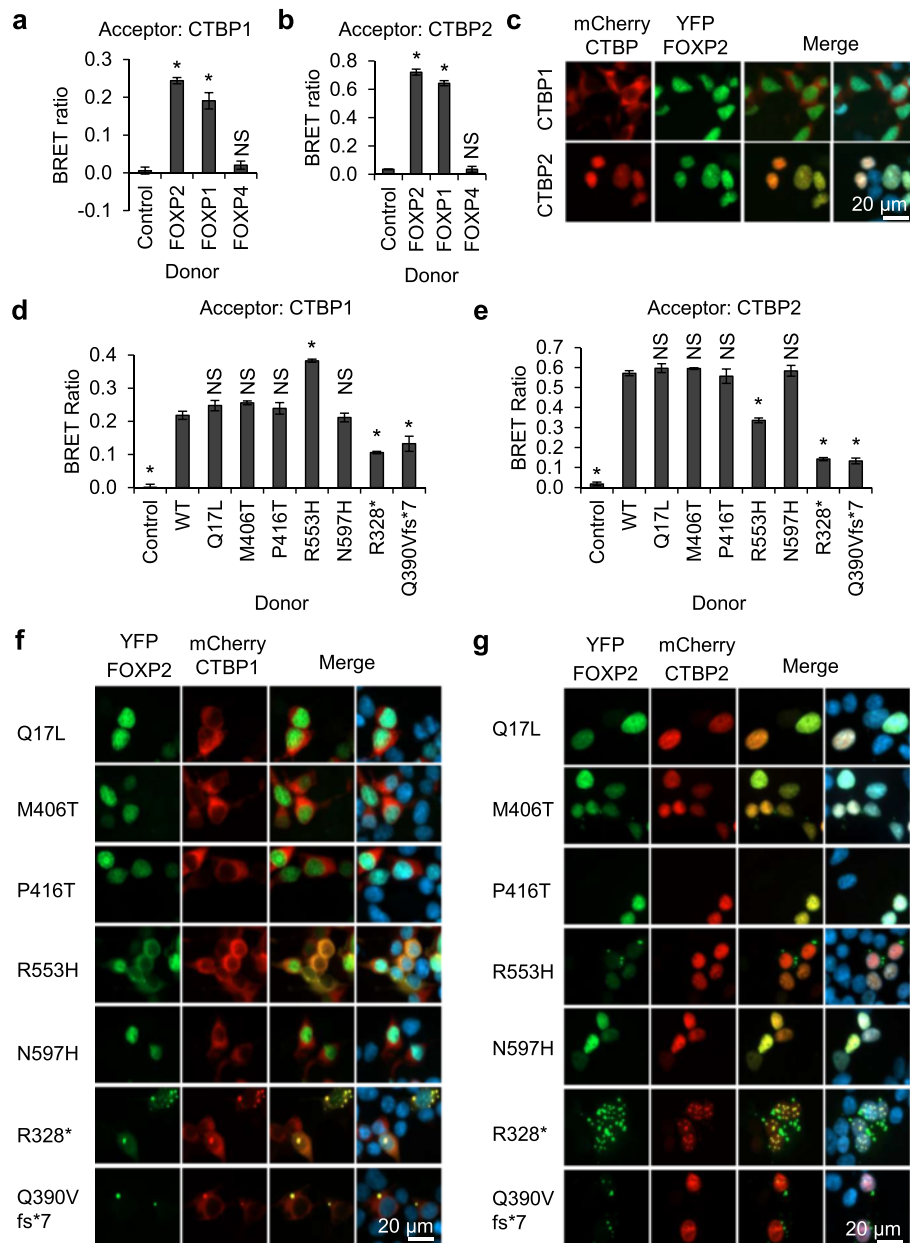


Fig. 3 Interaction of rare FOXP2 variants with CTBP1/2. **a, b** BRET assays for interaction of FOXP proteins with CTBP1 and CTBP2. HEK293 cells were transfected with FOXP1, FOXP2, or FOXP4 fused to *Renilla* luciferase (donor) and CTBP1 (**a**) or CTBP2 (**b**) fused to YFP (acceptor). The control donor protein is a nuclear-targeted luciferase. Values are mean corrected BRET ratios \pm S.D. ($n = 3$). Asterisks indicate significant differences compared to control ($p < 0.05$, one-way ANOVA followed by Bonferroni post hoc correction). NS not significant. Exact p values for both **a** and **b** are <0.0001 for FOXP2 and FOXP1 and >0.9999 for FOXP4. **c** Fluorescence micrographs of HEK293 cells transfected with mCherry-CTBP1 or mCherry-CTBP2, together with YFP-FOXP2. Nuclei were stained with Hoechst 33342. **d, e** BRET assays for interaction of rare FOXP2 variants with CTBP1 and CTBP2. HEK293 cells were transfected with FOXP2 variants fused to *Renilla* luciferase (donor) and CTBP1 (**d**) or CTBP2 (**e**) fused to YFP (acceptor). The control donor protein is a nuclear-targeted luciferase. Values are mean corrected BRET ratios \pm S.D. ($n = 3$). Asterisks indicate significant differences compared to wild-type (WT) FOXP2 ($p < 0.05$, one-way ANOVA followed by Bonferroni post hoc correction). NS not significant. Exact p values for **d** are <0.0001 for the control and the R553H, R328* and Q390Vfs*7 variants, 0.221 for the M406T variant, and >0.9999 for all other variants. Exact p values for **e** are <0.0001 for the control and the R553H, R328* and Q390Vfs*7 variants and >0.9999 for all other variants. **f, g** Fluorescence micrographs of HEK293 cells transfected with YFP-FOXP2 and mCherry-CTBP1 (**f**) or mCherry-CTBP2 (**g**). Nuclei were stained with Hoechst 33342

Table 2 Summary of functional characterization of rare FOXP2 variants in individuals with neurodevelopmental disorders

Variant ^a	Localization	Transcriptional repression		Dimerization with FOXP2s			Interaction with CTBPs		Role in disorder ^b
		SV40	SRPX2	FOXP2	FOXP1	FOXP4	CTBP1	CTBP2	
p.Q17L	+	+	+	+	+	+	+	+	Incidental
p.M406T	+	+	+	+	+	+	+	+	Incidental
p.P416T	+	+	+	+	+	+	+	+	Incidental
p.R553H	-	-	-	-	-	-	-	-	Causal
p.N597H	+	+	+	+	+	+	+	+	Incidental
p.R328*	-	-	-	-	-	-	-	-	Causal
P.Q390Vfs*7	-	-	-	-	-	-	-	-	Causal

+ Behaviour of the variant is comparable to that of the wild-type protein in this assay, - Behaviour of the variant differs from that of the wild-type protein in this assay

^aVariants are described in accordance with Human Genome Variation Society recommendations (www.hgvs.org/mutnomen, accessed June 2016) with reference to the major transcript NM_014491.3 (ENST00000350908)

^bProbable role of the variant in the disorder in the affected individual, based on the results of functional characterization

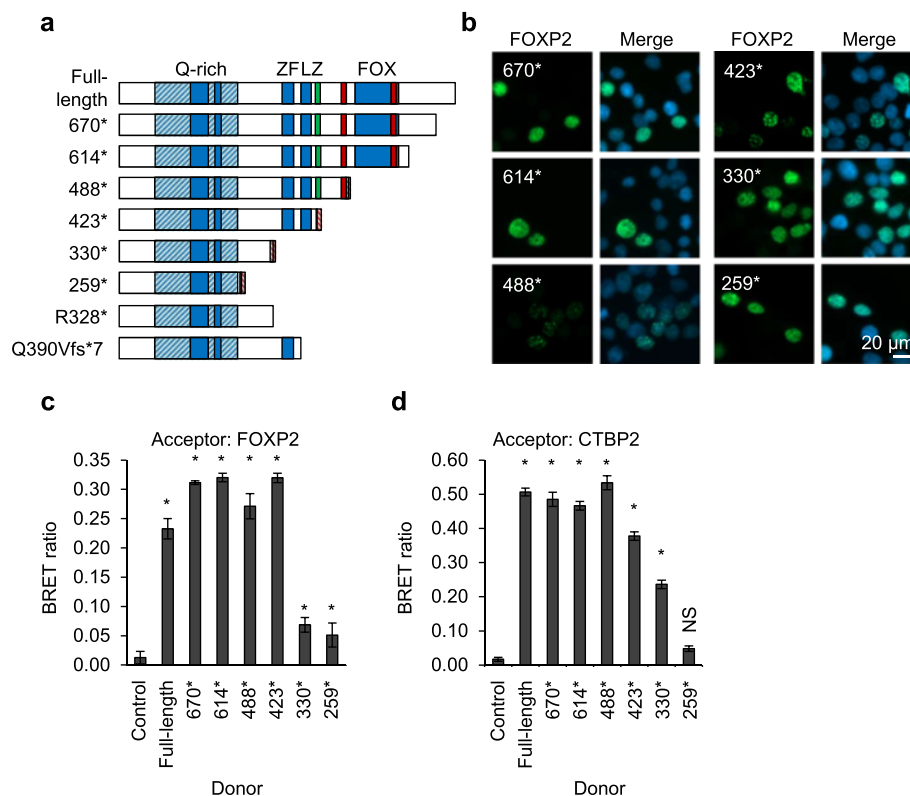


Fig. 4 Mapping of the CTBP binding site in FOXP2. **a** Schematic representation of synthetic truncated forms of FOXP2. The R328* and Q390Vfs*7 variants identified in patients are included for comparison. Known domains are labelled: glutamine-rich (Q-rich) region (hatched shading) including polyglutamine tracts (solid shading), zinc finger (ZF), leucine zipper (LZ), and forkhead domain (FOX). The PLNLV motif is indicated with a green bar. Nuclear localization signals are indicated with red bars. A synthetic nine-residue nuclear targeting sequence (hatched red bars) was appended to the C-terminus of variants which lack one or both of the endogenous nuclear localization signals. **b** Fluorescence micrographs of HEK293 cells transfected with synthetic truncated FOXP2 variants. Nuclei were stained with Hoechst 33342. **c, d** BRET assay for interaction of synthetic truncated FOXP2 variants with full-length FOXP2 and CTBP2. HEK293 cells were transfected with truncated FOXP2 variants fused to *Renilla* luciferase (donor) and FOXP2 (**c**) or CTBP2 (**d**) fused to YFP (acceptor). The control donor protein is a nuclear-targeted luciferase. Values are mean corrected BRET ratios \pm S.D. ($n = 3$). Asterisks indicate significant differences compared to control ($p < 0.05$, one-way ANOVA followed by Bonferroni post hoc correction). NS not significant. Exact p values for **c** are 0.002 for FOXP2.330*, 0.025 for FOXP2.259* and < 0.0001 for all other variants. Exact p values for **d** are 0.001 for FOXP2.423*, 0.012 for FOXP2.330*, 0.97 for FOXP2.259*, and < 0.0001 for all other variants

substantial retention of interaction between CTBP2 and FOXP2 forms truncated after residue 422 or residue 329 indicates that the PLNLV motif, zinc finger, and leucine zipper domains are not essential for interaction, although they may enhance it, and furthermore, FOXP2 dimerization is not necessary for CTBP2 binding (Fig. 4d). Only the shortest FOXP2 variant, which is truncated after residue 258, showed a complete or near-complete loss of interaction (Fig. 4d). The protein region including residues 259–329 may therefore be sufficient for interaction with CTBP2.

The region of FOXP2 encompassing residues 259–329 corresponds to a single exon in which non-synonymous nucleotide substitutions between different species are relatively common in comparison to the exons encoding highly conserved elements such as the FOX and leucine zipper domains [39–41]. In particular, this exon includes two non-synonymous nucleotide substitutions that occurred after the split between the human and chimpanzee lineages, but before the split between modern humans and Neanderthals (Fig. 5a) [39, 42]. Mice carrying a ‘humanized’ version of *Foxp2*, containing these two recent amino acid changes, exhibit altered neuronal morphology and electrophysiology and subtle behavioural changes [43–45]. Strikingly, the morphological and electrophysiological changes have a direction of effect opposite to that observed in mice with only one functional copy of *Foxp2*, suggesting that the changes might have been selected for due to an enhancement of protein function [1, 46].

The molecular mechanisms underlying the neurobiological changes observed in mice carrying humanized *Foxp2* are unknown, but might entail an alteration in the strength of the interaction between *Foxp2* and one or more proteins involved in *Foxp2*-mediated transcriptional regulation. Therefore, to test if the recent amino acid changes affect interaction with CTBPs, we introduced the ancestral amino acid residues, individually and in combination, into human FOXP2 (Fig. 5a). Versions of human FOXP2 carrying the ancestral amino acids were compared with the natural human and mouse proteins in functional assays. Introduction of the ancestral amino acids did not result in any change in protein localization or expression, and furthermore, the human and mouse versions of the protein were indistinguishable in these assays (Fig. 5b,c). Notably, the versions of FOXP2 carrying the ancestral amino acids also did not differ from the natural human protein in assays of transcriptional repression using the SV40 and *SRPX2* promoters, and there was no difference between the human and mouse orthologues in these assays (Fig. 5d,e). Homodimerization was also comparable for all variants tested (Fig. 5f). Finally, the versions of FOXP2 carrying the ancestral amino acids did not differ from the natural human protein in their interaction with CTBP1 and CTBP2 (Fig. 5g,h). Therefore, despite the CTBP binding site in FOXP2 encompassing the

protein region that includes the recent amino acid changes, it is unlikely that an alteration in the strength of the interaction between FOXP2 and CTBP1/2 is responsible for the neurobiological differences observed in mice carrying a partially humanized version of *Foxp2*.

The FOXP2 polyglutamine tract and neurodevelopmental disorder

FOXP2 contains a large polyglutamine tract of 40 residues (p.Q152-Q191) and a small tract of 10 residues (p.Q200-Q209), separated by 8 residues of intervening polypeptide (Fig. 6a). The tracts are located in the N-terminal portion of the protein, within a larger glutamine-rich region (Fig. 6a). Of note, the large tract in FOXP2 is the longest polyglutamine tract in any human protein in healthy individuals [47]. Abnormal expansions of polyglutamine tracts in at least nine different proteins result in neurodegenerative diseases [48]. However, FOXP2 is not regarded as a strong candidate for involvement in such disorders because its large polyglutamine tract shows relatively little length variation in the human population [47]. The mixture of CAG and CAA codons without long pure CAG repeats probably makes the tract less prone to expansion, and no expansions were detected in a sample of 142 individuals with progressive movement disorders [47, 49]. However, some rare variations in tract length have been observed in individuals with neurodevelopmental disorders (Additional file 4) [8, 11, 13, 14, 17, 18]. Most of the observed variants are deletions of 1–6 residues from the large tract (Additional file 4). Such deletions have also been observed in the general population (Exome Aggregation Consortium (ExAC), Cambridge, MA (<http://exac.broadinstitute.org>)), suggesting that they are probably not high-penetrance causal variants in cases of severe disorder, but might still represent risk factors, or produce a milder phenotype.

The significance of the polyglutamine tracts to the functions of the FOXP2 protein remains unknown. Polyglutamine tracts may be critical to the function of some proteins [50], and the absence of equivalent tracts in human FOXP1 and FOXP4 (despite presence of glutamine-rich regions) raises the possibility that tract expansion in FOXP2 may have contributed to divergence in FOXP protein function. To investigate the roles of the polyglutamine tracts in FOXP2, we engineered versions of the protein with reduced tract lengths reflecting the presumed ancestral protein sequence. The long tract in FOXP2 underwent expansion early in tetrapod evolution and has subsequently remained relatively stable: ray-finned fish species have a cluster of only 3 glutamine residues in place of the long tract in human FOXP2 [51, 52], whereas the coelacanth, an evolutionary intermediate between ray-finned fish and tetrapods, has a 25-residue tract,

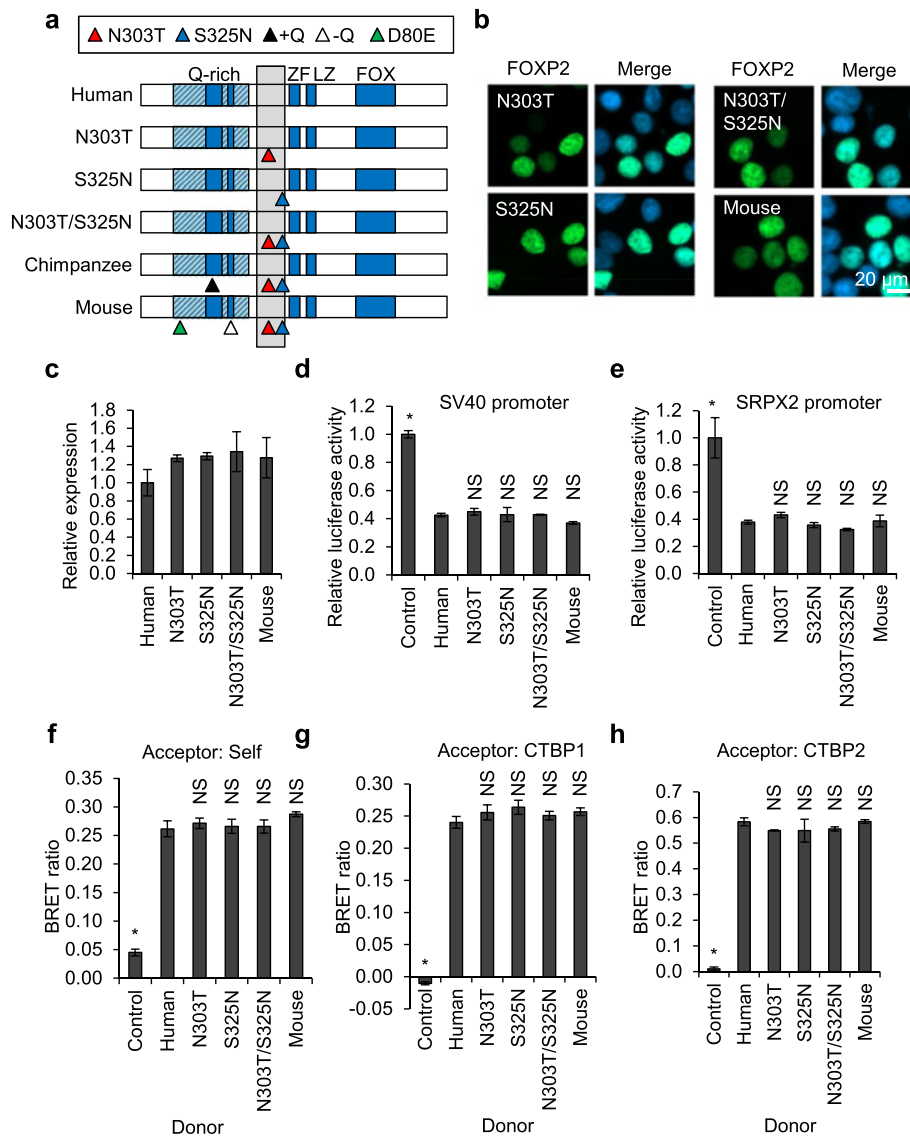


Fig. 5 (See legend on next page.)

(See figure on previous page.)

Fig. 5 Characterization of FOXP2 variants with ancestral amino acid substitutions. **a** Schematic representation of FOXP2 variants with ancestral amino acid substitutions. The N303T, S325N and N303T/S325N constructs are synthetic variants of human FOXP2 carrying ancestral amino acid substitutions. The chimpanzee and mouse orthologues of FOXP2 are included for comparison. Amino acid differences relative to human FOXP2 are indicated by arrowheads: N303T (red), S325N (blue), D80E (green), glutamine insertion (+Q, black), and glutamine deletion (–Q, white). Known domains are labelled: glutamine-rich (Q-rich) region (hatched shading) including polyglutamine tracts (solid shading), zinc finger (ZF), leucine zipper (LZ), and forkhead domain (FOX). The minimal CTBP interaction region determined using the BRET assay is indicated by the grey-shaded box. **b** Fluorescence micrographs of HEK293 cells transfected with FOXP2 variants with ancestral amino acid substitutions. Nuclei were stained with Hoechst 33342. **c** Fluorescence-based measurement of the expression level of FOXP2 variants with ancestral amino acid substitutions. HEK293 cells were transfected with YFP-FOXP2, together with mCherry for normalization. Fluorescence intensity was measured 48 h post-transfection. Values are mean YFP/mCherry fluorescence ratios \pm S.D. ($n = 3$), relative to the value for human FOXP2. **d, e** Luciferase reporter assays for transcriptional regulatory activity of FOXP2 variants with ancestral amino acid substitutions. HEK293 cells were transfected with a luciferase reporter vector containing the SV40 promoter (**d**) or the human *SRPX2* promoter (**e**), together with a *Renilla* luciferase normalization plasmid, and YFP-FOXP2 or YFP alone (control). Values are mean relative luciferase activity \pm S.D. ($n = 3$), expressed relative to the control. Asterisks indicate significant differences compared to human FOXP2 ($*p < 0.05$, one-way ANOVA followed by Bonferroni post hoc correction). NS not significant. Exact p values for **d** are <0.0001 for the control, >0.9999 for the N303T, S325N, and N303T/S325N variants and 0.2730 for the mouse protein. Exact p values for **e** are <0.0001 for the control and >0.9999 for the N303T, S325N, N303T/S325N and mouse variants. **f–h** BRET assays for protein-protein interactions of FOXP2 variants with ancestral amino acid substitutions. HEK293 cells were transfected with FOXP2 variants with ancestral amino acid substitutions fused to *Renilla* luciferase (donor), together with YFP (acceptor) fusions of the same FOXP2 variants (**f**), CTBP1 (**g**), or CTBP2 (**h**). The control donor protein is a nuclear-targeted luciferase and the control acceptor protein is a nuclear-targeted YFP. Values are mean corrected BRET ratios \pm S.D. ($n = 3$). Asterisks indicate significant differences compared to wild-type (WT) FOXP2 ($p < 0.05$, one-way ANOVA followed by Bonferroni post hoc correction). NS not significant. Exact p values for **f** are <0.0001 for the control, >0.9999 for the N303T, S325N, and N303T/S325N variants and 0.434 for the mouse protein. Exact p values for **g** are <0.0001 for the control and >0.9999 for the N303T, S325N, N303T/S325N and mouse variants. Exact p values for **h** are <0.0001 for the control, >0.9999 for the N303T, S325N, and mouse variants and 0.10 for the N303T/S325N variant

and tetrapod species typically have tracts of 35–41 residues (Additional file 5). We therefore reduced the length of the long tract in human FOXP2 to 3 residues. Similarly, the shorter tract was also reduced to 3 residues.

The two tract length reductions were engineered both individually and together (Fig. 6a). The resulting synthetic FOXP2 variants were assayed for a range of molecular properties including subcellular localization (Fig. 6b), expression level (Fig. 6c), transcriptional repression of the SV40 and *SRPX2* promoters (Fig. 6d,e), homodimerization (Fig. 6f), and interaction with CTBPs (Fig. 6g,h). Across these assays, there were no large differences between wild-type FOXP2 and versions of the protein with reduced polyglutamine tracts. Versions of FOXP2 lacking either the large tract or both tracts showed a small but statistically significant reduction in repressive ability in the SV40 luciferase reporter assay and also a small but statistically significant increase in homodimerization and CTBP1 interaction (Fig. 6e,f,g). However, these effects were not consistently observed across independent experiments and therefore may not represent genuine biological differences. The lack of any substantial differences between the wild-type protein and variants with reduced polyglutamine tracts in our assays suggests that the presence of the tracts may not have a marked impact on fundamental aspects of FOXP2 biology.

Discussion

Our detailed functional characterization of rare FOXP2 variants reported in individuals with neurodevelopmental disorders confirms the etiological role of the Q390Vfs*7 variant, which has similarly deleterious effects on protein

function to the two previously characterized pathogenic variants, p.R553H and p.R328*, in our assays. The data presented here point to potential diversity in molecular mechanisms in FOXP2-related speech/language disorder. Variants which abolish both transcriptional regulatory activity and protein dimerization, such as the p.R328* and p.Q390Vfs*7 variants, may act as null alleles, resulting in haploinsufficiency of FOXP2. In contrast, variants which show a loss of transcriptional regulatory activity but retain the ability to dimerize, like the p.R553H variant, may additionally interfere with the functioning of wild-type protein, as has been suggested for comparable variants in FOXP1 [31]. Variants producing these dominant-negative effects might result in a more severe phenotype. However, it will be necessary to identify further cases of FOXP2-related speech/language disorder resulting from missense variants in order to assess differences in phenotypic outcome.

In our assays, the p.Q17L, p.M406T, p.P416T, or p.N597H variants do not display any substantial effects on the core functions of the FOXP2 protein, suggesting that they may be incidental variants. Notably, the p.N597H variant was previously described as a likely causal variant in a child with CAS [7]. While we cannot rule out that the p.N597H variant disrupts an aspect of FOXP2 function not tested here, subtle effects on protein function are unlikely to lead to a speech/language disorder of comparable severity to complete loss-of-function variants. This case illustrates that, given the high level of rare coding variation in the general population (The 1000 Genomes Project Consortium, 2015), novel rare variants in known disorder-related genes cannot automatically be regarded

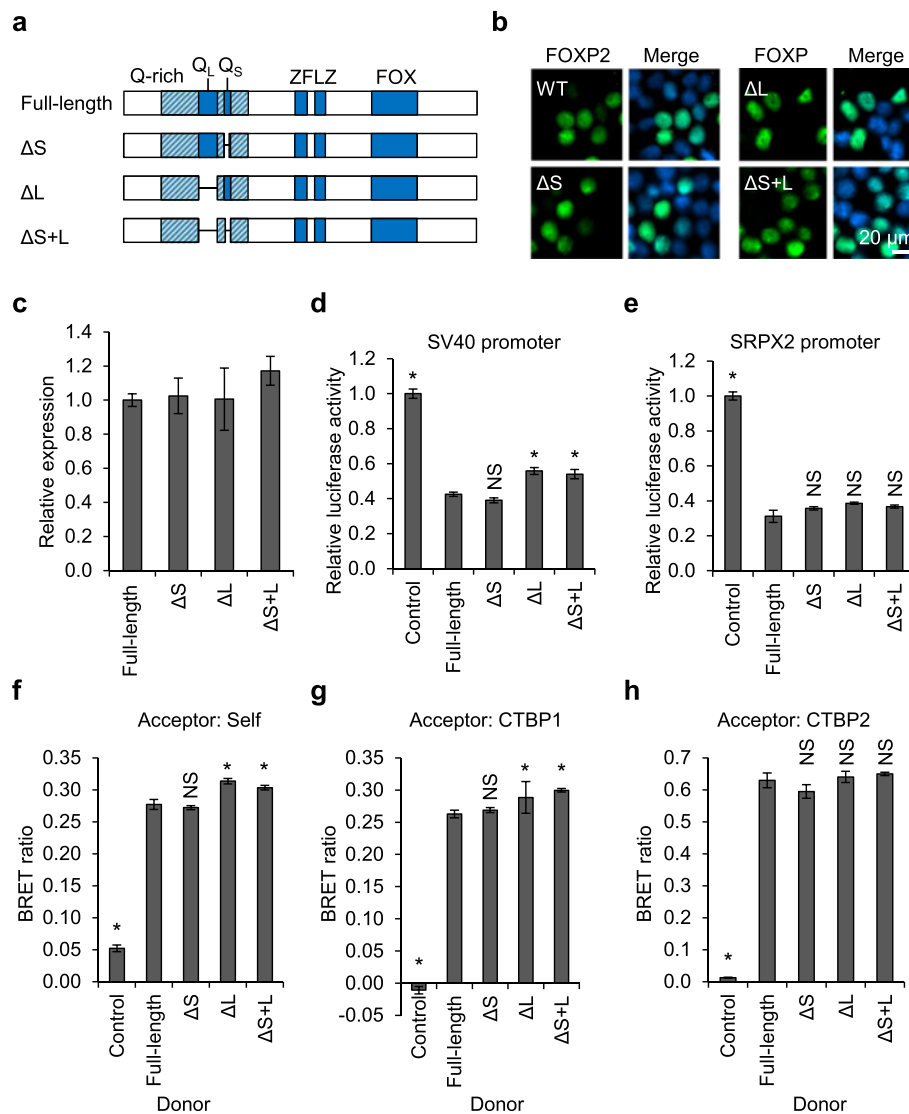


Fig. 6 Characterization of synthetic FOXP2 variants with reduced polyglutamine tracts. **a** Schematic representation of synthetic FOXP2 variants with reduced polyglutamine tracts. Known domains are labelled: glutamine-rich (*Q-rich*) region (hatched shading) including the long polyglutamine tract (*Q_L*) and short polyglutamine tract (*Q_S*), zinc finger (*ZF*), leucine zipper (*LZ*), and forkhead domain (*FOX*). The ΔS variant has a short polyglutamine tract reduced from 10 to 3 residues. The ΔL variant has a long polyglutamine tract reduced from 40 to 3 residues. The $\Delta S+L$ variant has shortened versions of both tracts. **b** Fluorescence micrographs of HEK293 cells transfected with FOXP2 variants with reduced polyglutamine tracts. Nuclei were stained with Hoechst 33342. **c** Fluorescence-based measurement of expression level for FOXP2 variants with reduced polyglutamine tracts. HEK293 cells were transfected with YFP-FOXP2 variants together with mCherry for normalization. Fluorescence intensity was measured 48 h post-transfection. Values are mean YFP/mCherry fluorescence ratios \pm S.D. ($n = 3$), expressed relative to the value for full-length FOXP2. **d, e** Luciferase reporter assays for transcriptional regulatory activity of FOXP2 variants with reduced polyglutamine tracts. Cells were transfected with a luciferase reporter vector containing the SV40 promoter (**d**) or the human *SRPX2* promoter (**e**), together with a *Renilla* luciferase normalization plasmid, and YFP-FOXP2 or YFP alone (control). Values are mean relative luciferase activity \pm S.D. ($n = 3$), expressed relative to the control. Asterisks indicate significant differences compared to full-length FOXP2 ($p < 0.05$, one-way ANOVA followed by Bonferroni post hoc correction). NS not significant. Exact p values for **d** are <0.0001 for the control, 0.4892 for ΔS , 0.0003 for ΔL , and 0.0010 for $\Delta S+L$. Exact p values for **e** are <0.0001 for the control, 0.0768 for ΔS , 0.0633 for ΔL , and 0.1506 for $\Delta S+L$. **f-h** BRET assays for protein-protein interactions of FOXP2 variants with reduced polyglutamine tracts. HEK293 cells were transfected with FOXP2 variants with reduced polyglutamine tracts fused to *Renilla* luciferase (donor) and YFP (acceptor) fusions of the same variants (**f**), CTBP1 (**g**) or CTBP2 (**h**). The control donor protein is a nuclear-targeted luciferase and the control acceptor protein is a nuclear-targeted YFP. Values are mean corrected BRET ratios \pm S.D. ($n = 3$). Asterisks indicate significant differences compared to wild-type (WT) FOXP2 ($p < 0.05$, one-way ANOVA followed by Bonferroni post hoc correction). NS not significant. Exact p values for **f** are <0.0001 for the control, ΔL and $\Delta S+L$, and >0.9999 for ΔS . Exact p values for **g** are <0.0001 for the control, >0.9999 for ΔS , 0.038 for ΔL , and 0.024 for $\Delta S+L$. Exact p values for **h** are <0.0001 for the control, 0.4 for ΔS , and >0.9999 for ΔL and $\Delta S+L$.

as causal. This is particularly true if it is not possible to establish segregation of the variant with disorder (or de novo occurrence of a variant in sporadic disorder) or if the disorder is genetically heterogeneous and many genetic risk factors remain unknown, as is the situation for CAS. Functional characterization of novel rare *FOXP2* variants is therefore essential to provide a concrete diagnosis of *FOXP2*-related disorder and to shed light on the aspects of protein function that are disrupted. Our experiments provide a framework for the characterization of novel *FOXP2* variants uncovered through future genetic analyses of individuals with speech/language disorder.

Concurrent with the present study, two further cases of probable *FOXP2*-related speech/language disorder were added to the DECIPHER database (patient IDs 271859 and 271246) [53]. One DECIPHER case presented with delayed speech and language development, dysarthria, and pulmonary stenosis and carries a de novo p.R553H variant identical to that observed in the KE family. The probability of recurrence for this particular variant is elevated because it involves mutation of a CpG site. Indeed, pathological mutation events have been observed at the equivalent CpG site in the FOX proteins *FOXC2*, *FOX E1*, *FOXF1*, *FOXL2*, and *FOXP1* [31, 54–57]. The second probable new case of *FOXP2*-related speech/language disorder in DECIPHER presented with delayed speech and language development, delayed fine motor development, strabismus and tall stature and carries a novel de novo stop-gain variant, p.R564*, truncating the *FOXP2* protein within the FOX domain. These two new DECIPHER cases underline the characteristic manifestation of heterozygous *FOXP2* disruption as a motor speech disorder. In addition, fine motor delays have now been reported in connection with two *FOXP2* variants (p.R564* and p.Q390Vfs*7), suggesting that the effects of *FOXP2* disruption on motor coordination may in some cases extend beyond the orofacial movements required for speech, consistent with the broader motor skill learning deficits observed in mice with heterozygous *FOXP2* disruption [1, 9].

Very recently, while the current manuscript was under review, eight further families with potential *FOXP2*-related speech/language disorder have been newly described, based on sequencing data [58]. Although functional characterizations were not carried out in that study, the new cases suggest that there is some variation in the nature and severity of the speech and language deficits resulting from *FOXP2* disruption and that there may also be variable cognitive impairments and behavioural anomalies in affected individuals. Of particular note, the newly described cohort included two novel missense variants in the FOX domain of *FOXP2*, but unlike the p.R553H variant found in the KE family, these missense variants do not occur in the DNA-binding helix of the domain. Future functional analyses of these newly identified variants would therefore be valuable

to reveal how they might interfere with protein function and to confirm that these variants play a causal role in the disorder. The description of these eight new families represents a step towards a ‘genotype-first’ approach to the characterization of the *FOXP2*-related phenotype [59]. However, sequencing of *FOXP2* in much larger cohorts of language-impaired individuals, preferably coupled to functional analyses, will be necessary to obtain a complete picture of the clinical spectrum associated with *FOXP2*-related disorder and to gain further insights into the role of *FOXP2* in neurodevelopment.

In addition to assessing the effects of rare variants on *FOXP2* protein function, we performed the first detailed characterization of the interactions between *FOXP2* and the transcriptional co-repressors *CTBP1* and *CTBP2*, finding that these interactions are independent of *FOXP2* homodimerization and DNA binding. Our results point to an extended *CTBP* binding site in *FOXP2* within residues 258–487, possibly encompassing multiple subsites. Previous yeast two-hybrid screens identified residues 122–382 in *FOXP2* as sufficient for interaction with *CTBPs* [34]. We propose a narrower critical region of 71 residues, from residue 259 to 329. This minimal region contains both of the amino acid substitutions that have occurred since the divergence of the human and chimpanzee lineages. However, we did not find any effect of these substitutions on *CTBP* interaction, and the human and mouse versions of *FOXP2* exhibit comparable levels of interaction with *CTBPs*. Changes in *CTBP* interaction do not therefore appear to underlie the neurobiological changes observed in mice carrying a partially humanized version of *Foxp2*. The amino acid changes on the human lineage may instead result in altered interaction with as yet unidentified protein interaction partners or in changes to post-translational modifications. In addition, we confirmed that the *FOXP2* paralog *FOXP1* also interacts with *CTBPs*, whereas *FOXP4* does not, pointing to a divergence in mechanisms of transcriptional regulation within the *FOXP* subfamily.

The *CTBPs* are conserved vertebrate proteins that interact with several different transcription factors and mediate transcriptional repression primarily through recruitment of the histone deacetylases *HDAC1/2* [36]. The *CTBPs* are widely expressed during embryonic development and have essential and partially overlapping roles in the development of multiple organs [35, 36]. They are expressed in the brain from a very early stage, although their specific roles in neurodevelopment have not been extensively investigated [35, 36]. *CTBPs* may therefore be key players in *FOXP2*- and *FOXP1*-mediated transcriptional repression across multiple organs and developmental stages, including in neurodevelopment. In addition, *CTBPs* may play a role in the post-translational modification of *FOXP2*. We and other groups have recently

demonstrated that FOXP2 is post-translationally modified by SUMOylation [21, 60, 61]. The essential SUMOylation enzyme UBC9 is a core component of the protein complex containing CTBP1, and CTBP1 has been suggested to function as a platform for protein SUMOylation [62]. Furthermore, the SUMO E3 ligase PIAS1, which promotes FOXP2 SUMOylation, is an auxiliary component of the CTBP1 complex [21, 62]. If CTBPs are among the key mediators of FOXP-mediated transcriptional repression, variants in FOXP2 or FOXP1 which disrupt the interaction with CTBPs might have a similar impact on neurodevelopment to variants which disrupt DNA binding, and an assessment of the effects on CTBP binding should therefore be included when characterizing novel, putatively pathogenic variants in FOXP proteins.

Finally, we reported the first detailed analysis of the role of the polyglutamine tracts in FOXP2 function. In our experiments, shortening the polyglutamine tracts in FOXP2 did not have a substantial impact on protein expression, nuclear localization, homodimerization, interaction with CTBPs, or transcriptional repression activity. The tracts do not therefore appear to be essential for core aspects of protein function in cellular assays. However, the tracts may have critical roles *in vivo* that are not apparent in cell models, and the generation of mice carrying a version of *Foxp2* lacking the polyglutamine tracts could therefore be informative. Deletion of the polyglutamine tracts in a small number of other proteins has been found to affect behaviour in mouse models. For example, deletion of the polyglutamine tract in mouse huntingtin results in subtle behavioural changes relating to learning and memory and motor coordination [63]. Replacement of mouse *POU3F2* with a *Xenopus* orthologue that lacks polyglutamine tracts and other homopolymer amino acid repeats resulted in neurochemical changes and a dramatic deterioration in pup retrieval behaviour [64]. The polyglutamine tract in mouse *SRY* has important roles in protein stabilization and transcriptional activation that are essential for sex determination [65]. It is also possible that the polyglutamine tracts in FOXP2 do not have critical roles in normal protein function. The tracts lie within a glutamine-rich protein region that is conserved in FOXP1 and FOXP4 and may be prone to developing expanded stretches of glutamine residues. In FOXP1, a different cluster of glutamine residues has undergone expansion in the rodent lineage to produce tracts of 37 residues in the mouse and 39 in rat, compared to 7 residues in primates. Expansions of up to ~40 glutamine residues within the glutamine-rich region of FOXP proteins may therefore have negligible effect on protein function.

The failure to observe any substantial effects of polyglutamine tract reduction on core aspects of FOXP2 function makes it less likely that tract length reductions contribute to risk for neurodevelopmental disorder.

However, one recent study suggested that a FOXP2 polyglutamine tract length variant tentatively associated with speech sound disorder (p.Q172del) exhibits altered transcriptional repression activity in a cell model [18, 66]. The p.Q172del variant was reported to increase transcript levels of the FOXP2 target gene *CNTNAP2* in transfected cells, whereas transcript levels were reduced in cells transfected with wild-type protein [66]. Nonetheless, both the variant and wild-type forms of FOXP2 produced increases in protein levels of *CNTNAP2*, with the variant having a stronger effect [66]. It is therefore unclear what effect the p.Q172del variant might have on transcriptional regulation *in vivo*, but our experiments suggest that the effect of a single glutamine deletion is likely to be extremely small. The p.Q172del variant is observed in the general population with a minor allele frequency of 0.6 % in the ExAC database (714 of 115484 chromosomes), and larger tract length reductions are also present in the ExAC database at very low frequency. Studies of larger numbers of individuals with FOXP2 proteins with standard-length and reduced-length tracts are therefore necessary to determine if tract length reduction might confer an increased risk for neurodevelopmental disorder or a reduction in linguistic abilities.

Conclusions

By performing detailed functional characterization of rare variants in FOXP2 found in individuals with neurodevelopmental disorders, we confirm the causal role of one recently reported uncharacterized variant and provide further characterization of two further causal variants. In addition, we highlight two variants which had been suggested to contribute to the neurodevelopmental disorders in the affected individuals, but which do not have detrimental effects on protein function, suggesting that they are in fact incidental. These findings underline the importance of performing functional characterization of novel variants identified in individuals with neurodevelopmental disorders, even when the affected gene has previously been implicated in disorder. The provision of accurate molecular diagnoses in cases of neurodevelopmental disorder will require consideration of protein function data together with detailed clinical observations and genetic characterization of affected individuals and their parents on a genome-wide scale.

Our research also provides important new insights into the biology of the FOXP2 protein, expanding our knowledge of the molecular functions of this key neural transcription factor. Our characterization of the interaction of FOXP2 with co-repressor proteins of the CTBP family provides additional mechanistic insight into FOXP2-mediated repression of transcription. We also report the first detailed examination of the role of the polyglutamine tracts in FOXP2. We find that these tracts are not essential to the core molecular functions of the

FOXP2 protein, suggesting that variations in tract length are unlikely to be a highly penetrant cause of neurodevelopmental disorder.

Additional files

Additional file 1: DNA sequences of primers used for molecular cloning. (PDF 239 kb)

Additional file 2: DNA sequences of primers used for site-directed mutagenesis. (PDF 234 kb)

Additional file 3: Western blots of YFP-tagged FOXP2 variants. (PDF 521 kb)

Additional file 4: Summary of studies examining FOXP2 polyglutamine tract length variation in individuals with neurodevelopmental and neuropsychiatric disorders. (PDF 251 kb)

Additional file 5: FOXP2 polyglutamine tract lengths in selected vertebrates. (PDF 325 kb)

Abbreviations

BRET: Bioluminescence resonance energy transfer; CAS: Childhood apraxia of speech; CTBP: C-terminal binding protein; FOX: Forkhead Box; MRI: Magnetic resonance imaging; YFP: Yellow fluorescent protein

Acknowledgements

We wish to thank Elliot Sollis for help generating DNA constructs.

Funding

This work was supported by the Max Planck Society.

Availability of data and materials

The data supporting the conclusions of this article are included within the article and its additional files.

Authors' contributions

SEF and SAG designed and supervised the study. SBE, SAG, SMC and PD performed the experiments. SAG, SBE and SEF wrote the manuscript. All authors were involved in data analysis and interpretation, and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Received: 29 June 2016 Accepted: 8 November 2016

Published online: 28 November 2016

References

- Groszer M, Keays DA, Deacon RM, de Bono JP, Prasad-Mulcare S, Gaub S, Baum MG, French CA, Nicod J, Coventry JA, Enard W, Fray M, Brown SD, Nolan PM, Paabo S, Channon KM, Costa RM, Eilers J, Ehret G, Rawlins JN, Fisher SE. Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits. *Curr Biol*. 2008;18:354–62. doi:10.1016/j.cub.2008.01.060.
- Shu W, Cho JY, Jiang Y, Zhang M, Weisz D, Elder GA, Schmeidler J, De Gasperi R, Sosa MA, Rabidou D, Santucci AC, Perl D, Morrisey E, Buxbaum JD. Altered ultrasonic vocalization in mice with a disruption in the *Foxp2* gene. *Proc Natl Acad Sci U S A*. 2005;102:9643–8. doi:10.1073/pnas.0503739102.
- Graham SA, Fisher SE. Understanding language from a genomic perspective. *Annu Rev Genet*. 2015;49:131–60. doi:10.1146/annurev-genet-120213-092236.
- Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*. 2001;413:519–23. doi:10.1038/35097076.
- Stroud JC, Wu Y, Bates DL, Han A, Nowick K, Paabo S, Tong H, Chen L. Structure of the forkhead domain of FOXP2 bound to DNA. *Structure*. 2006;14:159–66. doi:10.1016/j.str.2005.10.005.
- Vernes SC, Nicod J, Elahi FM, Coventry JA, Kenny N, Coupe AM, Bird LE, Davies KE, Fisher SE. Functional genetic analysis of mutations implicated in a human speech and language disorder. *Hum Mol Genet*. 2006;15:3154–67. doi:10.1093/hmg/ddl392.
- Laffin JJ, Raca G, Jackson CA, Strand EA, Jakielski KJ, Shriberg LD. Novel candidate genes and regions for childhood apraxia of speech identified by array comparative genomic hybridization. *Genet Med*. 2012;14:928–36. doi:10.1038/gim.2012.72.
- MacDermot KD, Bonora E, Sykes N, Coupe AM, Lai CS, Vernes SC, Vargha-Khadem F, McKenzie F, Smith RL, Monaco AP, Fisher SE. Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *Am J Hum Genet*. 2005;76:1074–80. doi:10.1086/430841.
- Turner SJ, Hildebrand MS, Block S, Damiano J, Fahey M, Reilly S, Bahlo M, Scheffer IE, Morgan AT. Small intragenic deletion in FOXP2 associated with childhood apraxia of speech and dysarthria. *Am J Med Genet A*. 2013;161A:2321–6. doi:10.1002/ajmg.a.36055.
- Gauthier J, Joobor R, Mottron L, Laurent S, Fuchs M, De Kimpe V, Rouleau GA. Mutation screening of FOXP2 in individuals diagnosed with autistic disorder. *Am J Med Genet A*. 2003;118A:172–5. doi:10.1002/ajmg.a.10105.
- Laroche F, Ramoz N, Leroy S, Fortin C, Rousselot-Paillet B, Philippe A, Colleaux L, Bresson JL, Mogenet A, Golse B, Mouren-Simeoni MC, Gorwood P, Galli T, Simonneau M, Krebs MO, Robel L. Polymorphisms of coding trinucleotide repeats of homeobox genes in neurodevelopmental psychiatric disorders. *Psychiatr Genet*. 2008;18:295–301. doi:10.1097/YPG.0b013e3283060fa5.
- Levchenko A, Davtian S, Petrova N, Malashichev Y. Sequencing of five left-right cerebral asymmetry genes in a cohort of schizophrenia and schizotypal disorder patients from Russia. *Psychiatr Genet*. 2014;24:75–80. doi:10.1097/YPG.000000000000021.
- Li H, Yamagata T, Mori M, Momoi MY. Absence of causative mutations and presence of autism-related allele in FOXP2 in Japanese autistic patients. *Brain Dev*. 2005;27:207–10. doi:10.1016/j.braindev.2004.06.002.
- Newbury DF, Bonora E, Lamb JA, Fisher SE, Lai CS, Baird G, Jannoun L, Slonims V, Stott CM, Merricks MJ, Bolton PF, Bailey AJ, Monaco AP, International Molecular Genetic Study of Autism C. FOXP2 is not a major susceptibility gene for autism or specific language impairment. *Am J Hum Genet*. 2002;70:1318–27. doi:10.1086/339931.
- Roll P, Vernes SC, Bruneau N, Cillario J, Ponsolle-Lenfant M, Massacrier A, Rudolf G, Khalife M, Hirsch E, Fisher SE, Szepietowski P. Molecular networks implicated in speech-related disorders: FOXP2 regulates the SRPX2/uPAR complex. *Hum Mol Genet*. 2010;19:4848–60. doi:10.1093/hmg/ddq415.
- Tolosa A, Sanjuan J, Dagnall AM, Molto MD, Herrero N, de Frutos R. FOXP2 gene and language impairment in schizophrenia: association and epigenetic studies. *BMC Med Genet*. 2010;11:114. doi:10.1186/1471-2350-11-114.
- Wassink TH, Piven J, Vieland VJ, Pietila J, Goedken RJ, Folstein SE, Sheffield VC. Evaluation of FOXP2 as an autism susceptibility gene. *Am J Med Genet*. 2002;114:566–9. doi:10.1002/ajmg.10415.
- Zhao Y, Ma H, Wang Y, Gao H, Xi C, Hua T, Zhao Y, Qiu G. Association between FOXP2 gene and speech sound disorder in Chinese population. *Psychiatry Clin Neurosci*. 2010;64:565–73. doi:10.1111/j.1440-1819.2010.02123.x.
- Deriziotis P, Graham SA, Estruch SB, Fisher SE. Investigating protein-protein interactions in live cells using bioluminescence resonance energy transfer. *J Vis Exp*. 2014. doi:10.3791/51438.
- Deriziotis P, O'Roak BJ, Graham SA, Estruch SB, Dimitropoulou D, Bernier RA, Gerdts J, Shendure J, Eichler EE, Fisher SE. De novo TBR1 mutations in sporadic autism disrupt protein functions. *Nat Commun*. 2014;5:4954. doi:10.1038/ncomms5954.
- Estruch SB, Graham SA, Deriziotis P, Fisher SE. The language-related transcription factor FOXP2 is post-translationally modified with small ubiquitin-like modifiers. *Sci Rep*. 2016;6:20911. doi:10.1038/srep20911.
- Mizutani A, Matsuzaki A, Momoi MY, Fujita E, Tanabe Y, Momoi T. Intracellular distribution of a speech/language disorder associated FOXP2 mutant. *Biochem Biophys Res Commun*. 2007;353:869–74. doi:10.1016/j.bbrc.2006.12.130.
- Wang B, Lin D, Li C, Tucker P. Multiple domains define the expression and regulatory properties of Foxp1 forkhead transcriptional repressors. *J Biol Chem*. 2003;278:24259–68. doi:10.1074/jbc.M207174200.
- Li S, Weidenfeld J, Morrisey EE. Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions. *Mol Cell Biol*. 2004;24:809–22.

25. Chatila TA, Blaeser F, Ho N, Lederman HM, Voulgaropoulos C, Helms C, Bowcock AM. JM2, encoding a fork head-related protein, is mutated in X-linked autoimmunity-allergic dysregulation syndrome. *J Clin Invest*. 2000;106:R75–81. doi:10.1172/JCI11679.
26. d'Hennezel E, Bin Dhuban K, Torgerson T, Piccirillo CA. The immunogenetics of immune dysregulation, polyendocrinopathy, enteropathy, X linked (IPEX) syndrome. *J Med Genet*. 2012;49:291–302. doi:10.1136/jmedgenet-2012-100759.
27. Ferland RJ, Cherry TJ, Preware PO, Morrissy EE, Walsh CA. Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain. *J Comp Neurol*. 2003;460:266–79. doi:10.1002/cne.10654.
28. Li X, Wang W, Wang J, Malovannaya A, Xi Y, Li W, Guerra R, Hawke DH, Qin J, Chen J. Proteomic analyses reveal distinct chromatin-associated and soluble transcription factor complexes. *Mol Syst Biol*. 2015;11:775. doi:10.15252/msb.20145504.
29. Takahashi K, Liu FC, Hirokawa K, Takahashi H. Expression of Foxp4 in the developing and adult rat forebrain. *J Neurosci Res*. 2008;86:3106–16. doi:10.1002/jnr.21770.
30. Bacon C, Rappold GA. The distinct and overlapping phenotypic spectra of FOXP1 and FOXP2 in cognitive disorders. *Hum Genet*. 2012;131:1687–98. doi:10.1007/s00439-012-1193-z.
31. Sollis E, Graham SA, Vino A, Froehlich H, Vreeburg M, Dimitropoulou D, Gilissen C, Pfundt R, Rappold GA, Brunner HG, Deriziotis P, Fisher SE. Identification and functional characterization of de novo FOXP1 variants provides novel insights into the etiology of neurodevelopmental disorder. *Hum Mol Genet*. 2015. doi:10.1093/hmg/ddv495.
32. Corominas R, Yang X, Lin GN, Kang S, Shen Y, Ghamsari L, Broly M, Rodriguez M, Tam S, Trigg SA, Fan C, Yi S, Tasan M, Lemmens I, Kuang X, Zhao N, Malhotra D, Michaelson JJ, Vacic V, Calderwood MA, Roth FP, Tavernier J, Horvath S, Salehi-Ashtiani K, Korkin D, Sebat J, Hill DE, Hao T, Vidal M, Lakoucheva LM. Protein interaction network of alternatively spliced isoforms from brain links genetic risk factors for autism. *Nat Commun*. 2014;5:3650. doi:10.1038/ncomms4650.
33. Rolland T, Tasan M, Charletoaux B, Pevzner SJ, Zhong Q, Sahni N, Yi S, Lemmens I, Fontanillo C, Mosca R, Kamburov A, Ghiassian SD, Yang X, Ghamsari L, Balcha D, Begg BE, Braun P, Brehme M, Broly MP, Carvunis AR, Convery-Zupan D, Corominas R, Coulombe-Huntington J, Dann E, Dreze M, Dricot A, Fan C, Franzosa E, Gebreab F, Gutierrez BJ, Hardy MF, Jin M, Kang S, Kirov R, Lin GN, Luck K, MacWilliams A, Menche J, Murray RR, Palagi A, Poulin MM, Rambout X, Rasla J, Reichert P, Romero V, Ruysinck E, Sahalie JM, Scholz A, Shah AA, Sharma A, Shen Y, Spirohn K, Tam S, Tejeda AO, Trigg SA, Twizere JC, Vega K, Walsh J, Cusick ME, Xia Y, Barabasi AL, lakoucheva LM, Aloy P, De Las RJ, Tavernier J, Calderwood MA, Hill DE, Hao T, Roth FP, Vidal M. A proteome-scale map of the human interactome network. *Cell*. 2014;159:1212–26. doi:10.1016/j.cell.2014.10.050.
34. Sakai Y, Shaw CA, Dawson BC, Dugas DV, Al-Mohtaseb Z, Hill DE, Zoghbi HY. Protein interactome reveals converging molecular pathways among autism disorders. *Sci Transl Med*. 2011;3:86ra49. doi:10.1126/scitranslmed.3002166.
35. Hildebrand JD, Soriano P. Overlapping and unique roles for C-terminal binding protein 1 (CtBP1) and CtBP2 during mouse development. *Mol Cell Biol*. 2002;22:5296–307.
36. Stankiewicz TR, Gray JJ, Winter AN, Linseman DA. C-terminal binding proteins: central players in development and disease. *Biomol Concepts*. 2014;5:489–511. doi:10.1515/bmc-2014-0027.
37. Bergman LM, Morris L, Darley M, Mirnezami AH, Gunatillake SC, Blydes JP. Role of the unique N-terminal domain of CtBP2 in determining the subcellular localisation of CtBP family proteins. *BMC Cell Biol*. 2006;7:35. doi:10.1186/1471-2121-7-35.
38. Verger A, Quinlan KG, Crofts LA, Spano S, Corda D, Kable EP, Braet F, Crossley M. Mechanisms directing the nuclear localization of the CtBP family proteins. *Mol Cell Biol*. 2006;26:4882–94. doi:10.1128/MCB.02402-05.
39. Enard W, Przeworski M, Fisher SE, Lai CS, Wiebe V, Kitano T, Monaco AP, Paabo S. Molecular evolution of FOXP2, a gene involved in speech and language. *Nature*. 2002;418:869–72. doi:10.1038/nature01025.
40. Li G, Wang J, Rossiter SJ, Jones G, Zhang S. Accelerated FoxP2 evolution in echolocating bats. *PLoS One*. 2007;2:e900. doi:10.1371/journal.pone.0000900.
41. Webb DM, Zhang J. FoxP2 in song-learning birds and vocal-learning mammals. *J Hered*. 2005;96:212–6. doi:10.1093/jhered/esi025.
42. Krause J, Lalueza-Fox C, Orlando L, Enard W, Green RE, Burbano HA, Hublin JJ, Hanni C, Fortea J, de la Rasilla M, Bertranpetit J, Rosas A, Paabo S. The derived FOXP2 variant of modern humans was shared with Neandertals. *Curr Biol*. 2007;17:1908–12. doi:10.1016/j.cub.2007.10.008.
43. Enard W, Gehre S, Hammerschmidt K, Holter SM, Blass T, Somel M, Bruckner MK, Schreiwies C, Winter C, Sohr R, Becker L, Wiebe V, Nickel B, Giger T, Muller U, Groszer M, Adler T, Aguilar A, Bolle I, Calzada-Wack J, Dalke C, Ehrhardt N, Favor J, Fuchs H, Gailus-Durner V, Hans W, Holzswimmer G, Javaheri A, Kalaydjiev S, Kallnik M, Kling E, Kunder S, Mossbrugger I, Naton B, Racz I, Rathkolb B, Rozman J, Schrewe A, Busch DH, Graw J, Ivandic B, Klingenspor M, Klopstock T, Ollert M, Quintanilla-Martinez L, Schulz H, Wolf E, Wurst W, Zimmer A, Fisher SE, Morgenstern R, Arendt T, de Angelis MH, Fischer J, Schwarz J, Paabo S. A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell*. 2009;137:961–71. doi:10.1016/j.cell.2009.03.041.
44. Reimers-Kipping S, Hevers W, Paabo S, Enard W. Humanized Foxp2 specifically affects cortico-basal ganglia circuits. *Neuroscience*. 2011;175:75–84. doi:10.1016/j.neuroscience.2010.11.042.
45. Schreiwies C, Bornschein U, Burguiere E, Kerimoglu C, Schreiter S, Dannemann M, Goyal S, Rea E, French CA, Puliyadi R, Groszer M, Fisher SE, Mundry R, Winter C, Hevers W, Paabo S, Enard W, Graybiel AM. Humanized Foxp2 accelerates learning by enhancing transitions from declarative to procedural performance. *Proc Natl Acad Sci U S A*. 2014;111:14253–8. doi:10.1073/pnas.1414542111.
46. Enard W. FOXP2 and the role of cortico-basal ganglia circuits in speech and language evolution. *Curr Opin Neurobiol*. 2011;21:415–24. doi:10.1016/j.conb.2011.04.008.
47. Butland SL, Devon RS, Huang Y, Mead CL, Meynert AM, Neal SJ, Lee SS, Wilkinson A, Yang GS, Yuen MM, Hayden MR, Holt RA, Leavitt BR, Ouellette BF. CAG-encoded polyglutamine length polymorphism in the human genome. *BMC Genomics*. 2007;8:126. doi:10.1186/1471-2164-8-126.
48. Riley BE, Orr HT. Polyglutamine neurodegenerative diseases and regulation of transcription: assembling the puzzle. *Genes Dev*. 2006;20:2183–92. doi:10.1101/gad.1436506.
49. Bruce HA, Margolis RL. FOXP2: novel exons, splice variants, and CAG repeat length stability. *Hum Genet*. 2002;111:136–44. doi:10.1007/s00439-002-0768-5.
50. Faux N. Single amino acid and trinucleotide repeats: function and evolution. *Adv Exp Med Biol*. 2012;769:26–40.
51. Bonkowski JL, Chien CB. Molecular cloning and developmental expression of foxP2 in zebrafish. *Dev Dyn*. 2005;234:740–6. doi:10.1002/dvdy.20504.
52. Itakura T, Chandra A, Yang Z, Xue X, Wang B, Kimura W, Hikosaka K, Inohaya K, Kudo A, Uezato T, Miura N. The medaka FoxP2, a homologue of human language gene FOXP2, has a diverged structure and function. *J Biochem*. 2008;143:407–16. doi:10.1093/jb/mvm235.
53. Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, Van Vooren S, Moreau Y, Pettett RM, Carter NP. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet*. 2009;84:524–33. doi:10.1016/j.ajhg.2009.03.010.
54. Baris I, Arisoy AE, Smith A, Agostini M, Mitchell CS, Park SM, Halefoglou AM, Zengin E, Chatterjee VK, Battaloglu E. A novel missense mutation in human TTF-2 (FKHL15) gene associated with congenital hypothyroidism but not athyrosis. *J Clin Endocrinol Metab*. 2006;91:4183–7. doi:10.1210/jc.2006-0405.
55. Beyens D, Moumne L, Veitia R, Peters H, Leroy BP, De Paeppe A, De Baere E. Missense mutations in the forkhead domain of FOXL2 lead to subcellular mislocalization, protein aggregation and impaired transactivation. *Hum Mol Genet*. 2008;17:2030–8. doi:10.1093/hmg/ddn100.
56. Brice G, Mansour S, Bell R, Collin JR, Child AH, Brady AF, Sarfarazi M, Burnand KG, Jeffery S, Mortimer P, Murday VA. Analysis of the phenotypic abnormalities in lymphoedema-distichiasis syndrome in 74 patients with FOXC2 mutations or linkage to 16q24. *J Med Genet*. 2002;39:478–83.
57. Sen P, Yang Y, Navarro C, Silva J, Szafranski P, Kolodziejka KE, Dharmadhikari AV, Mostafa H, Kozakewich H, Kearney D, Cahill JB, Whitt M, Bilic M, Margraf L, Charles A, Goldblatt J, Gibson K, Lantz PE, Garvin AJ, Petty J, Kiblawi Z, Zupan C, McConkie-Rosell A, McDonald MT, Peterson-Carmichael SL, Gaede JT, Shivanna B, Schady D, Friedlich PS, Hays SR, Palafoff IV, Siebers-Renelt U, Bohring A, Finn LS, Siebert JR, Galambos C, Nguyen L, Riley M, Chossaing N, Vigouroux A, Rocha G, Fernandes S, Brumbaugh J, Roberts K, Ho-Ming L, Lo IF, Lam S, Gerychova R, Jezova M, Valaskova I, Fellmann F, Afshar K, Giannoni E, Muhlethaler V, Liang J, Beckmann JS, Lioy J, Deshmukh H, Srinivasan L, Swarr DT, Sloman M, Shaw-Smith C, van Loon RL, Hagman C, Sznajder Y, Barrea C, Galant C, Detaille T, Wambach JA, Cole FS, Hamvas A, Prince LS, Diderich KE, Brooks AS, Verdij RM, Ravindranathan H, Sugo E, Mowat D, Baker ML, Langston C, Welty S, Stankiewicz P. Novel FOXF1 mutations in sporadic and familial cases of alveolar capillary dysplasia with misaligned pulmonary veins imply a role for its DNA binding domain. *Hum Mutat*. 2013;34:801–11. doi:10.1002/humu.22313.
58. Reuter MS, Riess A, Moog U, Briggs TA, Chandler KE, Rauch A, Stampfer M, Steindl K, Glaser D, Joset P, Study DDD, Krumbiegel M, Rabe H, Schulte-

- Mattler U, Bauer P, Beck-Wodl S, Kohlhase J, Reis A, Zweier C. FOXP2 variants in 14 individuals with developmental speech and language disorders broaden the mutational and clinical spectrum. *J Med Genet*. 2016. doi:10.1136/jmedgenet-2016-104094.
59. Stessman HA, Bernier R, Eichler EE. A genotype-first approach to defining the subtypes of a complex disease. *Cell*. 2014;156:872–7. doi:10.1016/j.cell.2014.02.002.
 60. Meredith LJ, Wang CM, Nascimento L, Liu R, Wang L, Yang WH. The key regulator for language and speech development, FOXP2, is a novel substrate for SUMOylation. *J Cell Biochem*. 2015. doi:10.1002/jcb.25288.
 61. Usui N, Co M, Harper M, Rieger MA, Dougherty JD, Konopka G. Sumoylation of FOXP2 regulates motor function and vocal communication through Purkinje cell development. *Biol Psychiatry*. 2016. doi:10.1016/j.biopsych.2016.02.008.
 62. Kuppuswamy M, Vijayalingam S, Zhao LJ, Zhou Y, Subramanian T, Ryerse J, Chinnadurai G. Role of the PLDLS-binding cleft region of CtBP1 in recruitment of core and auxiliary components of the corepressor complex. *Mol Cell Biol*. 2008;28:269–81. doi:10.1128/MCB.01077-07.
 63. Clabough EB, Zeitlin SO. Deletion of the triplet repeat encoding polyglutamine within the mouse Huntington's disease gene results in subtle behavioral/ motor phenotypes in vivo and elevated levels of ATP with cellular senescence in vitro. *Hum Mol Genet*. 2006;15:607–23. doi:10.1093/hmg/ddi477.
 64. Nasu M, Yada S, Igarashi A, Sutoo D, Akiyama K, Ito M, Yoshida N, Ueda S. Mammalian-specific sequences in pou3f2 contribute to maternal behavior. *Genome Biol Evol*. 2014;6:1145–56. doi:10.1093/gbe/evu072.
 65. Zhao L, Ng ET, Davidson TL, Longmuss E, Urschitz J, Elston M, Moisyadi S, Bowles J, Koopman P. Structure-function analysis of mouse *Sry* reveals dual essential roles of the C-terminal polyglutamine tract in sex determination. *Proc Natl Acad Sci U S A*. 2014;111:11768–73. doi:10.1073/pnas.1400666111.
 66. Zhao Y, Liu X, Sun H, Wang Y, Yang W, Ma H. Contactin-associated protein-like 2 expression in SHSY5Y cells is upregulated by a FOXP2 mutant with a shortened polyglutamine tract. *Mol Med Rep*. 2015;12:8162–8. doi:10.3892/mmr.2015.4483.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

