Human Molecular Genetics, 2006, Vol. 15, No. 21 ***3154–3167***
doi:10.1093/hmg/ddl392
Advance Access published on September 19, 2006

Functional genetic analysis of mutations implicated in a human speech and language disorder

Sonja C. Vernes[1,2,†], Jérôme Nicod[1,†], Fanny M. Elahi[1], Julie A. Coventry[1], Niamh Kenny[1], Anne-Marie Coupe[1], Louise E. Bird[1], Kay E. Davies[2] and Simon E. Fisher[1,*]

[1]Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford, OX3 7BN, UK and [2]Department of Physiology, Anatomy and Genetics, Le Gros Clark Building, South Parks Road, Oxford OX1 3QX, UK

Received July 13, 2006; Revised and Accepted September 11, 2006

Mutations in the *FOXP2* gene cause a severe communication disorder involving speech deficits (developmental verbal dyspraxia), accompanied by wide-ranging impairments in expressive and receptive language. The protein encoded by *FOXP2* belongs to a divergent subgroup of forkhead-box transcription factors, with a distinctive DNA-binding domain and motifs that mediate hetero- and homodimerization. Here we report the first direct functional genetic investigation of missense and nonsense mutations in FOXP2 using human cell-lines, including a well-established neuronal model system. We focused on three unusual FOXP2 coding variants, uniquely identified in cases of verbal dyspraxia, assessing expression, subcellular localization, DNA-binding and transactivation properties. Analysis of the R553H forkhead-box substitution, found in all affected members of a large three-generation family, indicated that it severely affects FOXP2 function, chiefly by disrupting nuclear localization and DNA-binding properties. The R328X truncation mutation, segregating with speech/language disorder in a second family, yields an unstable, predominantly cytoplasmic product that lacks transactivation capacity. A third coding variant (Q17L) observed in a single affected child did not have any detectable functional effect in the present study. In addition, we used the same systems to explore the properties of different isoforms of FOXP2, resulting from alternative splicing in human brain. Notably, one such isoform, FOXP2.10+, contains dimerization domains, but no DNA-binding domain, and displayed increased cytoplasmic localization, coupled with aggresome formation. We hypothesize that expression of alternative isoforms of FOXP2 may provide mechanisms for post-translational regulation of transcription factor function.

INTRODUCTION

Developmental disruptions of speech and language are highly heritable (1), but at present there is little known about the underlying neuromolecular mechanisms. Inheritance patterns are usually complex in families suffering from unexplained communication deficits, and it is likely that several—perhaps many—different genes are involved (2). We previously discovered that heterozygous point mutations in the *FOXP2* gene lead to a rare monogenic form of speech and language impairment (3,4), offering the first opportunities to study language-related neural pathways from a molecular perspective (5). Mutation of *FOXP2* impairs learning and production of the complex coordinated sequences of articulatory gestures that support speech (6,7); such difficulties are commonly referred to as Developmental Verbal Dyspraxia (or Childhood Apraxia of Speech); (OMIM 602081). These problems are accompanied by disruption of expressive and receptive linguistic skills, evident for both oral and written language (7). Deficits in aspects of general cognition—although found in some affected individuals (8)—do not appear to be central to the *FOXP2*-related disorder, and are more significant for the verbal domain (7). Moreover, people with disruption of *FOXP2* show abnormal activation of language-related regions

*To whom correspondence should be addressed. Tel: +44 1865 287 647; Fax: +44 287 533; Email: simon.fisher@well.ox.ac.uk
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

in the brain (including underactivation of Broca's area) when performing language-based tasks, even under 'covert' conditions, where responses are thought rather than spoken (9).

The protein encoded by *FOXP2* belongs to an important group of transcription factors defined by the presence of a characteristic 80–110 amino acid DNA-binding motif (known as the 'forkhead-box' or 'FOX' domain) (10). FOX proteins are critical regulators of gene expression that influence a diverse array of biological processes during development and in adulthood, in a wide range of eukaryotic species (11). Human *FOX* genes have been implicated in a number of different congenital disorders and disease states, and functional analyses of the relevant mutations have helped to shed light on aetiology (reviewed in 12). Consistent with other FOX proteins (11), FOXP2 itself is likely to perform multiple functions, depending on tissue and cellular context; for example, during embryonic development murine *Foxp2* is expressed in restricted parts of the lung, heart, intestines and central nervous system (CNS) (13,14). With regard to brain function, based on conserved CNS expression patterns observed in distantly related species from humans to fish (14–17), it has been hypothesized that orthologues of this gene may influence circuits involved in sensory-motor processing and motor-skill learning in a wide range of vertebrates (18,19).

FOXP2 is a member of a subfamily within the forkhead family of transcription factors, which also includes FOXP1, FOXP3 and FOXP4 (and their orthologues in other mammalian species) (3,13,20–22). FOXP is one of the most recently characterized of the forkhead subfamilies, and is notably divergent. The C-terminal end of the DNA-binding domain is truncated and shows significant structural differences from that found in other FOX proteins (23). In addition, although most FOX proteins are believed to function as monomers, the FOXP subgroup has been shown to form homo- and heterodimers (21,22). This interaction appears to be mediated by a highly conserved stretch of sequence, containing a zinc finger and a leucine zipper, which is common to all FOXP proteins. It has been suggested that FOXP dimerization via the leucine zipper is a necessary prerequisite for DNA-binding and transcriptional activity (22). However, recent structural studies indicate that an isolated FOXP forkhead domain is able to bind DNA in a specific manner, either as a monomer or a domain-swapped dimer (23). FOXP proteins are also characterized by an N-terminal region that is rich in glutamine residues, and some subfamily members—including FOXP2—contain polyglutamine tracts which may be important for protein–protein interactions (3). Crucially, the various unusual features of FOXP1–4 suggest that general findings that have been established for other FOX proteins may not be applicable to this recently discovered subfamily.

The primary aim of the present investigation was to directly assess the effects of aetiological mutations in *FOXP2* using functional genetic techniques, an approach that has previously been informative when studying disease-causing mutations in other forkhead genes (24–27). We focused on three different *FOXP2* coding changes that have been identified in people affected with verbal dyspraxia (3,4). The first to be described in the literature was a heterozygous missense mutation inherited by all of the fifteen affected members of a well-studied three generation family (known as KE), suffering from severe speech and language problems (3). The KE change yields an arginine-to-histidine substitution at a highly conserved residue within the DNA-binding domain (R553H). More recently, a novel heterozygous point mutation was identified in three individuals from an unrelated family segregating developmental verbal dyspraxia (4). This second change, found in two affected siblings and their mother (who has a history of speech problems), yields a stop codon at position 328 of the FOXP2 protein (R328X), leading to a dramatically truncated product lacking several critical domains, such as the zinc finger, leucine zipper and DNA-binding motif. The third coding variant investigated in the present study involves a substitution near to the N-terminus of the protein (Q17L) in a region of undetermined function. This coding change was previously found in a proband diagnosed with verbal dyspraxia, and was absent from a large number of control chromosomes, suggesting that it is unlikely to be a natural polymorphism (4). However, the proband has an affected sibling who does not carry the substitution (4), and thus functional data are important for clarifying a potential role for this change in susceptibility to speech problems.

The functional effects of coding variants of human *FOXP2* have not yet been addressed in any model system. Although the impact of some similar mutations have been examined in other FOX proteins (such as the R127H substitution of FOXC1 (26), which is analogous to R553H in FOXP2), it is essential to directly investigate the consequences for FOXP2 itself given the unusual nature of the FOXP subfamily, noted earlier. Here we describe the first investigations of the expression, subcellular localization, DNA-binding and transactivation properties of wild-type and mutant forms of FOXP2 using a range of human cell-lines, including SH-SY5Y cells—a common *in vitro* model for studying neuronal function. In addition, we explore the characteristics of different isoforms of the protein, thought to result from alternative splicing in human brain, and uncover a potential biological role for the truncated splice variant known as FOXP2.10+.

RESULTS

Protein expression of variants of human FOXP2

The functional consequences of three rare coding variants of FOXP2, found exclusively in patients affected with verbal dyspraxia (3,4), were evaluated in tissue culture using transfected human cell-lines. The full-length coding region of the major transcript of *FOXP2* (known as isoform I) (3) was isolated from a human foetal brain library and introduced into the pcDNA4/HisMax expression vector (Invitrogen), in-frame with an N-terminal vector-encoded Xpress™ tag. Site-directed mutagenesis was used to generate point mutations within *FOXP2*, yielding separate constructs encoding the R553H, Q17L and R328X changes found in cases of speech and language disorder (Fig. 1A).

Human cell-lines were transfected with wild-type (isoform I) and mutant constructs, and whole-cell extracts were resolved by SDS-PAGE and western blotting. Detection

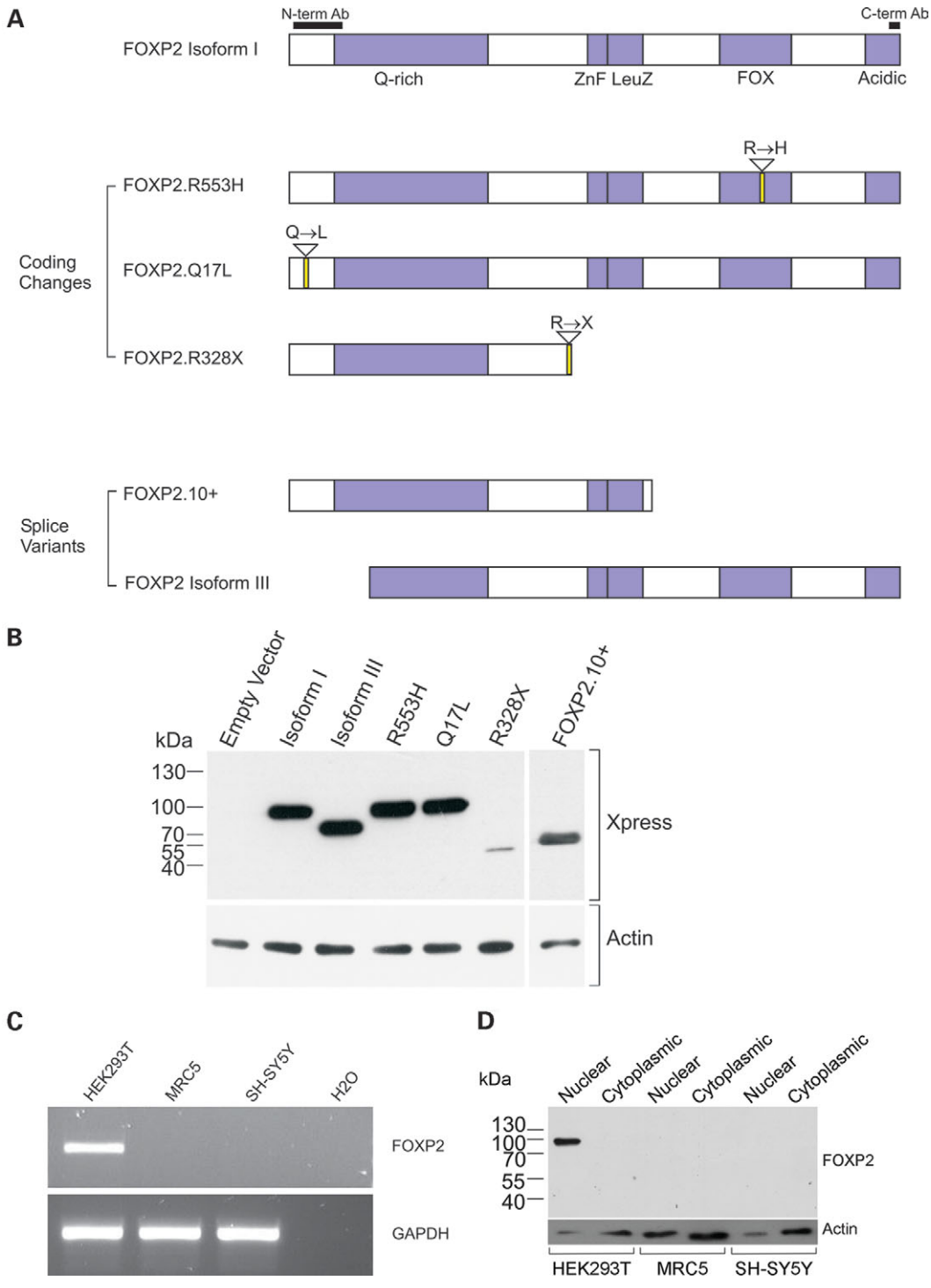

Figure 1. Expression of recombinant FOXP2. (**A**) Schematic representation of the full-length *FOXP2* isoform I containing the glutamine-rich region (Q-rich), zinc finger (ZnF), leucine zipper (LeuZ) and forkhead-box (FOX) domains and the C-terminal acidic tail region (Acidic). The predicted protein product of the coding changes identified in verbal dyspraxia patients (R553H, Q17L, R328X) and two forms of FOXP2 resulting from alternative splicing (FOXP2.10+, isoform III) are also given. Black bars indicate the approximate regions of FOXP2 recognized by the N-terminal and C-terminal FOXP2 antibodies. Note that the N-terminal antibody is not predicted to detect isoform III, whereas the C-terminal antibody is not predicted to detect FOXP2.10+ or FOXP2.R328X. (**B**) Western blot analysis of recombinant FOXP2 proteins. Proteins transiently expressed in SH-SY5Y cells from the pcDNA4/HisMax vector were resolved on SDS-PAGE and detected using an antibody to the N-terminal Xpress™ tag. Equivalent loading was confirmed using an actin internal loading control. (**C**) Endogenous FOXP2 expression was detected by RT-PCR in HEK293T cells but not in SH-SY5Y or MRC5 cells. Primers were designed to detect a region that is common to all the different isoforms investigated in this study. GAPDH was used as a control. (**D**) Western blot analysis of endogenous FOXP2 expression in HEK293T, MRC5 and SH-SY5Y cells. Nuclear and cytoplasmic extracts were resolved on SDS-PAGE and proteins were detected with the N-terminal FOXP2 antibody or an actin antibody as an internal loading control. Endogenous FOXP2 expression could only be detected in HEK293T extracts.

with a monoclonal antibody recognizing the N-terminal Xpress™ tag revealed the presence of recombinant proteins around the predicted molecular weight for all constructs (Fig. 1B). Corresponding data were obtained using polyclonal antibodies raised against N-terminal (Santa-Cruz) and C-terminal (Serotec) epitopes within the FOXP2 protein (Fig. 1A). As expected, the R328X product, which lacks the C-terminus, was only detected by N-terminal antibodies (data not shown). Findings were consistent across a range of transfected human cell-lines that differ in endogenous levels of FOXP2 expression. Although HEK293T cells appear to express endogenous FOXP2 (isoform I), we have not detected FOXP2 expression in our untransfected MRC5 or SH-SY5Y cells via RT-PCR (Fig. 1C), or western blotting using either of the N- or C-terminal antibodies described here (Fig. 1D).

Of note, cells transfected with the R328X construct contained substantially reduced amounts of encoded FOXP2 protein than those found for either the wild-type or other mutant constructs (Fig. 1B) in all cell-lines studied. Independently cloned R328X constructs each showed reduced protein, indicating that this phenomenon was not due to a cryptic mutation elsewhere in the vector. These data suggests that the R328X mutation may interfere with the stability of the mRNA or protein product.

In addition to the mutated variants of FOXP2, different isoforms of the protein resulting from alternative splicing were investigated using the same systems. One isoform, referred to here as FOXP2.10+, is encoded by a putative alternative mRNA transcript that contains a polyadenylation site in the intron following exon 10, and thus excludes exons 11–17 (28). As such, the predicted product lacks the C-terminal regions encoded by exons 11–17, most notably the forkhead DNA-binding motif and the acidic terminal domain (Fig. 1A); these are replaced by a short stretch of novel amino acids encoded by the slightly elongated version of exon 10. We verified the existence of the 10+ transcript in human foetal brain total RNA, cloned it, and introduced the open reading frame into pcDNA4/HisMax. Transfection of human cell-lines yielded robust expression of the tagged protein product at the predicted molecular weight of ~60 kDa (including N-terminal tag), as detected by western blotting of whole cell-lysates with antibodies recognizing the Xpress™ tag (Fig. 1B) or an N-terminal FOXP2 epitope. Like the R328X mutant form, the 10+ isoform lacks the normal C-terminus of isoform I, and hence was not detected by antibodies recognizing C-terminal epitopes (data not shown). Unlike R328X, the transfected 10+ protein product was present in similar amounts to that found for other transfected forms of FOXP2.

We also investigated a third isoform generated by alternative splicing, which is predicted to be translated from an internal initiation site in exon 4, yielding a protein that is truncated at the N-terminus by 92 amino acids (3). This isoform (known as isoform III) retains the characterized functional domains of the larger isoform, including polyglutamine tracts, zinc finger, leucine zipper, DNA-binding domain and acidic C-terminus (Fig. 1A). Expression of isoform III in mammalian cells yielded a stable product that was slightly smaller than isoform I (Fig. 1B) and not recognized by the N-terminal FOXP2 antibody, as predicted.

We observed similar results when studying the same set of wild-type and mutant constructs cloned into pcDNA3.1 vector systems (Invitrogen), featuring either tagged or untagged proteins expressed in significantly lower quantities than the pcDNA4/HisMax system (data not shown).

Intracellular localization of mutant FOXP2 proteins

Like other forkhead-box proteins, FOXP2 is expected to localize primarily to the nucleus of the cell. Indeed, endogenous FOXP2 protein detected in HEK293T cells is predominantly nuclear (Fig. 1D), in line with its putative role as a transcription factor. Transient transfection of human cell-lines with tagged expression constructs, followed by immunofluorescence with antibodies to the Xpress™ epitope, confirmed that FOXP2 isoform I is consistently localized to the nucleus (Fig. 2A). In addition, the protein appears to be generally excluded from nucleoli. Previous studies have shown that a subset of the disease-causing point mutations found in other human forkhead-box genes, including *FOXC1* and *FOXC2*, disrupt the nuclear localization of the encoded protein (25,26). Therefore, we used cell-lines transfected with our tagged mutant constructs of *FOXP2* to assess the impact of the R553H, Q17L and R328X changes on intracellular localization.

Immunofluorescence showed the intracellular localization of R553H to be either nuclear, or both nuclear and cytoplasmic (Fig. 2A). This variable pattern of staining appeared to be influenced partly by the type of cell-line that was transfected—for example, although transfected HEK293T cells tended to show faint staining outside the nucleus, cytoplasmic signals were usually more pronounced for MRC5 and SH-SY5Y cells (see example shown in Fig. 2A). As noted, these cell-lines displayed significant differences in endogenous expression of FOXP2; only HEK293T cells demonstrated detectable levels of endogenous FOXP2 (Fig. 1C–D). In contrast to the variable immunofluorescence results seen with R553H, the Q17L form of FOXP2 was consistently nuclear and indistinguishable from the wild-type, whereas the truncated R328X form showed a striking pattern of diffuse cytoplasmic staining in all transfected cells regardless of cell type (Fig. 2A).

Further investigation via western blotting of separated nuclear and cytoplasmic fractions (Fig. 2B) confirmed the immunofluorescence results and emphasized the effects of the R553H mutation on intracellular localization. Compared with transfected wild-type FOXP2 protein, recombinant R553H protein always displayed clearly reduced amounts in the nucleus, mirrored by concomitant increases in the cytoplasm. When tested by western analysis, the Q17L substitution had no detectable effect on nuclear localization, consistent with immunofluorescence results. The R328X product, which (as noted) showed dramatically lower abundance than other forms, was detected only in the cytoplasm and only following prolonged exposure times (Fig. 2B). The findings for nuclear-cytoplasmic distribution were similar whether using pcDNA4/HisMax (which gives maximal protein expression) or pcDNA3.1-based constructs (which give much milder levels of overexpression).

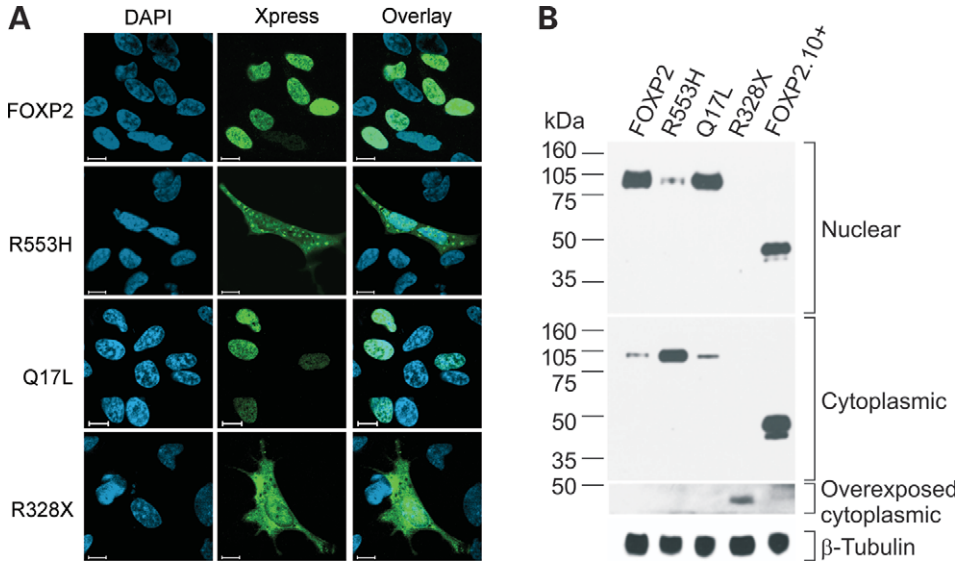

Figure 2. Intracellular localization of mutant FOXP2 proteins. (**A**) Immunofluorescence of transiently transfected SH-SY5Y cells. Recombinant FOXP2 was detected using an antibody to the N-terminal Xpress™ tag (green). DAPI counterstain (blue) indicates the location of nuclei. Wild-type FOXP2 displays predominantly nuclear localization. R553H shows both nuclear and cytoplasmic localization. An example of largely cytoplasmic staining is shown here; however, other cells were observed with a predominantly nuclear pattern of localization and in a small number of cases R553H could be seen to form small aggregates. Q17L staining is nuclear and indistinguishable from the wild-type protein, whereas R328X consistently shows weaker staining which is largely cytoplasmic. (Scale bar, 10 μM). (**B**) Western blot analysis of recombinant proteins transiently expressed in HEK293T cells. Cells were fractionated into nuclear and cytoplasmic compartments prior to SDS-PAGE and western blotting. Proteins were detected using an N-terminal FOXP2 Antibody (Santa Cruz Biotechnology) and equivalent loading was confirmed using the β-tubulin internal loading control. The R328X protein could only be detected following extended exposure (overexposed panel). Endogenous expression of FOXP2 in these cells could be detected with prolonged exposure (Fig. 1D). A similar pattern of results was observed for western blot analyses in MRC5 cells, which do not express detectable levels of endogenous FOXP2 (data not shown).

Mechanisms for nuclear localization of FOXP2

A number of mechanisms for nuclear targeting or exclusion have been documented in FOX proteins. These include nuclear localization or export signals within the forkhead-box domain, phosphorylation and ubiquitylation sites, and interactions with transport proteins (reviewed in 29). In particular, investigations of proteins such as FOXC1, FOXO1 and Foxp3 have highlighted the general importance of the forkhead-box domain for nuclear localization in this family of proteins (20,24,30,31). In our study, the predominantly cytoplasmic location of R328X mutant FOXP2 indicates that the N-terminal half of the protein is insufficient for effective nuclear localization, presumably due to the lack of a FOX domain. In addition, the altered nuclear-cytoplasmic balance of the R553H form suggests that substitutions within the FOXP2 forkhead-box domain can interfere with subcellular localization, in a similar manner to that found for disease-causing mutations in other forkhead proteins, such as FOXC1 (26) and FOXC2 (25).

We sought to investigate further the mechanisms of FOXP2 intracellular localization, to help explain disrupted patterns found for mutant forms of the protein. Previous studies identified the presence of a nuclear localization signal (NLS), enriched for basic residues, at the C-terminal end of the FOXC1 forkhead-box domain (24). Alignment of the FOXP2 forkhead-box domain with that of other FOX proteins indicates partial conservation of the basic NLS that was identified in FOXC1. In order to explore the role of this NLS in FOXP2, site-directed mutagenesis was used to generate two additional constructs in pcDNA4/HisMax. Y580X yields a product that is truncated immediately prior to the putative NLS, whereas G590X generates a product that is truncated just beyond it (Fig. 3A). On western analysis of nuclear and cytoplasmic fractions from cell-lines transfected with these constructs, each was found at appreciable levels in the nucleus. However, Y580X, which lacks the putative NLS, was more abundant in the cytoplasm compared with either the G590X form or full-length FOXP2 isoform I (Fig. 3B). Thus, it seems that the presence of the NLS at the C-terminus of the FOXP2 forkhead-box domain increases the efficiency of nuclear targeting. However, this is clearly not the only mechanism for targeting or retaining FOXP2 in the nucleus since the Y580X variant retained a substantial level of nuclear localization, despite lacking the putative NLS.

Intracellular localization of a brain-expressed isoform of FOXP2 that lacks the forkhead domain

We went on to investigate localization of different isoforms of FOXP2 that result from alternative splicing. The FOXP2.10+ isoform is of particular interest since it lacks the forkhead-box DNA-binding domain and acidic C-terminus, but retains zinc finger and leucine zipper domains that are suspected to help mediate homo- and heterodimerization. We compared the properties of this intriguing splice variant with isoforms I and III, as well as to the R328X mutant product, which has a more severe C-terminal truncation (Fig. 1A).

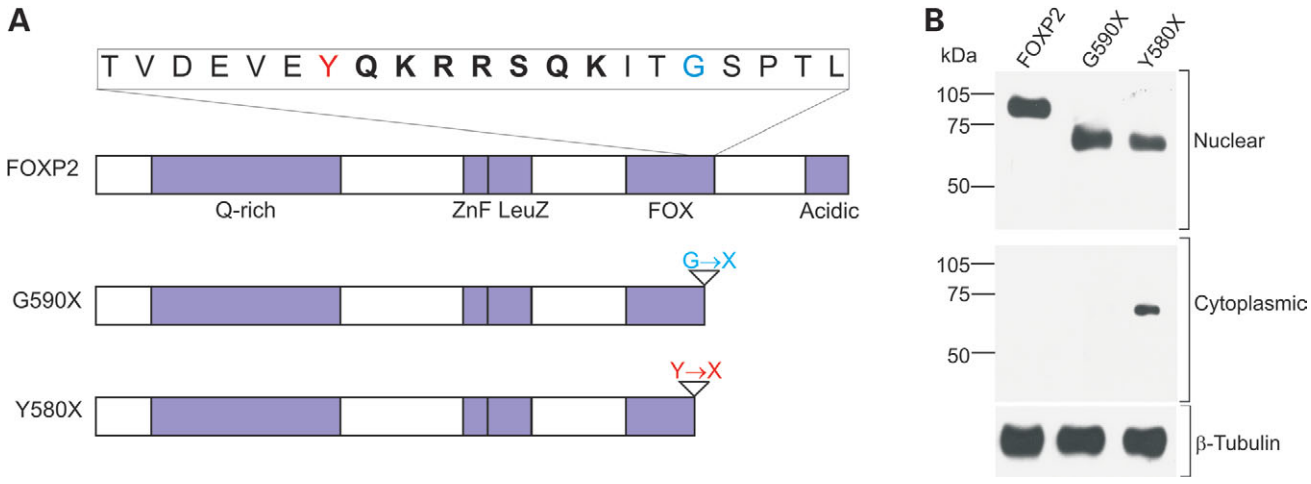

Figure 3. Mechanisms for nuclear localization of FOXP2. (**A**) Schematic diagram of FOXP2 constructs, with expanded amino acid sequence of the region flanking the putative NLS. The NLS is shown in bold. Two constructs were generated via site-directed mutagenesis, introducing stop codons at either Y580 or G590 (indicated in red and blue, respectively). (**B**) Western blot analysis of NLS constructs. HEK293T cells were fractionated into nuclear and cytoplasmic compartments prior to SDS-PAGE and western blotting. Proteins were detected using an N-terminal FOXP2 Antibody (Santa Cruz Biotechnology) and equivalent loading was confirmed using the β-tubulin internal loading control. Increased levels of Y580X could be detected in the cytoplasm compared with either G590X or full-length FOXP2. A similar pattern of results was observed for western blot analyses in MRC5 cells, which do not express detectable levels of endogenous FOXP2 (data not shown).

Immunofluorescence and western blotting of nuclear/cytoplasmic fractions in transfected cell-lines showed no differences in behaviour between isoforms I and III (Fig. 4A). However, the FOXP2.10+ isoform displayed increased cytoplasmic localization, consistent with the lack of a forkhead-box domain (and the associated NLS) (Figs. 2B and 4A). Despite this effect, FOXP2.10+ was still clearly detectable in the nucleus in all cell types tested (Fig. 2B and Supplementary Material), in contrast to the shorter mutant R328X form. Moreover, FOXP2.10+ further differed from R328X by forming brightly stained aggregations in the transfected cells (Fig. 4A). The appearance of these intracellular aggregations was strikingly similar to that previously reported for 'aggresomes'—cytoplasmic (frequently perinuclear) bodies that form via the accumulation of misfolded protein and often become heavily ubiquitylated. The formation of aggresomes is commonly associated with neurodegenerative disease states including Parkinson's and Huntington's disease (32,33).

Co-localization studies using aggresomal markers such as ubiquitin and γ-tubulin (Fig. 4B) supported the hypothesis that the FOXP2.10+-containing bodies in transfected cells do indeed represent ubiquitylated aggresomes (34–36). Markers for other intracellular compartments, such as lysosomes, did not co-localize with FOXP2.10+ (Fig. 4B). Moreover, western analysis using a ubiquitin specific antibody indicated that the FOXP2.10+ protein itself displays ubiquitin moieties under denaturing conditions (Fig. 4C). When other forms of FOXP2 were individually transfected and similarly subjected to western analysis, ubiquitylation was not detected (Fig. 4C, data not shown). The formation of aggresomes containing ubiquitylated FOXP2.10+ was robustly observed even when this isoform was expressed at lower levels (from pcDNA3.1-based constructs, rather than pcDNA4/HisMax).

DNA binding of variant forms of human FOXP2

We investigated the DNA-binding properties of wild-type and mutant forms of human FOXP2 via electrophoretic mobility shift assays (EMSAs), using an oligonucleotide probe that has been shown to be bound by mouse Foxp1 (21) (see Materials and Methods). We first studied DNA binding of purified human FOXP2 protein, obtained from bacterial expression systems. Several forms of FOXP2 protein were expressed and purified, ranging from the Δ-Q-rich protein lacking the glutamine-rich region, to a product containing only the forkhead-box domain (Fig. 5A). Proteins were resolved on SDS-PAGE and levels were standardized by Coomassie staining (Fig. 5B). All wild-type forms of the protein tested were found to bind the probe including the forkhead-box motif alone (Fig. 5B), consistent with recent structural studies suggesting that this domain is still able to bind DNA even if isolated from the rest of the protein (23). However, when the R553H mutation was introduced into the forkhead-box motif, protein–DNA complexes were not observed (Fig. 5B).

We further assessed the DNA binding of mutant forms of FOXP2 using recombinant protein expressed in human cultured cells. Nuclear lysates of transiently transfected cells were normalized via western blotting for equivalent levels of recombinant protein before use in EMSAs (Fig. 6). Full-length wild-type FOXP2 protein (isoform I) was found to bind the probe efficiently and specifically. We confirmed the identity of the protein causing the gel shift to be FOXP2 via the addition of an N-terminal FOXP2 antibody, which disrupted complex formation and produced a 'supershift'. Specificity of the interaction was established following efficient competition by the unlabelled competitor probe. Introduction of a Q17L change did not alter the DNA-binding capacity,

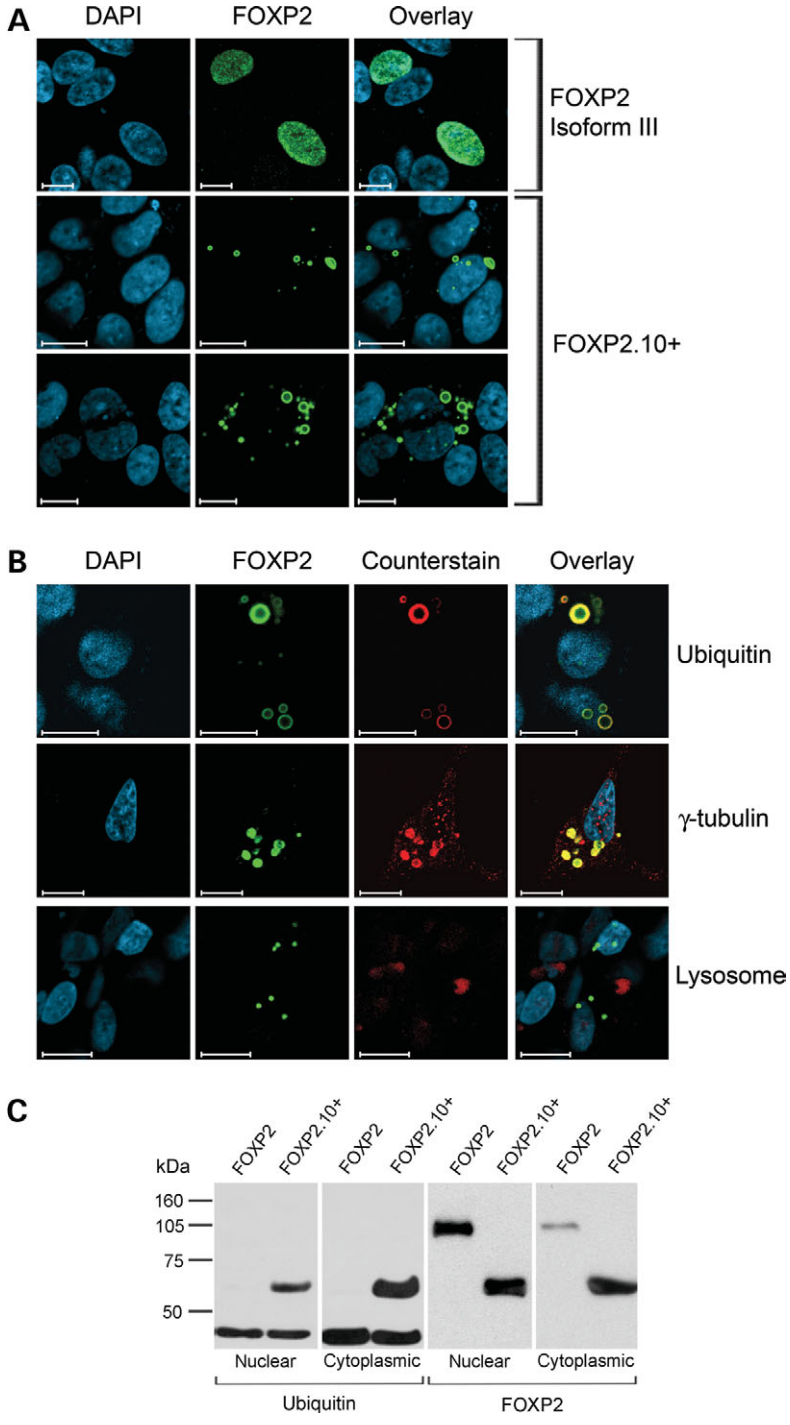

Figure 4. Characterization of FOXP2 Isoforms. Immunofluorescence of transiently transfected SH-SY5Y cells. (**A**) FOXP2 isoform III displays nuclear localization indistinguishable from that of isoform I, whereas FOXP2.10+ is predominantly localized to the cytoplasm and forms cytoplasmic aggregations. Recombinant proteins were detected with an antibody to the N-terminal Xpress™ tag (green) and counterstained with DAPI (blue) to indicate position of nuclei. Two examples of the pattern of staining observed for FOXP2.10+ are given. Aggregations observed in cells transfected with FOXP2.10+ are commonly observed showing both patterns of staining (i.e. small or large aggregations) that may represent different stages of aggresome formation. Due to the bright staining of the aggregations, the accompanying low-level diffuse staining in the nucleus and cytoplasm is not readily apparent here, but is shown in Supplementary Material. (**B**) FOXP2.10+ was detected with an antibody to the N-terminal Xpress™ tag or FOXP2 (green) and counterstained with antibodies to the aggresome markers ubiquitin and γ-tubulin or with a lysosomal counterstain (red). Nuclei are marked by DAPI staining (blue). Ubiquitin and γ-tubulin co-localize with FOXP2.10+ aggregates suggesting that these cellular bodies represent aggresomes. FOXP2.10+ does not co-localize with lysosomes. (Scale bar, 10 μM). (**C**) Western blot analysis of FOXP2.10+ ubiquitylation. Nuclear and cytoplasmic HEK293T extracts were resolved via SDS-PAGE and detected using either a ubiquitin-specific antibody (left) or the N-terminal FOXP2 antibody (right). FOXP2.10+ displays ubiquitylation, whereas ubiquitylation of FOXP2 isoform I was not detected. A non-specific low molecular weight band corresponding to ubiquitylation of other cellular proteins can be observed at approximately even levels in both lanes.

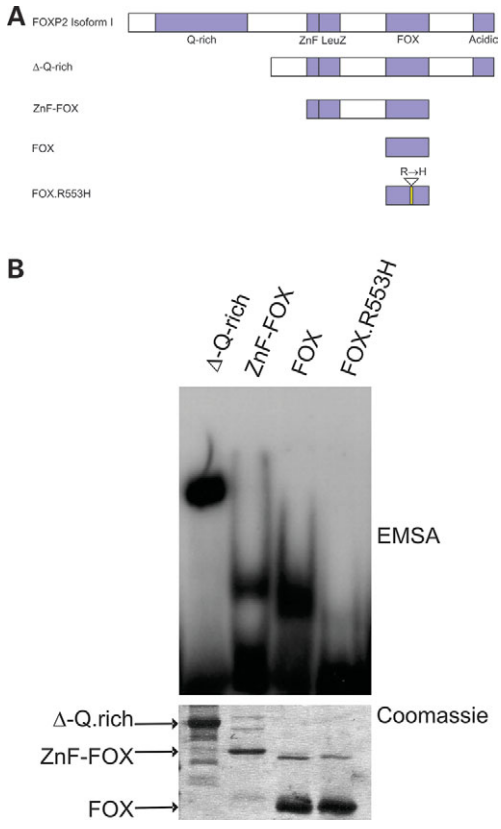

Figure 5. DNA binding of bacterially expressed FOXP2 protein. (**A**) Schematic representation of bacterially expressed FOXP2 proteins used in binding assays. (**B**) EMSA. All wild-type proteins/peptides were able to bind the consensus binding site, whereas binding could not be detected for FOX.R553H. Lack of binding by FOX.R553H was not due to reduced protein levels, as when bacterially expressed proteins were resolved on SDS-PAGE and detected with Coomassie stain, these proteins demonstrated equivalent loading. Arrows indicate position of recombinant proteins. Additional bands detected on Coomassie staining are likely to represent non-specific contaminating proteins.

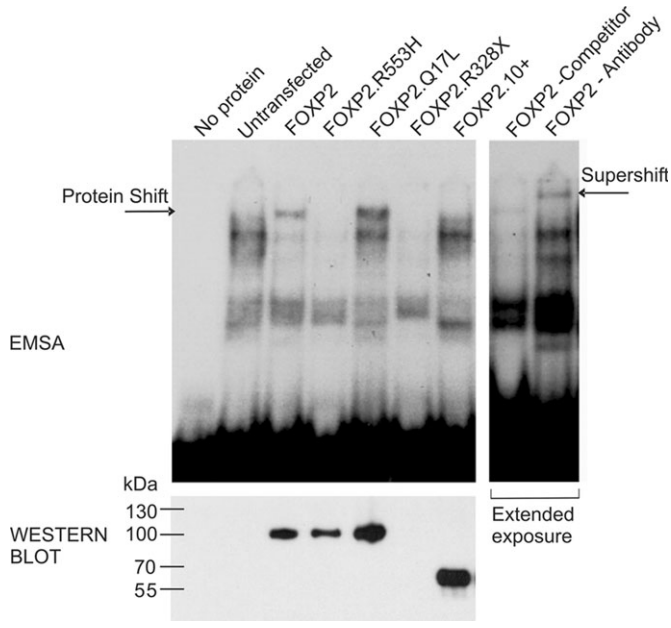

Figure 6. DNA binding of variant forms of human FOXP2. Nuclei from HEK293T cells expressing recombinant FOXP2 proteins were isolated and lysed. Nuclei from untransfected (control) cells were also prepared. Extracts were used in EMSA DNA-binding assays to assess binding capacity of each variant form (upper panel). Although wild-type FOXP2 protein bound the probe, no DNA binding could be observed for any of the variant forms except FOXP2.Q17L. Binding of the labelled probe by FOXP2 was efficiently impaired via competition with unlabelled competitor probe (right). Addition of an N-terminal FOXP2 antibody disrupted complex formation and extended exposure displayed the presence of a supershift, confirming the identity of the protein causing gel retardation to be FOXP2 (far right). Nuclear extracts used for EMSA's were resolved on SDS-PAGE and detected using an N-terminal FOXP2 antibody (lower panel). All proteins were present at comparable levels, except R328X, the presence of which in nuclear extracts is too low to make normalization viable in the context of this binding assay. As a further control, we also performed the assay without the addition of the nuclear extract.

whereas the R553H mutation abolished binding to the probe, consistent with the results obtained using purified protein (Fig. 5B). As expected, the proteins lacking a DNA-binding domain were unable to bind the target DNA. The R328X protein that lacks the zinc finger, leucine zipper and forkhead-box motifs, and the FOXP2.10+ protein that lacks the forkhead-box motif, did not display binding to the probe (Fig. 6).

Transactivation capacities of variant forms of human FOXP2

The identities of *in vivo* downstream targets of FOXP2 remain largely unknown. However, recent work by Wang *et al*. (21) has shown that murine Foxp1 and Foxp2 are able to strongly repress transcription from the SV40 promoter, via binding to a naturally occurring target site that fits a putative core consensus (TATTTRT) for Foxp DNA-binding. Given the extensive homology between human and murine FOXP proteins, we reasoned that human FOXP2 should show similar repression of this promoter via the same binding site. Therefore, we adopted the methods of the previous study, involving luciferase reporter assays in HEK293T cells (21), but used them here as an effective means for evaluating differences in behaviour of coding variants of human FOXP2.

We first verified the relevance of the experimental design to the study of human FOXP proteins, by co-transfecting cells with a pGL3-promoter construct (in which the SV40 promoter drives a firefly luciferase reporter) and a construct expressing either FOXP1, FOXP2 or FOXP4. Like mouse Foxp proteins, the human FOXP proteins tested were found to significantly repress the SV40 promoter 3–7-fold as compared with controls transfected with an empty expression construct ($P < 0.0005$), with magnitudes similar to those found for the corresponding mouse orthologues. FOXP4 showed an effect that was significantly ($P < 0.005$) stronger than FOXP1 or FOXP2 (Fig. 7A). Repression by FOXP2 was weakest, although the difference between this protein and FOXP1 was not statistically significant (Fig. 7A). The observed differences in repression for FOXP subfamily members are comparable to independent findings from another study of murine Foxp proteins that exploited a different cell-line and reporter construct (22). Mutation of the putative consensus target site in the SV40 promoter led to reduced basal activity of the reporter,

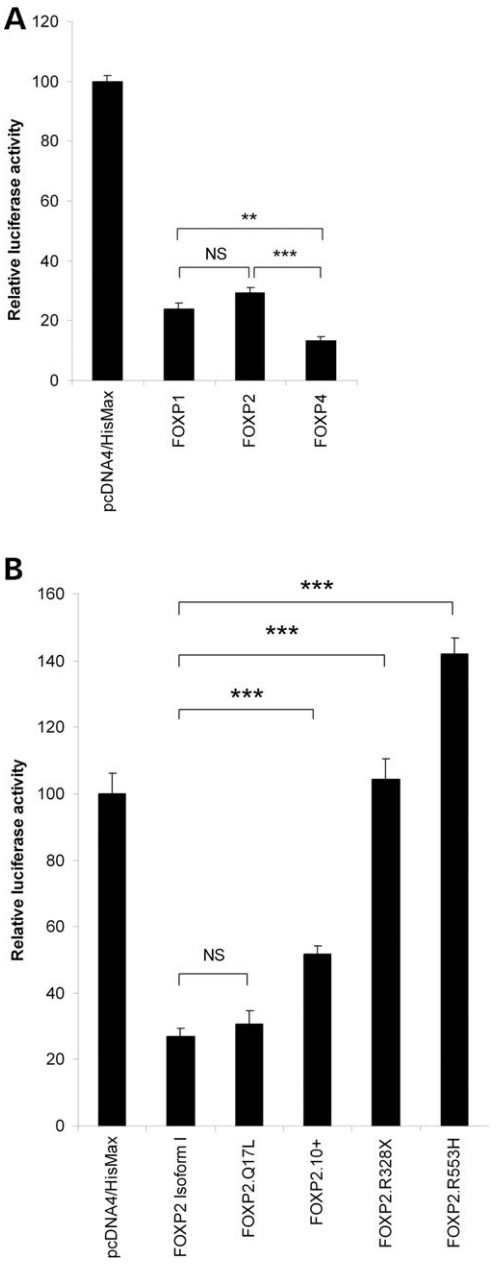

Figure 7. Luciferase assays determine transactivation properties of FOXP2 variants. (**A**) FOXP1, FOXP2 and FOXP4 significantly ($P < 0.0005$) repress pGL3-promoter transcriptional activity through a specific DNA-binding site in the SV40 promoter. Repression levels by FOXP1 and FOXP2 are similar, but FOXP4 is a significantly stronger repressor. The statistical significances of comparisons between the different FOXP proteins are indicated (NS not significant, $^{**}P < 0.005$, $^{***}P < 0.0005$). (**B**) Functional consequences of FOXP2 coding variants and the 10+ isoform, as compared with wild-type FOXP2 isoform I. The transactivation capacity of FOXP2.Q17L is indistinguishable from that of wild-type isoform I, but R553H and R328X fail to repress reporter expression. The introduction of FOXP2.10+ leads to a partial repression of the reporter gene, an effect that is significantly weaker than that observed for isoform I. Statistical significance on this figure is shown relative to values obtained with the wild-type FOXP2 isoform I (NS not significant, $^{***}P < 0.0005$). Results (mean ± SEM, three independent experiments performed in triplicate) are presented as relative luciferase activity, corrected for transfection by β-galactosidase activity of pHSV-LacZ. The control transfection value was obtained with the empty expression vector (pcDNA4/HisMax).

consistent with previous work, and reduced the levels of repression for all three FOXP proteins, with a complete abolition of repression by FOXP4 (data not shown).

We then used the same system to evaluate the transcriptional activity of the three coding variants of FOXP2 found in cases of verbal dyspraxia. Whereas the Q17L form of FOXP2 repressed the SV40 promoter to the same extent as wild-type isoform I, the R328X and R553H mutations each displayed a significant loss of repression ($P < 0.0005$; Fig. 7B). Thus, these latter mutations clearly interfere with the FOXP2 protein's behaviour as a transcription factor. Intriguingly, transfection with the R553H form led to substantial and significant ($P < 0.005$) increases in activity of the reporter as compared with the control situation in which an empty expression construct was transfected. Finally, we assessed the functions of alternative isoforms of FOXP2. Isoform III was found to repress at similar levels to isoform I (data not shown). Surprisingly, expression of FOXP2.10+ still led to significant repression of the SV40 promoter as compared with empty controls ($P < 0.0005$), despite the fact that this isoform lacks a DNA-binding domain. The level of repression by FOXP.10+ was weaker than that seen for isoform I ($P < 0.0005$; Fig. 7B). As discussed subsequently, the behaviour of the R553H and FOXP2-10+ proteins in this assay may potentially relate to interactions with endogenous wild-type FOXP2 isoform I present in HEK293T cells.

DISCUSSION

This paper provides the first direct functional characterization of point mutations implicated in a human speech and language disorder. We studied three unusual coding variants of FOXP2 that have been uniquely found in cases of verbal dyspraxia (3,4). Whereas one change (Q17L) did not substantially affect any aspect of protein function tested, the other two (R553H and R328X) interfered with nuclear localization, DNA-binding and transactivation capabilities of FOXP2. In addition, we investigated the functions of two largely unexplored splice variants suggested by previous RNA analyses, uncovering intriguing properties of the FOXP2.10+ isoform which may be relevant for *in vivo* neural function.

R553H affects multiple aspects of FOXP2 function

The R553H mutation was previously identified in a large three generation pedigree known as the KE family, half of whom suffer from a severe speech and language disorder (3). Individuals carrying this heterozygous mutation presented with verbal dyspraxia accompanied by expressive and receptive language difficulties (7). Structural neuroimaging studies have discovered that people who carry the R553H change have altered grey matter density in multiple regions, including parts of the cortex and in subcortical structures, most notably the striatum and cerebellum (37). Moreover, functional neuroimaging demonstrated that the mutation leads to anomalous patterns of neural activation during linguistic processing (9). The current study provides formal molecular analysis of the properties of the R553H form of the FOXP2 protein, shedding light on

why this simple substitution may have such profound consequences for brain development and function.

Our investigations indicate that the R553H change in FOXP2 impacts on the intracellular localization of the encoded protein. These findings are consistent with previous studies which reported dramatically reduced nuclear localization for pathological substitutions occurring at the paralogous position in the forkhead-box domains of FOXC proteins, including R127H in FOXC1 that causes Axenfeld–Reiger anomaly (26), and R121H in FOXC2 that causes hereditary lymphedema with distichiasis (25). It has been suggested that although the paralogous R-to-H changes occur at some distance from the conserved NLS, they may alter the electrostatic charge distribution of the forkhead-box domain, disrupting interaction of FOXC1/2 proteins with nuclear transport machinery. Alternatively, it is possible that this represents a site possessing auxiliary nuclear localization activity.

Of note, the effects of the R553H mutation on FOXP2 localization observed in the present study were much more subtle than those reported for the paralogous R-to-H changes in FOXC1 and FOXC2 (25,26). We also discovered that patterns of localization for the R553H form varied in different cell types, with stronger cytoplasmic signals in cells in which we could not detect any endogenous FOXP2 protein. These findings underscore the unusual nature of the FOXP subfamily of forkhead proteins. Unlike FOXC proteins, which are thought to act as monomers, the distantly related FOXP proteins have dimerization domains that appear crucial for function (21). This suggests that when FOXP2.R553H is expressed in the presence of endogenous FOXP2, dimerization can occur between mutant and wild-type proteins and such heterodimers may perhaps show more effective nuclear localization than dimers comprised only of mutant protein (i.e. a partial rescue of nuclear targeting). These observations could be relevant for understanding aetiological pathways in affected members of the KE family, who are heterozygous for the R553H mutation and may thus express both wild-type and mutant protein (discussed further subsequently).

The R553H substitution is localized to the third alpha helix of the FOXP2 forkhead-box domain, a region known to make direct contact with the major groove during DNA binding in FOX family members (38–40). This arginine residue is invariant across all known FOX family members and recent structural studies suggest that it plays a key role in both protein–DNA and intra-protein interactions made during binding of FOXP2 to a consensus target site (23). Our EMSA studies using purified proteins (ranging from Δ-Q-rich to isolated forkhead-box domain) and nuclear lysates, clearly demonstrated that the R553H change abolishes binding of FOXP2 to a previously defined FOXP consensus sequence. (It is worth noting the possibility that the R553H protein might still retain some capacity to bind to DNA, but perhaps with altered binding-site specificity.)

Moreover, we found that the transactivation capacity of FOXP2 is substantially altered by the R553H substitution. Using a previously established reporter assay for FOXP function, the R553H change led to the loss of repression of a promoter containing consensus binding sites. In fact, we observed increased expression from the target promoter when FOXP2–R553H was present, as compared with control experiments in which no FOXP2 was transfected. This effect could again relate to FOXP homo- and heterodimerization; the overexpressed mutant form may interfere with the transactivation capabilities of endogenous FOXP proteins present in the cell-lines. At this stage, it is unclear how relevant this situation is to aetiological pathways *in vivo*. It remains to be determined whether the R553H mutation in heterozygous KE members leads to speech/language disorder through a pure loss of function (i.e. reduced dosage of functional protein), or if the presence of the mutant protein also interferes with the activity of wild-type FOXP2 (an additional dominant negative effect).

R328X yields an unstable product that is predominantly cytoplasmic and appears largely non-functional

MacDermot *et al*. (4) reported the detection of a second heterozygous mutation in FOXP2 segregating with a speech and language disorder. The R328X mutation is presently the only nonsense mutation to have been identified for a human speech/language disorder and was present in the proband, affected sib and their mother. Here we explored the properties of the predicted protein product of R328X that lacks known functional domains such as the zinc finger, leucine zipper and forkhead-box. The R328X protein was predominantly located to the cytoplasm, and showed no DNA binding or transactivation capacity in our assays. This suggests that the severe truncation of the protein produces a largely non-functional transcription factor. It is possible that there may be nonsense-mediated RNA decay of the mutant allele *in vivo* in affected humans and data from our *in vitro* model indicate that the stability of the encoded protein may also be compromised. Taking these findings together, it is likely that the problems observed in affected people carrying the R328X mutation result from reduced dosage of FOXP2 protein, rather than a dominant negative effect of the truncated product. There are currently only limited available data describing the behavioural/cognitive phenotype associated with this nonsense mutation in humans, in contrast to the wealth of findings reported for the R553H mutation of the KE family. In future, it will be important to make a detailed comparison between the consequences of these two mutations for speech and language development, and correlate them with the functional genetic findings reported here.

Functional experiments suggest that Q17L is unlikely to have a pathological effect

Like R328X, the Q17L variant was identified by MacDermot *et al*. (4) during a screen of the entire coding region of *FOXP2* in probands diagnosed with verbal dyspraxia. This substitution was not detected at all in 366 random control chromosomes, and so is considered to be a very rare variant of FOXP2. However, in this case the proband had an affected sibling who did not carry the substitution in question, casting doubt on its aetiological significance. Functional experiments were therefore needed to clarify whether/how Q17L affects properties of FOXP2. No deviation from the wild-type status was observed for Q17L in our assays of localization, DNA binding or transactivation. However, we cannot discount the possibility that subtle effects on protein function may be occurring *in vivo*.

Alternative splicing of FOXP2 yields products with distinct functional properties

Differential splicing of a gene can substantially affect protein function via mechanisms such as altered levels of expression, subcellular localization or DNA-binding site recognition/affinity (41). A number of alternative isoforms are predicted to be expressed from the *FOXP2* locus (3,28). In the present study, we cloned and characterized two such alternative isoforms, known as FOXP2.10+ and isoform III. Crucially, while lacking either N- or C-terminal ends found in isoform I, both the alternative isoforms include the known dimerization domains of FOXP2, suggesting that they might be able to interact with isoform I and/or each other in regions of the brain where they are co-expressed. In particular, our functional analyses of the FOXP2.10+ isoform, encoded by a naturally occurring alternative transcript in human foetal brain, give intriguing clues to the potential biological significance of alternative splicing of FOXP2, which may have relevance for its role(s) in neural development/function.

Like the more severely truncated R328X mutant product, localization studies of FOXP.10+ protein indicated increased cytoplasmic location, most likely due to the lack of a forkhead-box domain. Importantly, we observed that transfected FOXP2.10+ was sequestered in cellular protein aggregates, rather than yielding the diffuse staining observed for the R328X product. We suggest that these bodies represent aggresomes, based on our finding that FOXP2.10+ co-localizes with classical aggresome markers such as ubiquitin and γ-tubulin. In addition, western blotting revealed the ubiquitylation of FOXP2.10+ recombinant protein. Ubiquitylation of the R328X protein could not be detected by either immunofluorescence or western analysis. Overall, it is interesting to note the strikingly different behaviour displayed by R328X and FOXP2.10+ products in these experiments. In terms of functional domains, a major difference between these proteins is the presence or absence of the zinc finger/leucine zipper regions that are known to mediate dimerization of FOXP family members.

Although the functions of aggresomes are not fully understood, recent evidence suggests that the sequestration of proteins into these intracellular bodies could play a regulatory role by controlling the quantity of active protein available in the nucleus (35,36,42,43). It should be emphasized that our FOXP2.10+ data come from an *in vitro* model involving protein overexpression, which does not directly reflect the *in vivo* environment. Nevertheless these data hint that the FOXP2.10+ isoform might play a post-translational role in regulating availability/functionality of longer FOXP2 isoforms via interactions involving the dimerization domains. Consistent with this, our luciferase assays in HEK293T cells suggest that transfected FOXP2.10+ may be able to influence expression from a reporter construct despite lacking a DNA-binding domain, perhaps by modulating the behaviour of endogenous FOXP2 present in these cells. The hypothesis that FOXP2.10+ might be able to regulate activity of other FOXP2 isoforms is currently being explored with further functional experiments.

Additional FOXP2 isoforms, predicted by RNA studies, are being cloned and will be investigated using the same techniques and assays described in the current report. FOXP2 has been shown to have a complex pattern of expression in the developing human brain. At present, it is not known whether alternative isoforms show distinctive expression profiles, but it is suspected that spatiotemporal regulation of differential splicing may allow FOXP2 to have diverse functions in distinct neural structures/circuits and at different time points. Thus, future studies will include expression analysis at the RNA and protein level as well as structure–function studies of the various isoforms *in vitro*. Moreover, work is currently underway to investigate the effects of point mutations and alternative splicing *in vivo* using animal models, in order to gain essential insights into how this gene impacts on distributed neural circuits in the developing and mature brain.

MATERIALS AND METHODS

Expression constructs and site-directed mutagenesis

FOXP2 isoforms analysed in the present study were cloned into the pcDNA4/HisMax expression vector (Invitrogen) following reverse transcription from a commercial Human Foetal Brain Total RNA (Clontech) and PCR amplification using primers specific for the full-length sequence, confirming the presence of each transcript *in vivo*. Each insert was fully sequenced to confirm it encodes the appropriate FOXP2 protein in frame with a His/ Xpress™ tag at the N-terminus. To assess the endogenous mRNA expression of FOXP2 in the different cell-lines, primers were design to amplify a 454 bp fragment spanning exons 6 to 9, a region common to all different isoforms studied (Fwd: 5′-CCTTCAGCGTCAGGGACTCA-3′, Rev: 5′-CTGCATTTGCACTCGACACTGAGC-3′).

Variant forms of FOXP2 were generated using the QuickChange Site-Directed Mutagenesis kit (Stratagene) as per manufacturer's instructions with the following primers: R553H; Fwd: 5′-CTTGGAAGAATGCAGTACATCATAAT CTTAGCCTGCAC-3′, Rev: 5′-GTGCAGGCTAAGATTA TGATGTACTGCATTCTTCCAAG-3′, Q17L; Fwd: 5′-GCAA CAGTTCAATGAATCTAAATGGAATGAGCACTCTAAG c-3′, Rev: 5′-GCTTAGAGTGCTCATTCCATTTAGATTCA TTGAACTGTTGC-3′; R328X; Fwd: 5′-CAGTTCTAAGT GCAAGATGAGACAGCTCGTCACATG-3′, Rev: 5′-C ATGTGACGAGCTGTCTCATCTTGCACTTAGAACTG-3′; G590X; Fwd: 5′-GCGAAGGTCACAAAAGATAACATGAA GTCCAACCTTAG-3′, Rev: 5′-CTAAGGTTGGACTTCAT GTTATCTTTTGTGACCTTCGC-3′; Y580X; Fwd: 5′-GT GGATGAAGTAGAATAGCAGAAGCGAAGGTCAC-3′ Rev: 5′-GTGACCTTCGCTTCTGCTATTCTACTTCATCCA C-3′. Direct sequencing was used to confirm presence of the desired mutations. pcDNA4/HisMax/FOXP1 and pcDNA 4/HisMax/FOXP4 were kindly provided by Dr Alison H Banham (University of Oxford, UK).

Cell culture and reagents

HEK293T and MRC5 cells were grown in Dulbecco's Modified Eagles Medium (DMEM) (Sigma) and SH-SY5Y cells in DMEM:F12 media (Sigma). Media was supplemented with 10% Foetal Calf Serum (Sigma), 2 mM L-glutamine (Sigma) and 2 mM Penicillin/Streptomycin (Sigma). Cells were

grown at 37°C in the presence of 5% CO_2. Transfections of HEK293T and MRC5 cells were performed using GeneJuice® (Novagen), whereas transfections of SH-SY5Y cells were carried out with Transfast® (Promega), according to manufacturers' instructions.

The Xpress™ mouse monoclonal antibody (Invitrogen) is directed towards an N-terminal tag present when proteins are expressed from pcDNA4/HisMax, enabling detection of any form of FOXP2 expressed using this vector. Antibodies directed towards FOXP2 included two different goat polyclonals, recognizing epitopes at either the N-terminus (Santa Cruz Biotechnology) or C-terminus (Serotec) of isoform I (Fig. 1). Anti-actin mouse polyclonal (Sigma) and anti-β-tubulin rabbit polyclonal (AbCam) antibodies were used as loading controls during western blotting. An anti-ubiquitin rabbit polyclonal antibody (AbGent) was used to detect ubiquitin moieties and an anti-γ-tubulin mouse monoclonal antibody (Santa Cruz Biotechnology) was used to detect γ-tubulin during immunofluorescence. Secondary antibodies included anti-mouse (BioRad), anti-goat (DAKO) or anti-rabbit (BioRad) HRP-conjugated secondaries for western blotting and anti-mouse Alexa Fluor 488 (Molecular Probes), anti-goat Alexa Fluor 568 (Molecular Probes) or anti-rabbit Alexa Fluor 568 (Molecular Probes) secondaries for immunofluorescence. Staining of lysosomes was via the LysoTracker® Red DND-99 stain (Molecular Probes).

Protein expression and purification from a bacterial system

Construct Δ-Q-rich (amino acids 262–715) was PCR amplified using the following primers; Fwd: 5′-AGTCCTGCT GAGATTCAGCAGTTATGG-3′ and Rev: 5′-GCATAT CAAGCTTTCATTCCAGATCTTCAGATAAAGGCTCTT CTTC-3′. The PCR product was phosphorylated using T4 polynucleotide kinase (NEB) and cut with HindIII (NEB). The restricted product was ligated with PmlI HindIII restricted pET45b, resulting in an N-terminally Hexa-histidine tagged product. For the remaining constructs, the following primers were used: ZnF-FOX (amino acids 337–596): Fwd: 5′-AAGTTCTGTTTCAGGGCCCGGGGGCCTCTCACACT CTCTATGGC-3′ and Rev: 5′-ATGGTCTAGAAAGCTTT ATTTTACTAAGGTTGGACTTCCTGTTATCTTTTG-3′, FOX and FOX.R553H (amino acids 500–596; FOX.R553H has the KE family mutation): Fwd: 5′-AAGTTCTGTTT CAGGGCCCGAATGCAGATGTCAGACCTCCATTTACTT ATGC-3′ and Rev: 5′-ATGGTCTAGAAAGCTTTATTTT ACTAAGGTTGGACTTCCTGTTATCTTTTG-3′. The PCR products were cloned into the pOPINB vector using the In-Fusion™ cloning system in the In-Fusion™ Dry-Down 96-well plate format (Clontech/BD Biosciences) and was performed as described in Berrow *et al*. (manuscript in preparation). The resulting products have an N-terminal hexa-histidine tag that is cleavable with Rhinovirus 3C-protease. All constructs were transformed into Rosetta plysS (Novagen) colonies and plated out on agar plates supplemented with the appropriate antibiotics. Colonies from the plates were used to inoculate 2 l of TB Overnight express (Novagen) supplemented with the appropriate antibiotics. The flasks were vigorously shaken at 240 r.p.m. at 37°C for 6 h, then 20°C for 22 h.

Cells were harvested by centrifugation and resuspended in 50 ml of 50 mM Tris pH 7.5, 45 mM Imidazole, 500 mM NaCl, 2000 units DNAseI (Sigma) and three cOmplete™ mini EDTA free protease inhibitor tablets (Roche Diagnostics Ltd). Cell suspensions were lysed using a Basic-Z Cell Disruptor (Constant Systems) at 30 kpsi and the supernatant was clarified by centrifugation at 48 000*g* for 15 min. All proteins were purified using His-Affinity chromatography on an Äkta-FPLC (GE Healthcare) as a first step. ZnF-FOX, FOX and FOX.R553H were cleaved off the His-column using a GST-3C protease fusion protein. The eluate was flowed through a GST column in series to remove the protease. In all cases, the purest peak fractions (by gel analysis) were pooled. FOX and FOX.R553H were injected directly onto a 16/60 HiLoad Superdex 75 column (GE Healthcare). Hi-Trap heparin columns (GE Healthcare) were utilized with a shallow gradient in order to separate Δ-Q-rich and ZnF-FOX proteins from proteolytic degradation products. The resultant peaks were analysed by SDS-PAGE and fractions containing un-proteolysed protein were pooled. The pools were applied to 16/60 HiLoad Superdex 200 column (GE Healthcare), fractions containing protein were analysed on SDS/PAGE gels and pooled. The pooled proteins were concentrated using either 5000 (FOX and FOX.R553H) or 30 000 (Δ-Q-rich and ZnF-FOX) molecular weight cutoff Vivaspin 6 concentrators (Vivascience). The mass of all purified proteins were confirmed by LC-ESI-MS (HPLC, Dionex) and electrospray ionization mass spectrometry.

Protein expression and purification in human cell culture systems

Cells were seeded at a density of 2×10^4 cells per cm^2, 24 h prior to transfection. Thirty six hours post-transfection cells were washed twice in PBS and proteins were extracted. For whole-cell protein extraction, cells were treated with Lysis Buffer (0.1 M Tris, 150 mM NaCl, 10 mM EDTA, 0.2% Triton X-100, 1% PMSF, protease inhibitor cocktail) at 4°C for 10 min. Cells were centrifuged at 10 000*g* for 30 min at 4°C, allowing cell debris to be pelleted and discarded. To obtain nuclear and cytoplasmic fractions, cells were lysed using a Nuclear Extraction Kit (Panomics) according to the manufacturers' instructions.

SDS-PAGE, western blotting and immunofluorescence

Proteins were resolved on 10% Polyacrylamide SDS gels and transferred to polyvinylidene fluoride (PVDF) membrane (Invitrogen) at 25 volts for 90 min using a wet transfer system (Invitrogen). Membranes were blocked in western blocking buffer (5% Skim Milk Powder, 0.1% Tween 20 in PBS) to prevent non-specific antibody interactions. Proteins were detected using primary antibodies at 4°C overnight. Secondary antibodies were applied for 1 h at room temperature. Proteins were visualized using 'ECL Plus' Enhanced Chemiluminescence Reagents (Amersham Biosciences) and Kodak MXB Film. For Coomassie staining, SDS-PAGE gels were

directly stained with Coomassie stain (0.25% w/v brilliant blue dye, 10% glacial acetic acid, 45% absolute ethanol) overnight and then washed in destain solution (10% absolute ethanol, 10% glacial acetic acid).

For immunofluorescence, cells were seeded onto poly-lysine (Sigma) coated coverslips and transfected as described earlier. Thirty six hours post-transfection cells were fixed using either 100% ice-cold methanol or 4% paraformaldehyde solution at room temperature. Antibodies were diluted in Blocking Solution (1% Fish Gelatine, 0.1% Triton X-100, 5% BSA in PBS). Primary antibodies were incubated at 4°C overnight before incubation with secondary antibodies at room temperature for 2 h under limited light exposure. Where relevant, living cells were incubated with the Lyso-Tracker stain at 37°C for 2 h prior to fixation. Nuclei were visualized using mounting media (VectaShield) containing a DAPI counterstain. Cells were viewed on a Zeiss LSM510 Confocal Microscope with Kr/Ar laser lines of 488 nm (green fluorescence) and 568 nm (red fluorescence). Images were captured using Zeiss LSM Image Software.

Electrophoretic mobility shift assays

Probes were designed as 25-nucleotide oligomers (5′-AGC TTAAACAAGACAACACAAATAA-3′), based on a previously described Foxp1 consensus binding sequence (21). Oligonucleotides and their complementary strands were annealed and 30 pmoles end-labelled via incubation with 2 μl α32P-CTP (Amersham Biosciences), 3 mM dATP, 3 mM dGTP, 3 mM dTTP, 2 U of Klenow enzyme (Roche) in 1X Klenow reaction buffer for 40 min at room temperature. Competitor probes were prepared in a similar manner, but 3 mM dCTP was used in place of α32P-CTP. Probes were diluted 1:1 with molecular biology grade H_2O and purified using Chroma Spin columns (BD Biosciences). Binding reactions were carried out with a 1/20 volume of probe in the presence of 20 mM HEPES, 2 mM EDTA, 2 mM EGTA, 25% glycerol, 0.1875 mg/ml poly(dI-dC) (Pharmacia). When using purified proteins, 1 μg protein was added, whereas for nuclear extracts, 2 μg of extract was used; in each case binding was allowed to continue for 15 minutes at room temperature. When an unlabelled competitor probe was used to confirm specificity of DNA binding, it was added in 5-fold excess and incubated at room temperature for 15 min before addition of labelled probe. Supershift assays were carried out via pre-incubation of polyclonal FOXP2 antibodies with the nuclear lysates for 15 min prior to the binding reaction. Protein–DNA interactions were resolved on a 5% polyacrylamide Tris/Borate/EDTA gel. Gels were dried under vacuum at 80°C.

Luciferase assays

HEK293T cells were seeded at a density of 4×10^4 cells per cm^2 in 24 well plates, 24 h prior to transfection. Experiments involved co-transfection with 25 ng of reporter construct [either pGL3-promoter (Promega) or pGL3-promoterMut; see below], 25 ng of pHSV-LacZ (gift from Dr R. Wade-Martins, University of Oxford, UK) as a normalization control and 200 ng of a construct expressing either FOXP1, FOXP2 (wild-type or mutant), or FOXP4 protein. Forty-eight hours post-transfection, firefly luciferase and β-galactosidase (LacZ) activities were measured as described (44). The activity of the pHSV-LacZ reporter was not affected by either wild-type or mutated FOXP proteins (data not shown). As described earlier (21), the hypothetical FOXP DNA-binding site in the SV40 promoter of the pGL3-promoter construct was modified by site-directed mutagenesis (TTATTTATG to cgAcTTATG) to generate the pGL3-promoterMut construct. Site-directed mutagenesis was performed using the QuickChange Site-Directed Mutagenesis kit as per the manufacturer's instructions (Stratagene). The statistical significance of differences in luciferase activity was assessed using Student's *t*-tests with a significance threshold of $P \leq 0.05$.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENTS

This work was supported by the Brain Sciences Initiative of the Medical Research Council. S.C.V. is a Christopher Welch Biological Sciences Scholar and was supported in part by the Clarendon Fund, Oxford, UK. J.N. was supported by the Swiss National Science Foundation and by a Marie Curie Intra-European Fellowship. F.M.E. is supported by an MRC studentship. S.E.F. is a Royal Society Research Fellow.

Conflict of Interest statement. The authors report that they have no involvements that might raise the question of bias in the work reported or in the conclusions, implications or opinions stated.

REFERENCES

1. Bishop, D.V. (2001) Genetic and environmental risks for specific language impairment in children. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **356**, 369–380.
2. Fisher, S.E., Lai, C.S. and Monaco, A.P. (2003) Deciphering the genetic basis of speech and language disorders. *Annu. Rev. Neurosci.*, **26**, 57–80.
3. Lai, C.S., Fisher, S.E., Hurst, J.A., Vargha-Khadem, F. and Monaco, A.P. (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, **413**, 519–523.
4. MacDermot, K.D., Bonora, E., Sykes, N., Coupe, A.M., Lai, C.S., Vernes, S.C., Vargha-Khadem, F., McKenzie, F., Smith, R.L., Monaco, A.P. *et al*. (2005) Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *Am. J. Hum. Genet.*, **76**, 1074–1080.
5. Marcus, G.F. and Fisher, S.E. (2003) FOXP2 in focus: what can genes tell us about speech and language? *Trends Cogn. Sci.*, **7**, 257–262.
6. Alcock, K.J., Passingham, R.E., Watkins, K.E. and Vargha-Khadem, F. (2000) Oral dyspraxia in inherited speech and language impairment and acquired dysphasia. *Brain Lang.*, **75**, 17–33.
7. Watkins, K.E., Dronkers, N.F. and Vargha-Khadem, F. (2002) Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain*, **125**, 452–464.
8. Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P. and Passingham, R. (1995) Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proc. Natl. Acad. Sci. USA*, **92**, 930–933.
9. Liegeois, F., Baldeweg, T., Connelly, A., Gadian, D.G., Mishkin, M. and Vargha-Khadem, F. (2003) Language fMRI abnormalities associated with FOXP2 gene mutation. *Nat. Neurosci.*, **6**, 1230–1237.

10. Clark, K.L., Halay, E.D., Lai, E. and Burley, S.K. (1993) Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature*, **364**, 412–420.
11. Carlsson, P. and Mahlapuu, M. (2002) Forkhead transcription factors: key players in development and metabolism. *Dev. Biol.*, **250**, 1–23.
12. Lehmann, O.J., Sowden, J.C., Carlsson, P., Jordan, T. and Bhattacharya, S.S. (2003) Fox's in development and disease. *Trends Genet.*, **19**, 339–344.
13. Shu, W., Yang, H., Zhang, L., Lu, M.M. and Morrisey, E.E. (2001) Characterization of a new subfamily of winged-helix/forkhead (Fox) genes that are expressed in the lung and act as transcriptional repressors. *J. Biol. Chem.*, **276**, 27488–27497.
14. Lai, C.S., Gerrelli, D., Monaco, A.P., Fisher, S.E. and Copp, A.J. (2003) FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain*, **126**, 2455–2462.
15. Haesler, S., Wada, K., Nshdejan, A., Morrisey, E.E., Lints, T., Jarvis, E.D. and Scharff, C. (2004) FoxP2 expression in avian vocal learners and non-learners. *J. Neurosci.*, **24**, 3164–3175.
16. Teramitsu, I., Kudo, L.C., London, S.E., Geschwind, D.H. and White, S.A. (2004) Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction. *J. Neurosci.*, **24**, 3152–3163.
17. Bonkowsky, J.L. and Chien, C.B. (2005) Molecular cloning and developmental expression of foxP2 in zebrafish. *Dev. Dyn.*, **234**, 740–746.
18. Fisher, S.E. and Marcus, G.F. (2006) The eloquent ape: genes, brains and the evolution of language. *Nature Rev. Genetics*, **7**, 9–20.
19. Scharff, C. and Haesler, S. (2005) An evolutionary perspective on FoxP2: strictly for the birds? *Curr. Opin. Neurobiol.*, **15**, 694–703.
20. Schubert, L.A., Jeffery, E., Zhang, Y., Ramsdell, F. and Ziegler, S.F. (2001) Scurfin (FOXP3) acts as a repressor of transcription and regulates T cell activation. *J. Biol. Chem.*, **276**, 37672–37679.
21. Wang, B., Lin, D., Li, C. and Tucker, P. (2003) Multiple domains define the expression and regulatory properties of Foxp1 forkhead transcriptional repressors. *J. Biol. Chem.*, **278**, 24259–24268.
22. Li, S., Weidenfeld, J. and Morrisey, E.E. (2004) Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions. *Mol. Cell. Biol.*, **24**, 809–822.
23. Stroud, J.C., Wu, Y., Bates, D.L., Han, A., Nowick, K., Paabo, S., Tong, H. and Chen, L. (2006) Structure of the Forkhead Domain of FOXP2 Bound to DNA. *Structure*, **14**, 159–166.
24. Berry, F.B., Saleem, R.A. and Walter, M.A. (2002) FOXC1 transcriptional regulation is mediated by N- and C-terminal activation domains and contains a phosphorylated transcriptional inhibitory domain. *J. Biol. Chem.*, **277**, 10292–10297.
25. Berry, F.B., Tamimi, Y., Carle, M.V., Lehmann, O.J. and Walter, M.A. (2005) The establishment of a predictive mutational model of the forkhead domain through the analyses of FOXC2 missense mutations identified in patients with hereditary lymphedema with distichiasis. *Hum. Mol. Genet.*, **14**, 2619–2627.
26. Saleem, R.A., Banerjee-Basu, S., Berry, F.B., Baxevanis, A.D. and Walter, M.A. (2003) Structural and functional analyses of disease-causing missense mutations in the forkhead domain of FOXC1. *Hum. Mol. Genet.*, **12**, 2993–3005.
27. Saleem, R.A., Banerjee-Basu, S., Berry, F.B., Baxevanis, A.D. and Walter, M.A. (2001) Analyses of the effects that disease-causing missense mutations have on the structure and function of the winged-helix protein FOXC1. *Am. J. Hum. Genet.*, **68**, 627–641.
28. Bruce, H.A. and Margolis, R.L. (2002) FOXP2: novel exons, splice variants, and CAG repeat length stability. *Hum. Genet.*, **111**, 136–144.
29. Van Der Heide, L.P., Hoekman, M.F. and Smidt, M.P. (2004) The ins and outs of FoxO shuttling: mechanisms of FoxO translocation and transcriptional regulation. *Biochem. J.*, **380**, 297–309.
30. Brownawell, A.M., Kops, G.J., Macara, I.G. and Burgering, B.M. (2001) Inhibition of nuclear import by protein kinase B (Akt) regulates the subcellular distribution and activity of the forkhead transcription factor AFX. *Mol. Cell. Biol.*, **21**, 3534–3546.
31. Zhang, X., Gan, L., Pan, H., Guo, S., He, X., Olson, S.T., Mesecar, A., Adam, S. and Unterman, T.G. (2002) Phosphorylation of serine 256 suppresses transactivation by FKHR (FOXO1) by multiple mechanisms. Direct and indirect effects on nuclear/cytoplasmic shuttling and DNA binding. *J. Biol. Chem.*, **277**, 45276–45284.
32. Ardley, H.C., Scott, G.B., Rose, S.A., Tan, N.G.S., Markham, A.F. and Robinson, P.A. (2003) Inhibition of proteasomal activity causes inclusion formation in neuronal and non-neuronal cells overexpressing parkin. *Mol. Biol. Cell*, **14**, 4541–4556.
33. Waelter, S., Boeddrich, A., Lurz, R., Scherzinger, E., Lueder, G., Lehrach, H. and Wanker, E.E. (2001) Accumulation of mutant Huntingtin fragments in aggresome-like inclusion bodies as a result of insufficient protein degradation. *Mol. Biol. Cell*, **12**, 1393–1407.
34. Garcia-Mata, R., Gao, Y.S. and Sztul, E. (2002) Hassles with taking out the garbage: aggravating aggresomes. *Traffic*, **3**, 388–396.
35. Kopito, R.R. (2000) Aggresomes, inclusion bodies and protein aggregation. *Trends Cell. Biol.*, **10**, 524–530.
36. Tanaka, M., Kim, Y.M., Lee, G., Junn, E., Iwatsubo, T. and Mouradian, M.M. (2004) Aggresomes formed by alpha-synuclein and synphilin-1 are cytoprotective. *J. Biol. Chem.*, **279**, 4625–4631.
37. Watkins, K.E., Vargha-Khadem, F., Ashburner, J., Passingham, R.E., Connelly, A., Friston, K.J., Frackowiak, R.S., Mishkin, M. and Gadian, D.G. (2002) MRI analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain*, **125**, 465–478.
38. Jin, C., Marsden, I., Chen, X. and Liao, X. (1999) Dynamic DNA contacts observed in the NMR structure of winged helix protein-DNA complex. *J. Mol. Biol.*, **289**, 683–690.
39. Kaufmann, E., Hoch, M. and Jackle, H. (1994) The interaction of DNA with the DNA-binding domain encoded by the Drosophila gene fork head. *Eur. J. Biochem.*, **223**, 329–337.
40. Kaufmann, E. and Knochel, W. (1996) 5 years on the wings of fork head. *Mech. Dev.*, **57**, 3–20.
41. Nakahata, S. and Kawamoto, S. (2005) Tissue-dependent isoforms of mammalian Fox-1 homologs are associated with tissue-specific splicing activities. *Nucleic Acids Res.*, **33**, 2078–2089.
42. Taylor, J.P., Tanaka, F., Robitschek, J., Sandoval, C.M., Taye, A., Markovic-Plese, S. and Fischbeck, K.H. (2003) Aggresomes protect cells by enhancing the degradation of toxic polyglutamine-containing protein. *Hum. Mol. Genet.*, **12**, 749–757.
43. Webb, J.L., Ravikumar, B. and Rubinsztein, D.C. (2004) Microtubule disruption inhibits autophagosome-lysosome fusion: implications for studying the roles of aggresomes in polyglutamine diseases. *Int. J. Biochem. Cell. Biol.*, **36**, 2541–2550.
44. Braganca, J., Eloranta, J.J., Bamforth, S.D., Ibbitt, J.C., Hurst, H.C. and Bhattacharya, S. (2003) Physical and functional interactions among AP-2 transcription factors, p300/CREB-binding protein, and CITED2. *J. Biol. Chem.*, **278**, 16021–16029.