Submitted as an *Original Research Paper* to *XXXXXX*

# Deletions and *de novo* variants involving *SOX11* are associated with a human neurodevelopmental disorder

**Abstract**

***SOX11* is a transcription factor which is proposed to play a role in brain development. A role for *SOX11* in human developmental disorders was suggested by a recent report of *SOX11* mutations in 2 patients with Coffin-Siris syndrome. Here we further investigate the role *of SOX11* variants in neurodevelopmental disorders. We identified 6 individuals with chromosome 2p25 deletions involving *SOX11*. These individuals had non-syndromic intellectual disability with microcephaly, developmental delay and shared dysmorphic features. We next utilised trio exome sequencing to identify 3 novel *de novo SOX11* variants. Two of these individuals had a phenotype compatible with Coffin-Siris syndrome while 1 had non-syndromic intellectual disability. The pathogenicity of these mutations was confirmed using an *in vitro* gene expression system. To further investigate the role of loss of *SOX11* in microcephaly we knocked down *SOX11* in xenopus. Morphants had significant reduction in head size compared to controls. This suggests that *SOX11* loss of function can be associated with microcephaly. We thus propose that *SOX11* deletion or mutation can present with either non-syndromic intellectual disability or a Coffin-Siris phenotype.**

## Introduction

The vertebrate *SOX* protein family consists of 30 genes [Pillai-Kastoori *et al*, 2015]. The *SOX* genes are classified into 8 subfamilies (*SOXA-SOXJ*) on the basis of sequence similarity. The *SOX* proteins are transcription factors with a shared motif called the SRY box, a high mobility group (HMG) DNA binding domain. The *SOX* proteins regulate gene expression, acting as either transcriptional activators or repressors, in multiple tissues and so play crucial roles in multiple developmental processes [Pillai-Kastoori *et al*, 2015]. The *SOXC* subfamily consists of *SOX4, SOX11* and *SOX12*. *SOX11* is thought to play a crucial role in brain development. In humans neuron production begins on embryonic day 42 and is largely complete by midgestation [Urban & Guillemot, 2014]. In the fetus the neuronal progenitors are located in the subventricular zone. After production in the subventricular zone neurons migrate outwards into the cortical layers and undergo differentiation into mature neurons. The linked processes of neuronal production from progenitor cells and differentiation into functioning neurons must be tightly regulated to ensure proper brain development [Urban & Guillemot, 2014]. *SOX11* null mice have reduced cortical neurogenesis secondary to reduced proliferation and abnormal differentiation of neuronal progenitor cells [Wang *et al*, 2013]. This results in *SOX11* null mice having reduced brain weights and thin cerebral cortices [Wang et al, 2013]. There is also evidence that *SOX11* plays a role in ocular development. *SOX11* knockdown in zebrafish and xenopus induces microphthalmia with or without iris coloboma [Pillai-Kastoori *et al*, 2014]. The *SOX11* gene thus represents a strong candidate gene for human neurodevelopmental disease.

Haploinsufficiency of other *SOX* genes is known to cause human disease. For example, mutations in *SOX10* are associated with Waardenburg-Hirschprung disease [Pingault *et al*, 1998], and *SOX9* mutations with campomelic dysplasia [Kwok *et al*, 1995]. In contrast, haploinsufficiency of *SOX5* is reported to cause intellectual disability (PMID: 23498568 and PMID: 23220431). Tsurusaki *et al* (2014) reported 2 children with Coffin-Siris syndrome (CSS, OMIM#125900) and *de novo* mutations in *SOX11*. CSS is characterised by developmental delay/intellectual disability, feeding difficulties, facial dysmorphology, microcephaly and hypoplastic nails of the fifth digits

[Santen & Clayton-Smith, 2014]. That *SOX11* mutations can be associated with CSS provides evidence of a role for *SOX11* in human brain development. Both of the mutations reported by Tsurusaki *et al* (2014) affected the HMG DNA binding domain and interfered with the ability of *SOX11* to induce gene transcription *in vitro*. This implicates regulation of gene expression as a mechanism by which *SOX11* contributes to human brain development. Multiple genes regulated at a transcriptional level by *SOX11* have been identified [Sha *et al*, 2012]. In HEK293 cells overexpressing *SOX11,* 251 genes upregulated by at least 1.5 fold were identified, including several genes relevant to neurogenesis (e.g. *TUBB3, CD24*)[Sha *et al*, 2012]. *SOX11* can also repress transcription of genes important for neurodevelopment. For example, in *SOX11* null mice *LIS1* was upregulated significantly [Wang *et al*, 2013] and in *SOX11* knockdown zebrafish *SHH* was upregulated [Pillai-Kastoori et al, 2014]. Altered levels of *SOX11* thus have the potential to cause dysregulation of multiple genetic pathways, with clear potential to disrupt developmental processes.

Here we further investigate a role for *SOX11* in human developmental disease. We identify a series of individuals with chromosome 2p25 deletions including *SOX11* and 3 individuals with *de novo SOX11* point mutations. These individuals had either non-syndromic intellectual disability or a mild CSS phenotype. An *in silic*o analysis demonstrated that expression of *SOX11* is highest in the brain during early fetal life and rapidly decreases into adulthood, suggesting a role for *SOX11* in human neurodevelopment. Lastly, knockdown of *SOX11* in xenopus larvae was associated with microcephaly in the morphants. Taken together, these data suggest that *SOX11* haploinsufficiency impairs brain development and is associated with human neurodevelopmental disorders.

## Materials and methods

### Ascertainment of *SOX11* deletion (2p25.2 deletions) and mutation cases

Individuals with deletion of chromosome 2p25.2, which included the *SOX11* gene, were identified through the DECIPHER collaboration and direct contact with investigators involved in microarray research programmes. Deletions were confirmed by fluorescent in situ hybridisation (FISH). None of the deletions identified were present in the recently published Copy Number Variation (CNVs) map of the human genome, which integrates CNV data from healthy individuals from multiple datasets [Zarrei *et al*, 2015]. Two individuals with *SOX11* mutations were identified as part of the Deciphering Developmental Disorders (DDD) exome sequencing study ((from ~~the~~ data freeze of 1133 children). The methods of DDD are fully described in previous publications [Wright *et al*, 2015]. A third individual with a *SOX11* mutation was identified by exome sequencing via the Genetics of Structural Brain Abnormalities and Learning Disabilities Study (Wales Research Ethics Committee 12/WA/0001) for which the methods have previously been described (PMID: 25855803) All mutations were confirmed by Sanger sequencing (supplementary figure 1).

### *In silico* assessment of pathogenicity of novel *SOX11* mutations

The predicted effect of the *SOX11* missense variants was examined using SIFT, Polyphen and the "Have Your Protein Explained" tool (http://www.cmbi.ru.nl/hope/home). Evolutionary conservation of mutated amino acids was assessed by aligning orthologues in Ensembl (http://www.ensembl.org/index.html). The presence of *SOX11* variants in normal control populations was queried using the ExAC browser (http://exac.broadinstitute.org/gene/ENSG00000176887).

### *In vitro* assessment of effect of *SOX11* mutations on transcriptional activity

#### Expression vectors

The *SOX11* open-reading frame clone was purchased from Promega (Tokyo, Japan) and *SOX11* mutants (c. 150G>C; p. Lys50Asn and c.359C>A; p. Pro120His) generated by site-directed

mutagenesis with the KOD-Plus-Mutagenesis Kit (TOYOBO, Osaka, Japan). WT and mutant *SOX11* cDNAs were PCR amplified and cloned into the p3xFLAG-CMV-14 mammalian expression vector (Sigma, St Louis, MO). The *GDF5* promoter 5′-flanking sequence (−448/+319) was PCR amplified and cloned into the pGL3-basic vector (Promega). All constructs were verified by Sanger sequencing. Human *SOX11* cDNA can be obtained from GenBank/EMBL/DDBJ nucleotide core database under the accession code AB028641.1.

**Luciferase reporter assays**

Transfection and luciferase reporter assays were performed as previously described. HeLa cells were cultured in Dulbecco's Modified Eagle's. Medium (DMEM) -high glucose supplemented with penicillin (50 units / ml), streptomycin (50 µg / ml) and 10% Fetal bovine serum (FBS). Cells were plated in 24-well plates, 24 h before transfection, and transfections performed using TransIT-LT1 (Takara, Ohtsu, Japan) with pGL3 reporter (500 ng / well), effector (250 ng / well) and pRL-SV40 internal control (6 ng / well) vectors. Twenty-four hours after transfection, cells were harvested and luciferase activities measured using the PicaGene Dual SeaPansy Luminescence Kit (TOYO B-Net, Tokyo, Japan). Production of WT and mutant SOX11 proteins was assessed by immunoblot analysis with monoclonal anti-FLAG M2 HRP antibody (1:3,000; Sigma), following the manufacturer's instructions.

***In silico* assessment of *SOX11* expression in developing brain**

Variation of *SOX11* expression levels in the human brain over different developmental stages was investigated by analysing RNA-sequencing data from the Brainspan atlas of the developing human brain (http://www.brainspan.org/). Normalised sample-level RNA-seq data was downloaded and matched to phenotype data for the brain region and age of each sample. The median and interquartile ranges of Log-2 RPKM (reads per kilobase per million) for *SOX11* were calculated for each trimester of pregnancy, age 0-10 years, age 10-20 years and age 30-40 years for each brain region. This was done for the cerebellum, hippocampus, striatum and prefrontal

cortex. Alterations in *SOX11* RPKM between developmental stages were sought using the Kruskall-Wallis test (PASW statistics 18, IBM computing).

Variation of *SOX11* expression in different anatomical areas of the fetal brain was examined using microarray data from the Brainspan atlas from a 15 week female fetus and a 21 week female fetus. *SOX11* expression (log2 normalised) was compared between brain areas with high levels of neurogenesis (ventricular zone and medial, lateral and caudal ganglionic eminences) and low levels (thalamus, cerebellum, brain stem) using the Mann-Whitney U-test (PASW statistics 18, IBM computing).

### ***SOX11* knockdown in xenopus larvae**

*Xenopus laevis* embryos were obtained and cultured according to standard protocols and staged as described previously. All morpholinos (MOs) were obtained by GeneTools, LLC, OR and resuspended in DEPC treated water. For loss of function experiments, Sox11 MO (30 ng per blastomere) was injected. For control experiments, the standard Control MO suggested by GeneTools was used. MOs were injected bilaterally into both dorso-animal blastomeres of *Xenopus* embryos at eight-cell stage to target anterior neural tissue. As a lineage tracer, 0.5 ng *gfp* RNA was co-injected in all experiments to ensure proper injections. For cephalic evaluations *Xenopus* embryos at stage 45 were fixed with formaldehyde and imaged using a Zeiss Axiophot microscope. The head area and papillary distance was measured using ImageJ software. Brains of representative embryos were dissected using fine tweezers. Head area and interpupillary distance in knockdown and control morphants was compared using the Mann-Whitney u-test (Graphpad prism).

## Results

### Clinical case reports

### Individuals with *SOX11* deletions

Case 1 is a 12 year old girl, the first child of healthy non-consanguineous parents. There was no notable family history. She was born at 41 weeks of gestation by caesarean section due to fetal bradycardia and oligohydramnios. Birth weight was 3685 g, length 50 cm (50$^{th}$ centile) and birth head circumference 35 cm (50$^{th}$ centile). At the age of 12 years her head circumference was 50.4 cm (<3$^{rd}$ centile), height 152 cm (25-50$^{th}$) and weight 44kg (75$^{th}$ centile). She had global developmental delay. She walked without support at 18 months and has poor fine motor skills. She started speaking single words at the age of 3 years and 6 months and has had no further language development. She manifests repetitive and stereotyped movements with hyperactivity and autism. She had a nonverbal IQ of 54. She was generally well apart from gastro-oesophageal reflux. On examination she has mild facial asymmetry with right microphthalmia and a wide mouth with thick lips and large, simplified ears. Bilateral 5$^{th}$ finger clinodactyly, cutaneous syndactyly of toes 2-3, scoliosis and inverted nipples were also noted. CGH was reported as: arr 2p25(4,269,295-6,823,106)(hg 18)x1 with deletion of the *SOX11* gene only. The deletion was not present in either parent.

Case 2 is a 5 years and 11 months old girl, the first child of healthy non-consanguineous parents. There was no notable family history. She was born at 36 weeks of gestation by caesarean section because of intrauterine growth restriction. Birth weight was 1,850 g (<10$^{th}$ centile), and head circumference 31 cm (<10$^{th}$ centile). She had developmental delay, first walking independently at 22 months old and speaking her first words at 18 months. At 5 years and 11 months old, weight was 14.5 kg (<3$^{rd}$ centile), height 102.5 cm (<3$^{rd}$ centile), and OFC 47 cm (<2nd centile). She had troublesome gastro-oesophageal reflux. On examination she was noted to have sparse hair, epicanthic folds and a wide mouth with thick lips. Bilateral clinodactyly of the 4$^{th}$ and 5$^{th}$ fingers with generalised joint laxity was noted. A renal ultrasound demonstrated bilateral renal hypoplasia. Brain MRI showed agenesis of the posterior third of the corpus callosum. Formal

ophthalmological examination detected a high degree hypermetropia. CGH was reported as arr 2p25(5,511,851-16,027,633)(hg 19)x1. The deletion encompassed the *SOX11 and MYCN* genes. The deletion was not present in either parent.

Case 3 is a 14 year old girl, the first child of healthy non-consanguineous parents. There was no notable family history. She was born at 37 weeks of gestation. Birth weight was 2 840 g (9$^{th}$ centile) and head circumference 33 cm (2$^{nd}$ centile). She had a weak suck and difficulty feeding in the neonatal period. She had developmental delay. She first walked at the age of 20 months. Her fine motor skills were impaired. She had no speech delay. At 14 years old her height was 158 cm (50$^{th}$) and head circumference was 52 cm (<2$^{nd}$). She was dysmorphic with a high forehead, narrow palpebral fissures, a smooth philtrum, micrognathia and a high arched palate. A scoliosis, inverted nipples, bilateral 5$^{th}$ finger clinodactyly and 2-3 toe syndactyly were noted. CGH was reported as: arr 2p25 (5,756,344-6,940,999)(hg 18)x1. The deletion included *SOX11, RSAD2* and *CMPK2*. The deletion was not found in the mother. The father was not available for testing.

Case 4 is a 25 year old woman, the 3$^{rd}$ child of healthy non-consanguineous parents. There was no family history of note. Pregnancy was complicated by reduced fetal movements. Birth weight was 3130g (25$^{th}$ centile). Choanal stenosis and persistant ductus arteriosus were present in the neonatal period. She had developmental delay; first walking at 4 years and 8 months and speaking her first words at 3 years old. At the age of 25 years old her height was 162 cm (25$^{th}$ centile), weight 50kg (10$^{th}$ centile) and head circumference 49cm (<0.4$^{th}$ centile). She has severe intellectual disability and aggressive behaviour. She is physically well but has hypermetropia. On examination she was noted to have a flat face with upslanting palpebral fissures and wide mouth with thick lips. Short 2$^{nd}$ fingers with bilateral 5$^{th}$ finger clinodactyly was present along with skin syndactyly of toes 2-3 and 4-5 bilaterally. CGH was reported as: arr 2p25 (5,535,091-16,398,225)(hg 19)x1 with deletion of *SOX11 and MYCN*. The deletion was not found in either parent.

Case 5 is a 34 year old man, the 3$^{rd}$ child of healthy non-consanguineous parents (previously published as patient 16 from the study of De Rocker *et al*). He was born at term. Birth weight was

3750g and length 52cm. He had intellectual disability, speech delay and epilepsy. On examination at age 34 years height was 150 cm (3rd centile) and head circumference 53 cm (3rd centile, corrected for height). He had facial dysmorphism: microretrognathia, short philtrum and mild trigonocephaly. Scoliosis and cryptorchidism were also noted. CGH was reported as: arr 2p25(2,231,163- 8,300,707)(hg 19)x1 with deletion of the *SOX11* and *MYT1L* genes. This was paternally inherited, a detailed phenotypic description of the father was not available, but he was reported to be epileptic and have intellectual disability and speech delay [De Rocker et al, 2014].

Case 6 is a 13 year old girl, who was ascertained via DECIPHER with limited phenotypic information available. She was born at term, small for gestational age. She presented as an infant with feeding difficulties, microcephaly and seizures. She is reported to have delayed speech and language development. On examination she was dysmorphic (low set ears, short nose and short philtrum), with clinodactyly of the 5th finger and a single transverse palmar crease. CGH was reported as: arr 2p25(5,209,876-8,078,809)(hg 19)x1. It was not possible to test parental samples.

**Individuals with *SOX11* mutations**

Case 7 is a 12 years and 6 months old girl, the first child of healthy non-consanguineous parents. There was no family history of note. She was born at 41 weeks of gestation. Birth weight was 3118g (26th centile). There was poor feeding in the neonatal period requiring nasogastric feeding. She had global developmental delay; first walking at 2 years and 6 months old. She has never spoken. At the age of 12 years and 6 months her height was 89.4 cm (-1.7 SD), weight 12.15 kg (-1.5 SD) and head circumference 46.5 cm (-3.3 SD). She has profound intellectual disability, severe autism, absence seizures, bruxism and a tendancy to pick at her skin. On examination facial dysmorphism (sunken eyes, depressed nasomaxillary area and wide mouth with thick lips), bilateral 5th finger clinodactyly and hypoplasia of the toe nails of her little toes were noted. She had oculomotor dyspraxia. Her phenotype was felt to be compatible with CSS. Brain MRI was normal. CGH did not demonstrate any pathogenic variants. Exome sequencing identified a *de novo*, heterozygous missense variant in SOX11 (c.359C>A, p. Pro120His; NM_003108.3). No other plausible pathogenic variants were identified on the exome.

Case 8 is an 11 year old boy, the first child of healthy non-consanguineous parents. There was no family history of note. Pregnancy was complicated by 3rd trimester bleeding. He was born at 40 weeks gestation. Birth weight was 3500g (-0.1SD) and head circumference was 34.5 cm (-0.5 SD). He was readmitted to hospital at 3 weeks of age with poor feeding (immature suck, choking and regurgitation of bottle feeds) and failure to thrive. He had global developmental delay; first walking at age 2 years and speaking his first words at age 3. At 11 years of age he was described as being aggressive with poor attention and no sense of danger. He spoke only in short sentences and was in a special needs school. On examination he was on the 9th centile for height and weight and had a head circumference under the 0.4th. He had bilateral conductive deafness, cryptorchidism, hypermetropia and squint. On examination he had facial dysmorphism (malar flattening, a short philtrum and tented upper lip) and a left preauricular skin tag. Bilateral 2-3 cutaneous toe syndactyly and 5th finger clinodactyly was also noted. He was felt to have a phenotype in keeping with mild CSS. CGH did not demonstrate any pathogenic variants. Exome

sequencing identified a *de novo*, missense variant in *SOX11* (c.150G>C, p. Lys50Asn). No other plausible pathogenic variants were identified on the exome.

Case 9 is a boy with microcephaly, cerebellar vermis atrophy and trigonocephaly. His parents were not reported to be consanguineous. His clinical phenotype was one of non-syndromic intellectual disability. Array CGH (180k, Agilent) did not demonstrate any likely pathogenic variants. Exome sequencing identified a *de novo*, nonsense variant in *SOX11* (c.87C>A, p. Cys29*). No other plausible pathogenic variants were identified on the exome.

### ***In silico*** **analysis of SOX11 missense variants**

The c.150G>C (p. Lys50Asn) variant was predicted by SIFT to be deleterious (score of 0) and Polyphen to be probably damaging (score of 1). The c.359C>A (p. Pro120His) variant was predicted by SIFT to be deleterious (score of 0) and by Polyphen to be probably damaging (score of 0.996). Both mutations localise to the HMG (DNA binding) domain. The “Have Your Protein Explained” tool identified that in the c.359C>A (p. Pro120His) variant the mutant residue (histidine) is larger and more hydrophilic than the wild type amino acid. This was predicted to interfere with DNA binding and protein-protein interaction. In the c.150G>C (p. Lys50Asn) variant, aspargine is noted to be of smaller size and neutral charge compared to the wild type amino acid. This was also predicted to interfere with DNA binding. Both variants are found at evolutionary conserved amino acids. Neither variant was found in the exAC database, nor was the c.87C>A, p. Cys29* mutation.

### *In vitro* assessment of effect of *SOX11* mutations on transcriptional activity

*SOX11* is known to regulate the GDF5 promotor. Luciferase reporter assays in HeLa cells indicate that both the p.-Lys50Asn and p.-Pro120His variants display reduced ability to activate the GDF5 promotor compared to wild type protein. This suggests that these *SOX11* variants interfere with the ability of *SOX11* to induce gene expression.

### Expression of *SOX11* in human brain

For all brain regions examined, *SOX11* expression levels (RNA-seq) were highest in the first trimester of pregnancy and then fell significantly (Kruskal-Wallis, $p<0.01$) to reach a nadir in the 4$^{th}$ decade of life (figure 5a-5d). *SOX11* expression levels (microarray)- were significantly greater in brain areas with high levels of neurogenesis compared to areas of low neurogenesis (Mann-Whitney U-test $p<0.01$)(figure 5e) at both 15 weeks and 21 weeks of gestation.

### Knockdown of *SOX11* in xenopus

Knockdown of SOX11 by morpholino injection resulted in a significant reduction in head area and interpupillary distance compared to controls (both $p<0.01$ on Mann-Whitney U-test). Figure 4 shows a representative morphant and control. There was no increased death rate amongst the morphants compared to controls.

## Discussion

Here we report 6 patients with chromosome 2p25 deletions, which include *SOX11,* and 3 individuals with *de novo SOX11* point mutations. The individuals with deletions had several phenotypic features in common (table 1). Microcephaly was reported in all individuals. All had global developmental delay, with particularly marked speech delay. A shared facial dysmorphology of wide mouth and thick lips could be discerned. Two of the patients reported here had trigonocephaly and so it is interesting to note that a 2004 report describes a boy with trigonocephaly, cleft palate and multiple minor anomalies with deletion of *SOX11* resulting from an unbalanced translocation 46,XY,t(2;17)(p25;q24) (PMID: 15540175). Fifth finger clinodactyly and cutaneous syndactyly of toes 2-3 was a frequent examination finding. None of the deletion cases had definite features of CSS such as hypertrichosis, or hypoplasia of the terminal phalanx or nail of the 5$^{th}$ digit or 5$^{th}$ toe. There was, however, considerable variabilty between cases in the number and severity of clinical features – likely reflecting the different deletion sizes. Two of the 3 individuals with *de novo SOX11* variants had features of CSS, similar to those described by Tsurusaki *et al* (2014), while the third had non-specific intellectual disability. Microcephaly, low birth weight and neonatal feeding difficulties associated with hypotonia were frequent findings. Developmental delay was reported in all cases apart from Patient 2 in Tsurusaki *et al's* report. Hypoplasia of the distal phalanx of the 5$^{th}$ finger, broad halluces, 2-3 toe syndactyly and hypoplasia of the nail of the 5$^{th}$ toe are also frequent findings. Ocular abnormalities were present in both deletion and mutation cases; case 1 had right microphthalmia, case 2, 4 and 7 had hypermetropia and squint while case 7 had oculomotor apraxia Our report confirms that *SOX11* mutations can be associated with a CSS phenotype and suggests that *SOX11* deletions and mutations can also be associated with non-syndromic intellectual disability.

The clinical features associated with *SOX11* deletion or mutation overlap with a number of other dysmorphic syndromes. Hypoplasia of the 5$^{th}$ fingers with dysmorphism and intellectual disability can be observed in conditions such as mosaic trisomy 9 [Burns & Campbell, 2015], Deafness-onychodystrophy-osteodystrophy-mental retardation syndrome [Campeau & Hennakem,

2014] and phenytoin embryopathy [Sabry & Farag, 1996]. There is also some clinical overlap between mild forms of Cornelia de Lange syndrome and individuals with *SOX11* mutations [Boyle *et al*, 2014]. However, certain features of CSS, such as the facial dysmorphology, enable distinction from these other disorders.

Our data suggests that there are 3 broad classes of deletion involving the *SOX11* gene. Firstly, relatively small deletions centred on SOX11 (case 1, case 3 and case 6). Secondly, there are larger terminal deletions of 2p25.2p24.3 which also involve the *MYT1L* gene (case 5). Thirdly, there are centromeric deletions which involve *SOX11* and *MYCN* (cases 2 and 4). The mechanism responsible for generating these deletions is unclear – but the lack of common breakpoints implies there is no shared genomic cause such as repetitive elements.

There are several other genes within the deleted regions which could contribute to the observed phenotypes. In the centromeric deletions (case 2 and 4) it is highly likely that *MYCN* deletion contributes to the phenotype. Deletions and mutations of *MYCN* are associated with Feingold syndrome [Cognet *et al*, 2011]. The classical features of Feingold syndrome are microcephaly, intestinal atresias and brachymesophalangy of the 2$^{nd}$ and 5$^{th}$ fingers. However, the severe intellectual disability reported in case 4 is unusual for individuals with Feingold syndrome [Cognet *et al*, 2011]. In the individual with a telomeric deletion (case 5) loss of *MYTL1* is likely to contribute to the phenotype. However case 5 has borderline microcephaly, whilst the other individuals in the report of De Rocker *et al* (2014) [De Rocker *et al*, 2014] who had similar deletions involving *MYT1L* but not *SOX11* tended to have macrocephaly. This suggests that *SOX11* haploinsufficiency may exert a powerful, negative influence on brain growth. Case 1 has no genes other than SOX11 in the deleted region, while the deletion in cases 3 and 6 also contains the *CMPK2*, *RSAD2* and *RNF144A* genes. Deletion of these genes is unlikely to lead to a neurodevelopmental disorder. *CMPK2* encodes a mitochondrial nucleoside monophosphate kinase and heterozygous loss is unlikely to have metabolic consequences [Xu *et al*, 2008]. *RSAD2* encodes viperin which is an antiviral protein [Upadhyay *et al*, 2014]. *RNF144A* is an E3 ubiquitin ligase involved in DNA damage repair and apotosis [Ho *et al*, 2014]. This provides evidence that

*SOX11* deletion alone can be associated with a neurodevelopmental phenotype.

The fact that both heterozygous deletions and mutations of *SOX11* are associated with microcephaly suggests that loss of function and haploinsufficiency may be the underlying mechanism. The 2 missense variants we describe are within the HMG DNA binding domain (as were the 2 previously reported missense variants) while the c.87 C>A variant would be predicted to lead to premature termination of translation prior to this domain. It is thus likely that they will interfere with the ability of *SOX11* to regulate its target genes. Our luciferase reporter gene assays provide further evidence in support of this as they indicate a reduced ability of mutant *SOX11* to induce gene expression. An in vitro study of *SOX11* overexpressing cells identified multiple genes upregulated by *SOX11* which are relevant to neurogenesis and brain development [Sha *et al*, 2012]. Examples include *FILIP1*, which regulates migration of neocortical neurons from the ventricular proliferative zone, and *GPC2*, which is expressed in axonal growth cones [Sha *et al*, 2012]. We hypothesise that haploinsufficiency of *SOX11* could potentially reduce expression of these target genes at critical points in brain development, resulting in a neurodevelopmental disorder.

Data from *SOX11* null mice indicate that *SOX11* plays an important role in regulating neurogenesis [Wang *et al*, 2013]. The expression pattern of *SOX11* in the human brain is also in keeping with the gene playing a role in neurogenesis during embryonic development. We show that *SOX11* expression peaks in the first 3 months of *in utero* life and declines thereafter. Since neurogenesis in the fetal brain is largely completed by mid-gestation [Urban & Guillemot, 2014], this temporal expression pattern fits with *SOX11* being involved in neurogenesis in the fetal brain. The spatial expression pattern of *SOX11* in fetal brain also suggests that *SOX11* is involved in neurogenesis, since *SOX11* expression was significantly higher in the ventricular zone than in areas with relatively low levels of neurogenesis (cerebellum, thalamus, brain stem). Data from animal models also suggests that *SOX11* plays an important role in eye development, since *SOX11* knockdown in zebrafish can cause ocular malformations [Pillai-Kastoori et al, 2014]. In addition, variants in *SOX11* have also been identified in 2 individuals with iris coloboma and no

neurodevelopmental phenotype [Pillai-Kastoori et al, 2014]. This may be explained by the fact that the sequence variants in these individuals were not located in the DNA binding HMG domain of SOX11 protein, while variants reported herein associated with a neurodevelopmental disorder were predicted to interfere with DNA binding.

Our experiments in xenopus larvae indicate that loss of *SOX11* is associated with microcephaly. There was no increased death rate amongst the morphants ; indicating that microcephaly was not a non-specific toxic effect of SOX11 knockdown. The precise mechanism by which loss of *SOX11* results in microcephaly in xenopus is unclear. Our previous work on xenopus eye development indicates that *SOX11* knockdown does not alter proliferation but is associated with increased neuronal apoptosis. This suggests that *SOX11* may function as a neuronal survival factor in brain development.

In conclusion, we describe a series of individuals with *SOX11* deletions or *de novo* mutations. Initially *SOX11* mutations were described as being associated with a CSS phenotype, and the individuals described here with *SOX11* mutations confirm this. The 2 *SOX11* missense variants reported here are the only plausibly pathogenic *SOX11* variants identified from over 1 000 exomes performed in the DDD study. This provides evidence that *SOX11* variants are a rare cause of neurodevelopmental disorders. We also report that individuals with *SOX11* deletion do not have typical clinical features of CSS. This leads us to suggest that *SOX11* deletion or mutation can be associated with a phenotype of non-specific intellectual disability or CSS. Similar findings have been reported for the other CSS genes *ARID1A* and *ARID1B* [Sim *et al*, 2015]. Deletion mapping, in our small cohort, suggests that *SOX11* deletion, although likely pathogenic in itself, can also act as part of a contiguous gene deletion along with loss of *MYT1L* or *MYCN*. The mechanism is likely to be *SOX11* haploinsufficiency with dysregulation of the *SOX11* target genes and consequent disruption of brain development. In conclusion, the current study provides evidence that *SOX11* is a further member of the *SOX* protein family associated with human neurodevelopmental disease.

**References**

Boyle MI, Jespersgaard C, Brondum-Neilsen K, Bisgaard AM, Tumer Z. Cornelia de Lange syndrome. *Clin Genet* 2014; [epub ahead of print]

Burns DA, Campbell E. Twenty-five additional cases of trisomy 9 mosaic: birth information, medical conditions and developmental status. *Am J Med Genet A* 2015; **167**: 997-1007.

Campeau PM, Hennekam RC; DOORS syndrome collaborative group. DOORS syndrome: phenotype, genotype and comparison with Coffin-Siris syndrome. *Am J Med Genet C Semin Med Genet* 2014; **166C**: 327-332.

Cizelsky W, Hemple A, Metzig M, Tao S, Holleman T, Kuhl M, Kuhl SJ. SOX4 and SOX11 function during Xenopus laevis eye development. *PLoS One* 2013; **8**: e69372.

Cognet M, Nougayrede A, Malan V, *et al*. Dissection of the MYCN locus in Feingold syndrome and isolated oesophageal atresia. *Eur J Hum Genet* 2011; **19**: 602-606.

De Rocker N, Vergult S, Koolen D, *et al*. Refinement of the critical 2p25.3 deletion region: the role of MYT1L in intellectual disability and obesity. *Genet Med* 2014; Sep 18. Doi 10.1038/gim.**2014**.124. [Epub ahead of print].

Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. *Nature* 2015; **519**: 223-228.

Ho SR, Mahanic CS, Lee YJ, Lin WC. RNF44A an E3 ubiquiting ligase for DNA-PKcs promotes apoptosis during DNA damage. *Proc Natl Acad Sci* 2014; **111**: E2646-2655.

Kowk C, Weller PA, Guioli S, *et al*. Mutations in SOX9 the gene responsible for Campomelic dysplasia and autosomal sex reversal. *Am J Hum Genet* 1995; **57**: 1028-1036.

Lo-Castro A, Giana G, Fichera M, *et al*. Deletion 2p25.2: a cryptic chromosome rearrangement in a patient with autism and mental retardation detected using aCGH. *Eur J Med Genet* 2009; **52**: 67-70.

Pillai-Kastoori L, Wen W, Morris AC. Keeping an eye on SOXC proteins. *Dev Dyn* 2015; **244**: 367-376.

Pillai-Kastoori L, Wen W, Wilson SG, et al. SOX11 is required to maintain proper levels of

hedgehog signalling during vertebrate ocular morphogenesis. *PLoS Genet* 2014; **7**: e1004491.

Pingault V, Bondurand N, Kuhlbrodt K, *et al*. SOX10 mutations in patients with Waardenburg-Hirschprung disease. *Nat Genet* 1998; **18**: 171-173.

Sabry MA, Farag TI. Hand anomalies in fetal-hydantoin syndrome: from nail/phalangeal hypoplasia to unilateral acheiria. *Am J Med Genet* 1996; **62**: 410-412.

Santen GW, Clayton-Smith J; ARID1B-CSS consortium. The ARID1B phenotype: what we have learned so far. *Am J Med Genet C Semin Med Genet* 2014; **166C**: 276-289.

Sha L, Porteus D, Blackwood D, Muir W, Pickard B. SOX11 target genes: implications for neurogenesis and psychiatric illness. *Acta Neuropsych* 2012; **24**:16-25.

Sim JC, White SM, Lockhart PJ. ARID1B-mediated disorders: mutations and possible mechanisms. *Intractable Rare Dis Res* 2015; **4**: 17-23.

Tsurusaki Y, Koshimizu E, Ohashi H, et al. De Novo SOX11 mutations cause Coffin-Siris syndrome. *Nat Comm* 2014; **5**: 4011.

Upadhyay AS, Vonderstein K, Pichlmair A, et al. Viperin is an iron-sulfur protein that inhibits genome synthesis of tick-born encephalitis virus via radical SAM domain activity. *Cell Microbiol* 2014; **16**: 834-848.

Urban N, Guillemot F. Neurogenesis in the embryonic and adult brain: same regulators, different roles. *Front Cell Neurosci* 2014; **8**: 396.

Wang Y, Lin L, Lai H, et al. Transcription factor SOX11 is essential for both embryonic and adult neurogenesis. *Dev Dyn* 2013; **242**: 638-653.

Wright CF, Fitzgerald TW, Jones WD, *et al*. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet* 2015; **385**: 1305-1314.

Xu Y, Johansson M, Karlsson A. Human UMP-CMP kinase 2, a novel nucleoside monophosphate kinase localized in mitochondria. *J Biol Chem* 2008; **283**: 1563-1571.

Zarrei M, MacDonald JR, Merico D, Scherer SW. A copy number variation map of the human genome. *Nat Rev Genet* 2015; **16**: 172 – 183.

**Acknowledgments**

We thank the families who kindly agreed to be part of this study. This study makes use of data generated by the DECIPHER Consortium. A full list of centres who contributed to the generation of the data is available from http://decipher.sanger.ac.uk and via email from decipher@sanger.ac.uk. Funding for the project was provided by the Wellcome Trust. We declare that those who collected data and deposited it in the DECIPHER database bear no responsibility for its use and interpretation in the current work. The DDD study presents independent research commissioned by the Health Innovation Challenge Fund [grant number HICF-1009-003], a parallel funding partnership between the Wellcome Trust and the Department of Health, and the Wellcome Trust Sanger Institute [grant number WT098051]. The views expressed in this publication are those of the author(s) and not necessarily those of the Wellcome Trust or the Department of Health. The study has UK Research Ethics Committee approval (10/H0305/83, granted by the Cambridge South REC, and GEN/284/12 granted by the Republic of Ireland REC). The research team acknowledges the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network. Dr Alisdair McNeill is supported by the INSIGNEO collaboration for *in silico* medicine at Sheffield University.

**Figure 1. Schematic diagram of deletions.**

The deletions are displayed as a custom track in UCSC genome browser with refseq genes using hg 19. The black bars represent the deletion in each patient and the 2 dashed vertical lines represent the smallest region of overlap. The smallest region of overlap contains only the *SOX11* gene.

**Figure 2. Clinical photographs of study participants.**

2a. Case 2. Facial photograph in top panel, photograph of hands in lower panel (note 5$^{th}$ finger clinodactyly). 2b. Case 4. Facial photograph in top panel, photograph of hands in lower panel (note 5$^{th}$ finger clinodactyly). 2c. Case 3. Facial photograph in top panel, photograph of hands in lower panel (note 5$^{th}$ finger clinodactyly). 2d. Case 1. Facial photograph in top panel.

2e. Boy with c.150G>C (p. Lys50Asn) *SOX11* mutation (case 7). Facial photograph in top panel, hand in middle panel (note 5$^{th}$ finger clinodactyly) and foot in lowermost panel (note broad hallux and 2-3 toe syndactyly). 2f. Boy with c.87C>A (p. Cys29*) SOX11 mutation (case 9). Facial photograph in top panel, hand in middle panel (note 5$^{th}$ finger clinodactyly) and foot in lowermost panel (note broad hallux, 2-3 toe syndactyly and hypoplasia of nail of 5$^{th}$ toe). 2g. Girl with c.359C>A (p. Pro120His) *SOX11* mutation (case 8). Facial photograph in top panel, hand in middle panel (note 5$^{th}$ finger clinodactyly) and foot in lowermost panel (note broad hallux and hypoplasia of nail of 5$^{th}$ toe).

**Figure 3. SOX11 variants identified in current study.**

3a. Schematic diagram of SOX11 protein demonstrating location of 3 reported sequence variants.

3b. 3D models demonstrating alteration of SOX11 protein structure associated with the 2 missense variants. Green areas represent the wild type residue while the red area indicates the structure adopted by the mutant amino acid. Both the missense variants were in the DNA binding domain of SOX11 and predicted to alter its structure, thus interfering with DNA binding.

3c. Bar chart demonstrating that the 2 *SOX11* missense variants had reduce ability to activate the GDF5 promotor in an *in vitro* reporter system. The adjacent Western blot confirms that the mutant proteins were stably expressed during the experiment.

**Figure 4. Sox11 knockdown leads to microcephaly in *Xenopus laevis*.**

4a. Bilateral injection of Sox11 MO results in significant smaller heads measured by the head area (white dotted circles) and the pupillary distance (red lines) compared to bilateral Control MO injections. In addition, Sox11 morphants show an eye phenotype as previously described (red arrowhead; Cizelsky et al., 2013).

4b. Statistical evaluation of the measured head area.

4c. Statistical evaluation of the measured pupillary distance. N, number of individual embryos analyzed.

****, $P \leq 0.0001$. P-values were calculated by a non-parametric Mann-Whitney rank sum test.

**Figure 5. Expression of SOX11 in developing human brain.**

5a-5d show changes in SOX11 expression levels as measured by rna-sequencing in the cerebellum, hippocampus, prefrontal cortex and striatum respectively. Columns labelled first, second and third refer to trimesters of pregnancy. Column 10 represents the first decade of life, 20 the second decade of life and 40 the third and fourth decades. There was a significant decline in SOX11 expression levels with increasing age as assessed by the Kruskall-Wallis test ($p<0.01$).

5e. Microarray data demonstrating that SOX11 is expression is higher in brain regions with high levels of neurogenesis (periventricular) compared to areas with low levels of neurogenesis (cerebellum, thalamus, brain stem), * = $p<0.01$. The first 2 columns represent 15 weeks of gestation while the second 2 columns represent 21 weeks of gestationg.