Edgerley, K., Bryson, L., Hanington, L., Irving, R., Joss, S., Lampe, A., Maystadt, I., Osio, D., Richardson, R., Splitt, M., Sansbury, F. H., Scurr, I., Stewart, H., McNeil, A., & Low, K. J. (2023). SOX5: Lamb–Shaffer syndrome—A case series further expanding the phenotypic spectrum. *American Journal of Medical Genetics, Part A*, *191*(5), 1447-1458. https://doi.org/10.1002/ajmg.a.63124

SOX5: Lamb-Shaffer syndrome – A case series further expanding the phenotypic spectrum

Authors (with institutional affiliations)**:**

Katharine Edgerley[1]
Lisa Bryson[2]
Lucy Hanington[3]
Rachel Irving[4]
Shelagh Joss[2]
Anne Lampe[5]
Isabelle Maystadt[6]
Deborah Osio[7]
Ruth Richardson[8]
Miranda Split[8]
Francis H Sansbury[4]
Ingrid Scurr[1]
Helen Stewart[3]
Alisdair McNeil[9]
Karen Low[1]

[1]Department of Clinical Genetics, University Hospitals Bristol and Weston NHS Foundation Trust, Bristol, UK
[2] Department of Clinical Genetics, NHS Greater Glasgow and Clyde, Glasgow, UK.
[3]Department of Clinical Genetics, Oxford Regional Genetics Service, Oxford, UK
[4]Department of All Wales Medical Genomics Service, NHS Wales Cardiff and Vale University Health Board, Cardiff, UK
[5] Department of Clinical Genetics, South East of Scotland Clinical Genetics Service, Edinburgh, UK
[6] Department of Clinical Genetics, Institute of Pathology and Genetics, Belgium
[7] Department of Clinical Genetics, West Midlands Regional Genetics Service, Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK
[8] Northern Genetics Service, The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK
[9] Department of Clinical Genetics, University of Sheffield, Sheffield, UK

Author contribution:
Edgerley K and Low K designed the study.
Author, recruitment of study subjects, data collection and analysis: Edgerley, K.
Data collection via phenotyping proforma: all other authors.
Supervision: Low, K.
All authors reviewed the results and approved the final version of the manuscript.

Availability of data:
The data that support the findings of this study are openly available in DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl at https://www.deciphergenomics.org

Abstract
Objective: To delineate further the clinical phenotype of Lamb-Shaffer Syndrome (LSS)
Methods: 16 unpublished patients with heterozygous variation in SOX5 were identified either through the UK Decipher database or the study team was contacted by clinicians directly. Clinical

phenotyping tables were completed for each patient by their responsible clinical geneticist. Photos and clinical features were compared to assess key phenotypes and genotype-phenotype correlation.
Results: We report 16 *SOX5* variants all of which meet American College of Medical Genetics/Association for Clinical Genomic Science ACMG/ACGS criteria class IV or V. 7/16 have intragenic deletions of *SOX5* and 9/16 have single nucleotide variants (including both truncating and missense variants). The cohort includes two sets of monozygotic twins and parental gonadal mosaicism is noted in one family.
Conclusions: This cohort of 16 patients is compared with the 7160 previously reported cases and corroborates previous phenotypic findings. As expected, the most common findings include global developmental delay with prominent speech delay, mild to moderate intellectual disability, behavioural abnormalities and sometimes subtle characteristic facial features. We expand in more detail on the behavioural phenotype and observe that there is a greater tendency towards lower growth parameters and microcephaly in patients with single nucleotide variants. This cohort provides further evidence of gonadal mosaicism in *SOX5* variants; this should be considered when providing genetic counselling for couples with one affected child and an apparently *de novo* variant.

Key Words: up to 7
SOX5, Lamb-Shaffer, neurodevelopment, global delay, speech delay

Introduction

The SOX protein family consists of 20 transcription factors (SOXA to SOXH) which play an essential role in cell differentiation, proliferation, terminal maturation and cell survival [Zawerton A et al, 2019]. They mediate important embryonic developmental processes such as sex determination, neurogenesis, chondrogenesis and skeletogenesis [Kamachi et al, 2013; Lamb et al, 2012]. They play a particularly important role in tissue specialization by regulating the timing of cell differentiation [Nesbitt et al, 2015]. *SOX5* resides on chromosome 12p12.1 and is one of the *SOXD* genes. The essential role of *SOXD* genes (*SOX5, SOX6* and *SOX13)* in key developmental pathways has been well established in mouse models [Lefebvre, 2010] and predictive modelling shows haploinsufficiency of SOX5 and SOX6 as being causative [Huang et al, 2010]. *SOX5* codes for at least five transcript isoforms but with three major ones reported. These include two long transcripts (L1 and L2) and one short transcript (S). These different transcripts are tissue specific; for example S-SOX5 is only transcribed in the testes. The longest of these isoforms encodes for a large, 763-amino acid, protein which is the predominant brain isoform and a transcription factor integral to neurogenic cell differentiation. [Ikeda et al, 2002]. Variants in *SOX5* can cause significant neurodevelopmental delay and behavioural traits described in patients with Lamb-Shaffer syndrome, a condition first described in 2012 [Lamb et al, 2012].

Most pathogenic variants in *SOX5* are *de novo* and are either deletions or single nucleotide variants (SNVs). Sometimes, more rarely, splice site variants or translocations cause a dominant disorder due to haploinsufficiency. These variants result in global developmental delay with prominent speech delay, intellectual disability, behavioural abnormalities and some dysmorphic features. The dysmorphic features include down-slanting palpebral fissures, frontal bossing, crowded teeth, auricular abnormalities, prominent philtral ridges, hypotonia and strabismus. [Lee et al, 2013; Schanze et al, 2013]

Previous studies have shown there to be a phenotypic spectrum associated with *SOX5* pathogenic variants. Sometimes, haploinsufficiency of *SOX5* may be compensated for by other members of the *SOXD* family of genes (e.g *SOX6*) resulting in a surprisingly mild phenotype. [Lefebvre, 2010]. Complete haploinsufficiency of *SOX5* has been occasionally shown to cause skeletal abnormalities (including butterfly vertebrae and scoliosis) in addition to the neurodevelopmental features mentioned above. Deletions affecting the short isoform of *SOX5* have been identified in phenotypically normal

individuals (either in unaffected relatives or in a control group) [Lamb et al, 2012]. On the other hand, individuals with large deletions spanning multiple genes tend to be more severely affected, showing more dysmorphic features and having a more severe phenotype (including musculoskeletal anomalies). [Lamb et al, 2012] Prior to *SOX5* having a definitive association with disease, Lamb et al reported seventeen patients with genetic alterations involving *SOX5*, seven of whom had larger 12p12 deletions also encompassing other genes as well. All seven individuals reported with larger 12p12 deletions had behavioural abnormalities. By comparison, most individuals (5/9) with alterations limited to *SOX5* had behavioural abnormalities. Likewise, all individuals with whole *SOX5* gene deletions had anomalies of the hands or feet (including brachydactyly, clinodactyly, contractures) compared to just 4/9 of the patients with intragenic alterations. Individuals with larger deletions also tended to have smaller growth parameters.

To date, there is no evidence of an overt genotype-phenotype correlation observed in intragenic alterations [Zawerton et al, 2019]. For the purposes of this report, we will be focusing solely on intragenic alterations, both SNVs and intragenic deletions, in order to isolate relevant phenotypic information specific to *SOX5.*

We present here a further sixteen patients from thirteen families with Lamb-Shaffer syndrome, where 7/16 have intragenic deletions and 9/16 have SNVs of *SOX5*. We will describe their clinical course and associated genetic findings, comparing these to previously reported cases. These cases strongly correlate to the phenotype previously reported in related to *SOX5* variants but also extend our current understanding of the condition, and will provide essential information for future genetic counselling.

Methods

Patient ascertainment

Subjects were identified through the UK Decipher database, under the Deciphering Developmental Disorders (DDD) study complementary analysis project (CAP) ethical approval process. We selected only those reported to have a pathogenic or likely pathogenic *SOX5* variant. A total of 28 were selected. Their corresponding clinicians were contacted and consent was obtained to be part of the study. Consent was received for 12 patients including two sets of siblings (one of which was a pair of monozygotic twins). Further to this, A CAP was applied for and authorised, indicating to others that we were collecting *SOX5* cases for publication. As a result, several clinicians made personal communications with four additional suitable patients including a further pair of monozygotic twins) who were not yet on the Decipher database. This made a total of sixteen patients.

Clinical information was gathered through completion of a standardised phenotyping questionnaire. to partake in the study and for publication of the subject photographs shown in *Figure 1*. A summary of the genetic and clinical data can be found in *Table 1*. Informed consent was obtained

Genomic analysis

P3, P4, P5, P14 and P15 were recruited through the national DDD study, therefore trio exome analysis was performed. P6, P10, P11 and P16 also had trio exome analysis with application of a large virtual gene panel but were not recruited via the DDD project. All of these SNVs were confirmed with Sanger sequencing. Parental testing was also done by targeted Sanger sequencing.

These sequence variants were annotated using standard databases, filtered based on established criteria and validated by targeted Sanger sequencing. The remaining patients (P1, P2, P7, P8, P9) had copy number aberrations involving chromosome 12p and the *SOX5* gene, established using high resolution quantitative chromosome analysis via array-based comparative genomic hybridisation (aCGH) or, in the case of P12 and P13, by SNP array. Subsequent validation was variable; P12 and 13 were both validated by FISH, P9 by real-time quantitative PCR (qPCR) and in P2 the copy number

variant was first identified by next generation sequencing and validated by aCGH. The remaining three patients with microdeletions (P1, P7 and P8) did not have their result validated by further studies.

Results

16 patients, from 13 families, with ACMG/ACGS class IV or V variants (likely pathogenic or pathogenic) [Richards et al, 2015; Ellard et al, 2020; Riggs et al, 2020] respectively were collated along with clinical phenotypic information using a standardised phenotype proforma. The clinical features and *SOX5* variants in this cohort of 16 patients was compared with the previously reported cases. This data is presented in Table 1. The age of this cohort ranges from 6years 3months to 18years 6months with an mean average age of 11years 7months. The dataset includes 3 sets of siblings: two pairs of monozygotic twins (P3 and P4, P10 and P11) and a further pair of siblings (P7 and P8).

1. Genotype

In this cohort of 16 patients, genetic variation in *SOX5* included point mutations (stop gain, missense and frameshift) and intragenic microdeletions. Most are predicted to cause loss of function re-affirming haploinsufficiency as the mutational mechanism.

Seven patients (P3, P4, P5, P6, P10, P11, P16) had truncating variants (nonsense or frameshift variants). This includes two sets of monozygotic twins (P3 and P4, P10 and P11). P5 and P6 are unrelated individuals with identical nonsense variants. Most of these truncating variants were *de novo*. P5's father was not available to test, though we know the variant was not maternally inherited. The majority of these truncating variants were classified as 'Pathogenic' (ACMG class 5). P5's variant which, despite being the same as P6, could only be classified as ACMG class 4 as it was not possible to establish that this was *de novo*. All six of these truncating variants spare the last exon, and likely trigger nonsense-mediated mRNA decay, limiting protein translation and causing haploinsufficiency.

P14 and P15 had different likely pathogenic missense *SOX5* variants. These *de novo* variants are located in proximity to each other in a mutational hotspot within a high-mobility-group (HMG) domain and are absent from controls (gnomAD). Please refer to Figure 2 in supplementary material for an illustrated diagram [Decipher]. In addition, for P15 multiple lines of computational evidence indicate a deleterious effect of the variant on the gene.

The remaining 7 patients had intragenic microdeletions (P1, P2, P7, P8, P12 and P13). These microdeletions ranged in size from 99kb to 556kb and involved different breakpoints in the gene. All are intragenic (i.e restricted to segments of *SOX5*) and do not involve whole gene deletions nor other genes. All have been reported as pathogenic or likely pathogenic.

Most of these variants (11/16) were undetected in parental blood, suggesting *de novo* occurrence. Inheritance studies were not available for P1 and P5. P12's variant was inherited from her affected mother. The variants in P7 and P8, two siblings with the same alteration, were not detected in either unaffected parent suggesting likely parental gonadal mosaicism.

Three patients have other significant genetic diagnoses alongside LSS. P3 and P4 (twins) both have Klinefelter's syndrome whilst P15 also has a maternally inherited 17q12 deletion resulting in a diagnosis of herediatry neuropathy with pressure palsies (HNPP).

2. Phenotype

Clinical information is presented for Patients 1-16 in Table 1 and compared to the previous publications in Table 2 (please see supplementary information). This patient series consists of 9 males and 7 females. Overall, this cohort corroborates previous findings in the literature with major features including developmental delay and intellectual disability (16/16), speech delay (16/16), behaviour

problems (15/16) and mild dysmorphic features (14/15). As seen in previous cases, there is considerable variability in the clinical features and in their severity.

Perinatal history:
All patients were conceived naturally. The pregnancy of P3 and P4 (twins) was complicated by polyhydramnios, cytomegalovirus infection and obstetric cholestasis requiring ursodeoxycholic acid in the last trimester. Despite these complications, growth parameters of both babies were within normal range. Intrauterine growth parameters were small for P5 (OFC), P6 and P15 (all parameters). P15 was induced at 37 weeks gestation as a result. P11 was the only individual who was born with a birthweight of <-2 SDS. This growth restriction was due to the pregnancy of P10 and P11 (twins) being complicated by twin-twin transfusion syndrome. All patients were born at term except P3 and P4 who were premature (35 weeks gestation).
P2 and P15 had some initial difficulties with feeding but the remaining 14 individuals had no postnatal complications.

Developmental delay and Intellectual disability:
Major features present in all cases include global developmental delay, particularly of speech and language, and intellectual disability. Intellectual disability is generally in the mild to moderate range. All individuals have a statement of educational needs. 12/16 individuals attend a special needs school and the remaining four require a level of additional support in mainstream schooling, including speech and language support. P2, one of the older individuals in the cohort, lives in supported housing with 24hr support. She is independent for some activities of daily living including cooking and volunteering in a community kitchen.

The main challenge around learning (and in managing behaviour) for patients with LSS is usually around speech and communication which clearly compounds any underlying learning difficulty. In this cohort of 16 patients, there is a high burden of speech difficulty; impacting every individual to varying degrees. Many of these patients continue to use sign language and Makaton [Walker M et al, 1981] to aid their communication through to later in childhood and adolescence. Some individuals have very significant speech impairment . For example, P10 and P11 (twins) have no words in their vocabulary other than “no” at the age of 8 years.

Most individuals in this cohort had gross and fine motor developmental, delay, although to a lesser extent than their speech delay. All individuals can walk independently. For 7/16 patients, walking was delayed until after 3 years of age. The age at which independent walking was achieved ranges from 15months to 4years1month with an average age of 2years5months (Figure 3 in supplementary material). P6 developed new fatigue and progressive weakness aged 9, and consequently uses a wheelchair intermittently. Several patients have a diagnosis of dyspraxia and are known to have difficulties holding a pen. Many LSS patients progress into late childhood or early adolescence still struggling to do up buttons and laces or may need help cutting up their food.

4/16 individuals had mild hearing impairment. P9 required a long-term prescription for recurrent middle ear infection and P6 required grommets for glue ear.

Behaviour
Prominent behavioural difficulties are also a significant feature across the majority of individuals in this patient cohort, and often present some of the more challenging aspects of care for parents. 3/16 have a diagnosis of Attention Deficit Hyperactivity Disorder (ADHD), 6/16 have a formal diagnosis of Autism Spectrum Disorder and a further 7/16 have some autistic traits. Most of these patients are intolerant to change in routine, exhibit some repetitive or ritualistic behaviour and have a tendency

to outbursts or tantrums (often due to frustration associated with communication issues). Fussy eating is common and some individuals display quirky behaviours around eating often resulting in poor weight gain. These unusual behaviours extended beyond the content of the food to include the context in which it is being eaten. P5 will only eat at home and this has to be by himself, usually in his bedroom. Similarly, P6 becomes very anxious and distressed around eating in public places or around other people.. She needs a lot of encouragement around mealtimes, has poor weight gain and has subsequently been referred to a specialist feeding clinic.

P5, P10, P11 and P12 all exhibit self-injurious behaviour as part of these outbursts (head banging, throwing oneself against a wall or the floor). Anxiety is also an associated feature and was reported in 4/16 individuals. P5 exhibits some other unusual behaviours including refusing to have his nails cut and having a phobia of sand. He also becomes anxious when asked to sit on chairs out in public, therefore insisting on sitting down on the floor (for example in hospital waiting rooms). He also likes to sleep on the floor as opposed to a bed.

Growth

To stratify growth data for height, weight and head circumference into centiles or SDS, the UK-WHO Growth Chart app was used [Apple app store]. Microcephaly (SDS <-2) is present in 4/16 cases (P3,P 4, P5 and P15) but with no associated abnormality on imaging. All patients with microcephaly have sequence variants, rather than intragenic microdeletions. The head circumference was not recorded in a further three cases.

Our data shows that LSS patients also tend to be short of stature with 6/16 having significant short stature with a height ≤ -2SDS. Interestingly, this is also more prevalent in patients with a single nucleotide variants (5/9 compared to those with an intragenic microdeletion 1/7).

Overall, 10/16 of patients in this cohort were reported to have feeding difficulties and most tended to be on the lower range of weight for their age. Three patients have confirmed lactose intolerance (P10, P11, P12), and 4/16 were diagnosed with gastro oesophageal reflux. Four patients (P3, P4, P5, P6) were markedly underweight with a weight ≤ -2SDS and this was associated with unusual eating behaviours. These patients were all described as 'fussy' eaters by their parents, with very restrictive dietary preferences.

Musculoskeletal

Scoliosis was identified in 3 patients (P4, P11 and P15); a feature which has been noted before in the literature. P15 had a mild pectus deformity. Hypermobility was noted in 9/16 individuals. Several (5/16) patients were found to have long fingers and many (8/16) patients had unusual toes; curled, overlapping, clinodactyly or laterally deviated. See Figure 1 for associated photographs.

Neurology

Hypotonia was noted in many of these individuals (8/16), some of which resolved as the patient became older. Other features include an unsteady, broad-based gait (6/16) and issues with sleep (11/16). Sleep issues included disrupted sleep and difficulty falling asleep in the first instance. Three patients were prescribed medication temporarily to help with sleep difficulties (P1; Methylphenidate and P6 and P15; Melatonin). In two patients (P3 and P8) poor sleep was noted to improve as they reached late childhood/early teenage years. Bed wetting until 6 years was seen in one individual (P13). None of this cohort has a formal diagnosis of epilepsy but P15 is reported to have probable seizures. P5 has changes on EEG (paroxysmal fast activity) but no associated clinical seizures .

MRI brain scans were undertaken for 8/16 patients, five of which indicated some structurally abnormal findings. P5 had right frontal lobe cortical dysplasia. P9 had increased perivascular spaces,

enlargement of lateral ventricles and thinning of the corpus callosum. P10 and P11 (twins) both had arachnoid cysts and ventriculomegaly. P16 had a hypoplastic optic nerve and small optic chiasm.

Ocular

Ocular features are frequently observed in LSS patients. 12/16 individuals from this cohort have ophthalmic abnormalities, making a total of 53% of all LSS patients published to date (see Table 2). Features in this cohort include strabismus (8/16), acuity problems (6/16), nystagmus (3/16), myopia (7/16), astigmatism (2/16), hypermetropia (1/16) and Duane anomaly (1/16).

Cardiac and Renal

The only cardiac abnormality observed in this cohort was a diagnosis of supraventricular tachycardia in P6. P4 and P5 had small kidneys identified on ultrasound scan. There were no other significant renal abnormalities observed.

Skin and Teeth

There is no consistent dermatological phenotype associated with LSS. A variety of skin changes were noted in this cohort. P3 has a fungal nail infection where his toes overlap, P5 has dark patches on his hands, P6 has warts on her feet. P8 and P15 have a large café au lait macule. P11 has eczema and livedo reticularis. P11, 12, 13 and 15 all have dry skin. P12 also has a stork mark on her occiput.

4/16 individuals have problems with dentition. These include overcrowding, poor enamel, multiple caries, ulcers and teeth extractions.

Facial features

Please see Table 1 for detailed phenotypic description of facial features and Figure 1 for associated photographs. Dysmorphic features were variable and typically fairly subtle with no recognisable facial gestalt. Features in common across several patients include; high forehead, epicanthic folds, down-slanting palpebral fissures, prominent ears and a wide mouth. Although this cohort represents a relatively narrow age window (6- 18years) and all are under 19 years of age, previous reports indicate that LSS has an evolving facial phenotype with coarsening of features with age (along with gingival hyperplasia). [G. Innella et al, 2020]. This is seen in P5 as shown in Figure 1.

Discussion:

Table 2 shows a comprehensive overview of all published LSS patients to date. Here we compare our patient data with the phenotypic data of previous reports. The data from Lamb et al 2012 is stratified into intragenic deletions (Column 1) and large deletions involving *SOX5*, five of which include whole gene *SOX5* deletions (Column 2). Refer to Table 3 in supplementary material for comparison of our patient genotypic data with that of previously published LSS reports.

LSS is caused by a range of variation in *SOX5.* Initially, the literature predominantly identified this syndrome as being caused by a critical microdeletion region (12p12.1), the critical gene being *SOX5* [Lamb et al, 2012; Lee et al 2013; Schanze et al 2013; Quintela et al 2015]. This included both large multigenic deletions as well as smaller intragenic deletions. Overall, the literature review identified 5/87 patients with whole deletions of *SOX5* [Lamb et al, 2012], 2/87 patients with large deletions involving partial deletion of *SOX5* and 27/87 patients with intragenic *SOX5* deletions. Subsequently, a *de novo* nonsense point mutation introducing premature stop codon was identified through whole exome sequencing [Nesbitt et al 2015]. Following this, with the improvement of next generation sequencing technologies, additional papers have demonstrated other intragenic point mutations in the *SOX5* gene, such as missense variants and splice site variants [Zawerton A et al, 2019]. Two cases involving balanced reciprocal translocations within *SOX5* have also been reported [Fukushi et al 2017, Lamb et al 2012]. Our cohort reinforces this variant spectrum with a mix of intragenic deletions,

truncating and missense variants. From our literature review we noted that overall, 47% of published LSS patients were found to have a deletion (of which 6% were whole gene deletions), 51% a point mutation and 2% a balanced translocation involving the *SOX5* gene. Of the point mutations, the majority were nonsense (23%) or missense (21%) with fewer reported frameshift (5%) and splice mutations (2%).

Both of the missense variants seen in this cohort (P14, P15) are located in the *SOX5* HMG domain (Figure 2). This domain is important for several cellular functions including DNA binding and bending, nuclear trafficking and protein-protein interactions [Zawerton A et al, 2019]. 15/16 of the pathogenic missense mutations reported in Zawerton et al are also located within this HMG domain with only one located outside it. The high prevalence of missense variants located within the *SOX5* HMG domain in LSS patients, compared to the relatively low proportion of HMG domain missense variants reported in gnomAD (7.5% compared to 21-33% in other regions of the gene) [Zawerton A et al, 2019] indicates it to be a tightly constrained region of *SOX5* and therefore of functional significance. Furthermore, *SOX5* is a gene which exhibits significant missense constraint in control populations overall (Z score: 3.21 with 427 missense variants predicted but only 244 observed) [Zawerton A et al, 2019] providing additional evidence of pathogenicity for missense variants in this gene.

The majority of pathogenic or likely pathogenic *SOX5* variants are *de novo;* 62% of all previously reported cases being confirmed *de novo* and only 5% confirmed to be inherited (3% maternal 2% paternal). 17% of all reported cases have a sibling with the same diagnosis and a significant proportion of these do not have a confirmed *SOX5* variant identified in a parent, indicating a relatively high rate of parental mosaicism in this cohort of patients. At least 14% of the cohort of 41 LSS patients published in Zawerton et al exhibited parental mosaicism [Zawerton A et al, 2019]. This study adds to the evidence of germline mosaicism in *SOX5* variants as P7 and P8 (siblings of different ages) have both inherited an identical *SOX5* alteration from one of their parents, despite parental testing in DNA extracted from blood being negative. Therefore, the overall rate of germline mosaicism seen in all LSS patients published to date is approximately 9%. This is a significantly high proportion of cases and an important aspect to consider when counselling families on recurrence risk for this condition in the presence of an apparent de novo variant.

This cohort contains two sets of monozygotic twins with single nucleotide variants in *SOX5*, a particular pedigree not previously reported in conjunction with LSS. This is most likely due to chance. However, given the small sample size, perhaps there is an underlying mechanism which associates both *SOX5* variants and monozygous twinning. This remains speculative as a literature review did not yield any supportive evidence for this theory and the authors are aware of other ID monogenic syndromes such as KBG Syndrome in which a number of affected monozygous twins have been diagnosed (personal communication, Dr. Karen Low).

Previous reports did not find evidence for any genotype-phenotype correlation for LSS and suggested that the spectrum of clinical features is likely explained by variable gene expressivity. Although there is no obvious correlation with severity of intellectual disability or developmental delay, there does seem to be some evidence that truncating sequence variants are more likely to result in smaller growth parameters than intragenic deletions. All of the microcephalic (≤ -2SDS) patients in this cohort (4/16) have truncating single nucleotide variants. Furthermore, five of the six patients in our cohort with significant short stature (≤ -2SDS) have *SOX5* sequence variants. This finding is reflected in, though not commented on, previously reported cases [Supplementary table; Zawerton A et al, 2019, Fukushi et al 2017]. Previous literature indicates that the point mutation (or balanced translocation) patients are the ones who are more likely to be small with small heads (but not exclusively so). There are patients with 12p microdeletions who are also reported to have short stature and microcephaly

(SDS<-2), but proportionally seem to be fewer than in the sequence variant cohort of LSS patients. Zawerton et al suggest that brain growth is frequently altered in the patient cohort with sequence variation. [Zawerton et al, 2019]. There is currently no diagnosis of epilepsy in our cohort, other than P12's mother, despite *SOX5* pathogenic variants being shown to predispose to epilepsy. [Zawerton et al, 2019; Innella et al, 2020].

Reviewing the published literature and drawing from data in this paper, these is currently no evidence to suggest incomplete penetrance associated with pathogenic SOX5 variants, although there is marked variable expression associated with LSS. In our cohort, two families demonstrated inheritance of an intragenic *SOX5* variant from a parent, though neither suggests incomplete penetrance of the condition. P12's mother is affected herself, with mild learning difficulties and epilepsy. P7 and P8 have both inherited their *SOX5* alteration from an unaffected parent, but this was due to germline mosaicism. Similarly, in Zawerton et al, the LSS variants which were inherited either came from an affected parent (1/34) or an unaffected germline mosaic parent (5/34).

In summary, this cohort further illustrates the genetic and clinical spectrum associated with LSS. The genotype incorporates deletions, stop-gain and missense variants that cause loss of function. These mechanisms of pathogenesis are well established for this condition. Our case series further corroborates previous clinical findings associated with LSS whilst expanding the phenotype further. In particular, we have described in more detail some of the unusual behavioural and feeding features found in this condition. There is variable phenotypic expression and no clear clinical indication for any specific recommended screening protocols. Our cohort supports evidence of gonadal mosaicism in LSS which ought to be reflected in the counselling around recurrence risk of an apparently de novo *SOX5* variant. We also highlight for the first time the suggestion that *SOX5* truncating sequence variants appear to be associated with smaller growth parameters and head circumferences than 12p microdeletions.

Acknowledgements:

This study makes use of data generated by the DECIPHER community. A full list of centres who contributed to the generation of the data is available from https://deciphergenomics.org/about/stats and via email from contact@deciphergenomics.org. Funding for the DECIPHER project was provided by Wellcome [grant number WT223718/Z/21/Z].

The DDD study presents independent research commissioned by the Health Innovation Challenge Fund [grant number HICF-1009-003], a parallel funding partnership between Wellcome and the Department of Health, and the Wellcome Sanger Institute [grant number WT098051]. The views expressed in this publication are those of the author(s) and not necessarily those of Wellcome 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.

References:

Apple app store. (Accessed 2020, Nov 11)
https://apps.apple.com/gb/app/growth-charts-uk-who/id916579608

Decipher. (Accessed 2022, November 14)
https://www.deciphergenomics.org/patient/264625/genotype/193098/protein

DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources. Firth, H.V. *et al.*, 2009. Am.J.Hum.Genet 84, 524-533
(DOI: dx.doi.org/10/1016/j.ajhg.2009.03.010)

Ellard S, et al. 2020. ACGS Best Practice Guidelines for Variant Classification in Rare Disease. *Association for Clinical Genomic Science (ACGS*)

Fukushi D, Kenichiro Y et al. 2018. Clinical and genetic characterization of a patient with *SOX5* haploinsufficiency caused by a *de novo* balanced reciprocal translocation. *Gene* 655: 65-70

Huang N, Lee I, Marcotte EM, Hurles ME. 2010. Characterising and predicting haploinsufficiency in the human genome. *Plos Genetics* 6:e1001154

Ikeda et al. 2002. Identification and characterization of the human long form of Sox5 (*L-SOX5*) gene. *Gene* 298: 59-68

Kamachi Y KondohH. 2013. Sox proteins: regulators of cell fate specification and differentiation. *Development* 140:4129-4144.

Lachner, A., Backfisch, I., Hoogerheide, V., van Gog, T., & Renkl, A. 2020. Timing matters! Explaining between study phases enhances students' learning. *Journal of Educational Psychology*, 112(4), 841–853.

Lamb AN, Rosenfeld JA, Neill NJ, et al. 2012 Haploinsufficiency of SOX5 at 12p12.1 is associated with developmental delays with prominent language delay, behaviour problems, and mild dysmorphic features*. Human Mutation* 33:728-740.

Lee et al. 2013. Deletion 12q12 Involving *SOX5* in Two Children With Developmental Delay and Dysmorphic Features. *Pediatric Neurology* 48:4: 317-320

Lefebvre V. 2010 The SOxD transcription factors-Sox5, Sox6, and Sox13—are key cell fate modulators. *The International Journal of Biochemistry and Cell Biology*. 42:429-32

Nesbitt A, Bhoj EJ et al. 2015. Exome sequencing expands the mechanism of SOX5-associated intellectual disability: A case presentation with review of sox-related disorders. *American Journal of medical genetics AJMG* 167A: 11: 2548-2554

Richards, S et al. 2015. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, 17 (5); 405-424.

Riggs et al. 2020. Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American college of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genetics in Medicine*. 22 (2); 245-257

Schanze et al. 2013. Haploinsufficiency of SOX5, a member of the SOX (SRY-related HMG-box) family of transcription factors is a cause of intellectual disability. *European Journal of Medical Genetics*. 56: 2: 108-113

Walker M, Armfield A. What is the Makaton vocabulary? Spec Educ Forward Trends. 1981 Sep;8(3):19-20. PMID: 6458105)

Zawerton A et al. 2019. Widening of the genetic and clinical spectrum of Lamb-Shaffer syndrome, a neurodevelopmental disorder due to *SOX5* haploinsufficiency. *Genetics in Medicine*. 22: 524-537

SOX5: Lamb-Shaffer syndrome – A case series further expanding the phenotypic spectrum

No conflict of interest

This is a statement to confirm that the authors of this paper have no conflict of interest to declare.