CHAPTER 4

MODELLING A DELETION OF THE HUMAN REGION 5q35.2–q35.3 IN MICE

4.1 GENERAL OVERVIEW

Haploinsufficiency of the human 5q35 region spanning the *NSD1* gene results in a rare condition known as Sotos syndrome. Patients with Sotos syndrome display a variety of clinical features, including pre- and postnatal overgrowth, advanced bone age, intellectual disability, and hypotonia, as well as facial, cardiovascular and/or urinary/renal abnormalities. To search for dosage-sensitive genes involved in this disease, we used chromosome engineering to generate a monosomic mouse model carrying a deletion of the *4732471D19Rik* to *B4galt7* interval, syntenic with 5q35.2–q35.3 in humans. Haploinsufficiency for the 36 genes in this interval resulted in no gross morphological defects. In particular, no overgrowth was observed. However, histopathological analysis showed dilation of the pelvicalyceal system in the kidneys of monosomic mice, which mimics the renal abnormality observed in patients with Sotos syndrome. Thus haploinsufficiency of a gene (or genes) within the *4732471D19Rik–B4galt7* interval successfully models the renal abnormality observed in patients with Sotos syndrome.

4.2 INTRODUCTION

4.2.1 OVERVIEW OF SOTOS SYNDROME

Sotos syndrome (SoS OMIM ID: 117550), also known as cerebral gigantism, is named after Dr Juan Sotos who in 1964, first described a group of children with overgrowth, acromegaly, and intellectual disability (Sotos 1964). Indeed, SoS is one of the most commonly diagnosed overgrowth disorders after Beckwith-Wiedeman syndrome (Baujat 2007). Patients with SoS are also diagnosed with prenatal overgrowth, facial abnormalities, macrocephaly, advanced bone age, large hands and feet, and coordination problems. Moreover, hypotonia, poor feeding, jaundice, frequent ear and chest infections, cardiac and urinary/renal defects, seizures, scoliosis, and behavioural problems have also been observed in some SoS patients. However, the variety and degree of severity of SoS symptoms decrease with age, and some clinical features, including overgrowth, hypotonia, poor feeding and infections, become much less apparent in children after the first few years of life (Cole 1994; Tatton-Brown 2007). Patients with SoS might also have an elevated level of predisposition to cancer as some SoS individuals have been diagnosed with neuroblastoma, ganglioglioma, sacrococcygeal teratoma, hepatoblastoma, acute lymphoblatic leukeamia or small cell lung carcinoma (Saugier-Veber 2007). However, the overall risk of tumourgenesis is estimated to be lower than 3% in SoS individuals (Tatton-Brown 2007).

4.2.2 THE MOLECULAR AND GENETIC BASIS OF SOTOS SYNDROME

SoS shows an autosomal dominant mode of inheritance (Baujat 2007). In 2002, the nuclear receptor SET domain containing protein 1 (*NSD1*) was identified as a causative gene for SoS after *NSD1* was found to be disrupted in a child carrying an apparently balanced *de novo* t(5;8)(q35;q24.1) translocation (Kurotaki 2002). Subsequent analysis of patients clinically diagnosed with SoS showed that haploinsufficiency of *NSD1* due to intragenic *NSD1* mutations, partial *NSD1* deletions or chromosomal microdeletions

spanning the 5q35 region encompassing the entire *NSD1* gene accounted for more than 90% of cases (Kurotaki 2002; Cecconi 2005; Douglas 2005; Tatton-Brown 2005; Saugier-Veber 2007; Tatton-Brown 2007), while deletions and uniparental disomies of the 11p15 region or deletions of the *GPC3* gene have been identified in some SoS individuals that did not show any *NSD1* anomalies (Douglas 2005; Tatton-Brown 2007).

The prevalence of intragenic *NSD1* mutations and 5q35 microdeletions encompassing the *NSD1* gene depends greatly on the ethnic origin. Namely, up to 80% of non-Japanese SoS cases but only approximately 12% of Japanese SoS cases are identified with intragenic *NSD1* mutations, while more than 50% of Japanese SoS cases but only approximately 10% of non-Japanese SoS cases are detected with 5q35 microdeletions (Tatton-Brown 2005; Baujat 2007; Saugier-Veber 2007). The increased frequency of 5q35 microdeletions in Japanese SoS cases might be attributed to an inversion polymorphism causing a predisposition to microdeletions that is commonly found in the Japanese (Visser 2005).

More than 200 different intragenic *NSD1* mutations have been described so far in patients with SoS (Kurotaki 2003; Faravelli 2005; Tatton-Brown 2005; Saugier-Veber 2007). The large number of distinct mutations and the low number of recurrent mutations might be attributed to the lack of mutational hot spots in the *NSD1* gene (Faravelli 2005; Saugier-Veber 2007). Also, the vast majority of intragenic *NSD1* mutations occur *de novo*, and only a few familial cases have been reported (Kurotaki 2003; Tatton-Brown 2005). Two types of intragenic *NSD1* mutations, namely truncating and missense mutations, have been identified both in Japanese and non-Japanese SoS individuals (Kurotaki 2003; Faravelli 2005; Tatton-Brown 2005; Saugier-Veber 2007). Interestingly, truncating mutations are detected throughout the *NSD1* gene, while missense mutations are identified only in the functional domains of the *NSD1* gene (Faravelli 2005; Tatton-Brown 2005; Tatton-Brown 2007).

Partial *NSD1* deletions are identified in approximately 5% of patients with SoS (Douglas 2005). Partial *NSD1* deletions have unique breakpoints and occur throughout the gene. However, an increased tendency towards deletions spanning exon 1 and 2 of the *NSD1* gene is observed (Douglas

2005). This is most likely due to the high density of Alu repeats flanking these exons that mediate non-allelic homologous recombination (NAHR). In contrast, partial deletions that do not have breakpoints within Alu repeats seem to be mediated by non-homologous end joining (NHEJ) (Douglas 2005).

Interestingly, the type of chromosomal rearrangement and the size of 5q35 microdeletions differ between SoS individuals of different ethnic origin. In particular, a 1.9 Mb microdeletion spanning the 5q35.2–q35.3 region (**Figure 4.1**) has been identified in the majority of Japanese SoS cases (Kurotaki 2003; Visser 2005; Saugier-Veber 2007; Tatton-Brown 2007). This deletion results from intrachromosomal rearrangements, and is mediated through NAHR between low-copy repeats (LCRs) flanking the *NSD1* gene (**Figure 4.1**) (Kurotaki 2003; Tatton-Brown 2005). In contrast, most of the 5q35 microdeletions identified in non-Japanese SoS cases have unique breakpoints, and thus differ in size, and most likely result from interchromosomal rearrangements (Tatton-Brown 2005). Notably, in both Japanese and non-Japanese SoS cases, a significant bias towards microdeletions of the paternal 5q35 region is observed. However, this might be at least partially attributed to the increased rate of recombination at the 5q telomere in males compared to females (Tatton-Brown 2005).

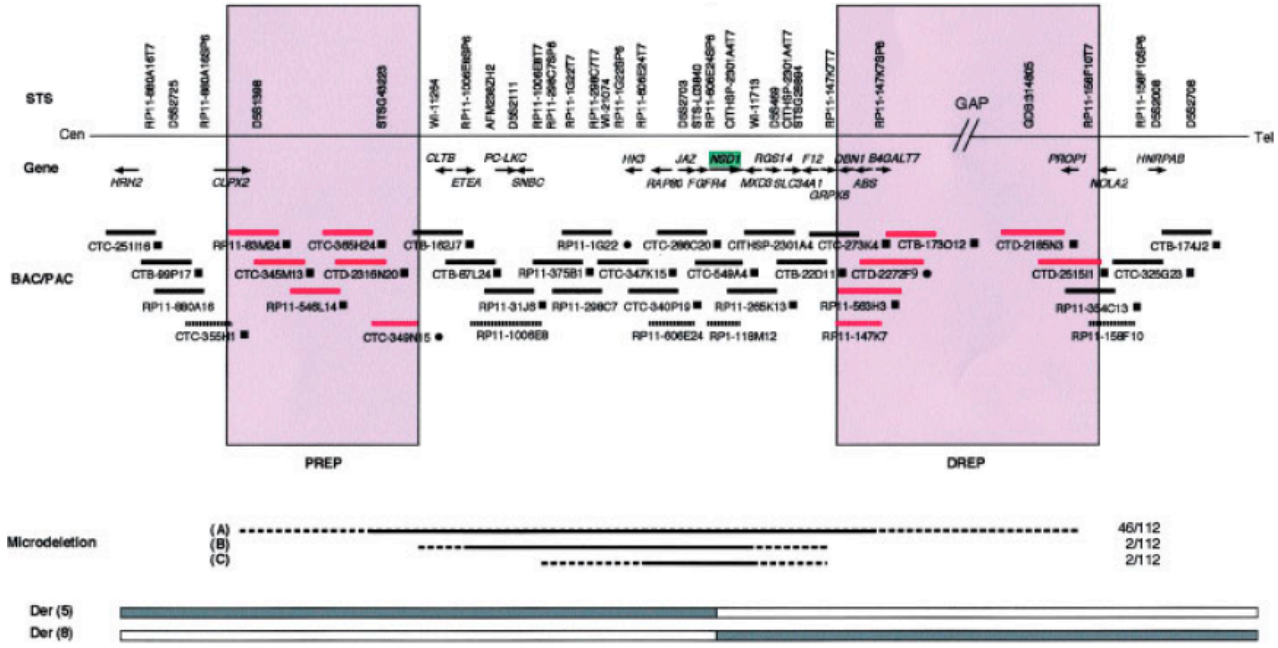

Figure 4.1. Physical map covering microdeletions and their flanking regions, and a summary of 50 microdeletions identified by FISH or microarray CGH analysis. Thirty-seven BAC/PAC clones were incorporated in this map. Small horizontal arrows show the orientation of a gene. Microdeletions A, B, and C are found in 46, 2, and 2 of 112 SoS cases respectively. Long horizontal lines indicate confirmed (solid lines) and possibly deleted (dotted lines) regions respectively. Purple boxes indicate regions of LCRs (low copy repeats), PREP (proximal-repeat) and DREP (distal-repeat). Figure taken from Kurotaki *et al*., 2003.

4.2.3 PHENOTYPIC DIFFERENCES BETWEEN SOTOS SYNDROME CASES CARRYING INTRAGENIC MUTATIONS AND 5q35 MICRODELETIONS

The major clinical symptoms of SoS, including pre- and postnatal overgrowth, facial abnormalities, hypotonia, and intellectual disability, are diagnosed both in SoS patients carrying intragenic *NSD1* mutations and in SoS individuals carrying 5q35 microdeletions (Tatton-Brown 2005). However, in contrast to SoS individuals with intragenic *NSD1*, SoS patients with the 5q35 microdeletions tend to show less pronounced overgrowth, yet more profound intellectual disability (Kurotaki 2003; Nagai 2003; Tatton-Brown 2005; Saugier-Veber 2007). Moreover, several studies have reported increased frequency of cardiovascular and urinary/renal abnormalities in SoS patients carrying 5q35 microdeletions (Kurotaki 2003; Nagai 2003; Saugier-Veber 2007). Thus it is possible that genes other than the *NSD1* gene

could be dosage-sensitive, and therefore responsible for the extended variability and degree of severity of phenotypes observed in SoS patients with the 5q35 microdeletion.

4.2.4 THE *NSD1* GENE

The *NSD1* gene maps to the 5q35 region, consists of 23 exons and is expressed in a variety of tissues, including human foetal and adult brain, skeletal muscle, kidney, spleen, thymus, and lung, as well as adult peripheral blood leukocytes and placenta (Kurotaki 2001).

NSD1 encodes a 2,696 amino acid protein that belongs to an emerging family of proteins that includes NSD2 and NSD3 (Kurotaki 2001). NSD1 contains multiple functional domains, including a Su(var)3-9, Enhancer-of-Zeste, Trithorax (SET) domain, SET-associated cysteine-rich (SAC) domain, five zinc-finger plant homeodomains (PHDs), two proline-tryptophan-tryptophan-proline (PWWP) domains, and two distinct nuclear receptor (NR) interaction domains (NID^{-L} and NID^{+L}) (**Figure 4.2**) (Kurotaki 2001; Tatton-Brown 2007). SET domains are found in proteins that regulate gene transcription, play a role in position-effect-variegation (PEV), and are involved in the establishment and maintenance of chromatin structure (Jenuwein 1998; Kurotaki 2001), while SAC domains are present in proteins that function as histone-methyltransferases (HMTases) (Huang 1998). Interestingly, both SET and SAC domains have been shown to be involved in methylation of lysine 36 of histone 3 and lysine 20 of histone 4 (Tatton-Brown 2007). Also, PHD domains are found in proteins that interact with chromatin and are involved in chromatin-mediated regulation of gene transcription and modulation of chromatin structure (Huang 1998). Thus, altogether, the presence of SAC, SET and PHD domains in the NSD1 protein suggests an involvement of the *NSD1* gene in modulation of chromatin structure and regulation of gene transcription. The function of PWWP domains has not been fully elucidated. However, they are believed to be involved in protein-protein interaction (Stec 2000). Interestingly, NID^{-L} and NID^{+L} domains possess binding properties of NR interaction domains found in NR corepressors and coactivators that suggests that the NSD1 protein functions as a bifunctional

transcriptional co-factor that can either activate or repress transcription in response to ligand binding (Huang 1998), which further supports a putative role of the *NSD1* gene in regulation of gene transcription.

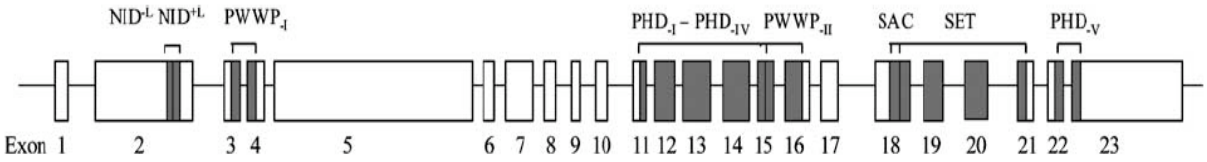

Figure 4.2. Schematic representation of the *NSD1* gene. Figure taken from Tatton-Brown *et al*., 2007. NID^{-L} and NID^{+L} - nuclear receptor (NR) interaction domains; $PWWP_{-I}$ and $PWWP_{-II}$ - proline-tryptophan-tryptophan-proline domains; PHD_{-I} - PHD_{-V} - zinc-finger plant homeodomains; SAC - SET-associated cysteine-rich domain; SET - Su(var)3-9, Enhancer-of-Zeste, Trithorax domain.

4.2.5 A MOUSE MODEL OF SOTOS SYNDROME

The mouse orthologue of the *NSD1* gene was identified in 1998 during a search for factors mediating transcriptional responses induced by binding of ligands to nuclear receptors (Huang *et al*., 1998). The murine *Nsd1* gene is highly conserved with the human *NSD1* gene, showing 86% homology at the nucleotide level and 83% homology at amino acid level (Kurotaki 2001).

In 2003 a mouse knock-out of the *Nsd1* gene was developed by gene disruption (Vani Rayasam 2003). The targeted intragenic deletion of exon 2 of the *Nsd1* gene resulted in a frameshift mutation with a premature stop codon in exon 3. The truncated NSD1 protein was devoid of NID^{-L} and NID^{+L} and all conserved domains. Heterozygous *Nsd1* mice were viable, fertile and did not display any gross phenotypic abnormalities. Most notably, heterozygous *Nsd1* mice showed a normal growth rate and did not display any signs of overgrowth, and thus did not recapitulate the major clinical feature of SoS patients. However, vani Rayasam *et al* speculated that if the overgrowth in the heterozygous *Nsd1* mice is less pronounced than in human SoS individuals, this phenotype might have been missed. Alternatively, analysis of the mice at earlier timepoints might have shown an overgrowth phenotype. No homozygous *Nsd1* mice were present among progeny from heterozygote

Nsd1 intercrosses, showing that the *Nsd1* mutation resulted in prenatal death. Collection of embryos from heterozygote *Nsd1* intercrosses at different timepoints during gestation revealed the absence of homozygous *Nsd1* embryos at or after E10.5, which indicates that the *Nsd1* gene is essential for early post-implantation development. Further analysis showed that homozygous *Nsd1* embryos displayed increased apoptosis in the embryonic ectodermal cells, were able to initiate mesodermal formation, but failed to complete gastrulation (presumably due to abnormal ectodermal development). Moreover, vani Rayasam *et al* showed that the Nsd1 protein has an intrinsic histone methylotransferase activity that was capable of methylating lysine 36 of histone 3 and lysine 20 of histone 4, which provided further evidence supporting the involvement of the *Nsd1* gene in modulation of chromatin structure and regulation of gene transcription.

Interestingly, none of the other mouse models carrying knock-out mutations of individual genes located within the most frequently detected human SoS 5q35 microdeletion have shown a dominant phenotype, suggesting that *Sncb*, *Unc5a*, *Fgfr4*, *Mxd3*, *Rgs14*, *Slc34a1*, *F12*, *Grk6*, *Dbn1* and *Dok3* are not dosage-sensitive, and therefore, at least individually, are not likely to be responsible for the extended variability and degree of severity of clinical features observed in SoS patients with the 5q35 microdeletions (Mouse Genome Informatics - http://www.informatics.jax.org).

4.2.6 GENERATION OF A NEW MOUSE MODEL OF SOTOS SYNDROME

To our knowledge only one mouse model of SoS has been generated to date (Vani Rayasam 2003). This mouse model provides phenotypic data only on the deletion of the *Nsd1* gene, and thus the contribution of other genes within the SoS deletion to the phenotype remains unclear. To this end, we used chromosome engineering to generate a new mouse model of SoS, $Df^{4732471D19Rik\text{-}B4galt7}$, carrying a deletion syntenic to the human region 5q35.2–q35.3, in order to allow investigation of the contribution of the 36 genes at the distal end of human chromosome 5 to the development of clinical features commonly diagnosed in SoS patients.

4.3 RESULTS

4.3.1 GENERATION OF MONOSOMIC MICE FOR THE 1.5 Mb *4732471D19Rik–B4galt7* REGION

The *4732471D19Rik* and *B4galt7* genes are located at the proximal and distal ends, respectively, of a 1.5 Mb region in the B1 band of mouse chromosome 13 (MMU13), which is syntenic to the human region 5q35.2–q35.3 (**Figure 4.3**). This region on human chromosome 5 (HSA5) contains 36 genes, and all of these genes have orthologous counterparts in the syntenic region of MMU13 (**Table 4.1**). The *4732471D19Rik–B4galt7* deletion was generated using chromosomal engineering (Zheng 1999) as described in Material and Methods (**Figure 4.4A**). Briefly, AB2.2 ES cells were sequentially electroporated with targeting vectors containing a portion of the *Hprt* selection cassette (5' or 3'*Hprt*), a *loxP* site and a coat colour marker (agouti or tyrosinase). The targeting vector containing the 5' *Hprt* cassette (MICER clone: MHPN55m07) (Adams 2004) was inserted proximal to *4732471D19Rik* and the targeting vector containing the 3' *Hprt* cassette (MICER clone: MHPP265c24) (Adams 2004) was inserted distal to *B4galt7* (**Figure 4.4A**). The correct insertion of both targeting vectors was confirmed by Southern blot analysis on *Bst*EII- or *Spe*I-digested genomic DNA extracted from ES clones, selected either in G418 or puromycin using a 5′ and 3′ Southern external probe respectively (**Figure 4.4A and 4.4B**). Double-targeted clones, in which both targeting vectors were inserted on the same chromosome (*cis*), were electroporated with a Cre-expression vector, and subsequently selected in medium containing HAT to isolate ES clones carrying a chromosomal deletion generated via recombination of the *loxP* sites (**Figure 4.4A**). The deletion allele was designated $Df^{4732471D19Rik\text{-}B4galt7}$ (alternatively named *Ms(4732471D19Rik-B4galt7)2Dja* and abbreviated as *Ms2Dja*). The presence of the deletion in *Hprt*-resistant ES clones was confirmed by FISH (**Figure 4.4C**). The positive ES clones were used to generate chimaeras, which transmitted $Df^{4732471D19Rik\text{-}B4galt7}$ to their progeny.

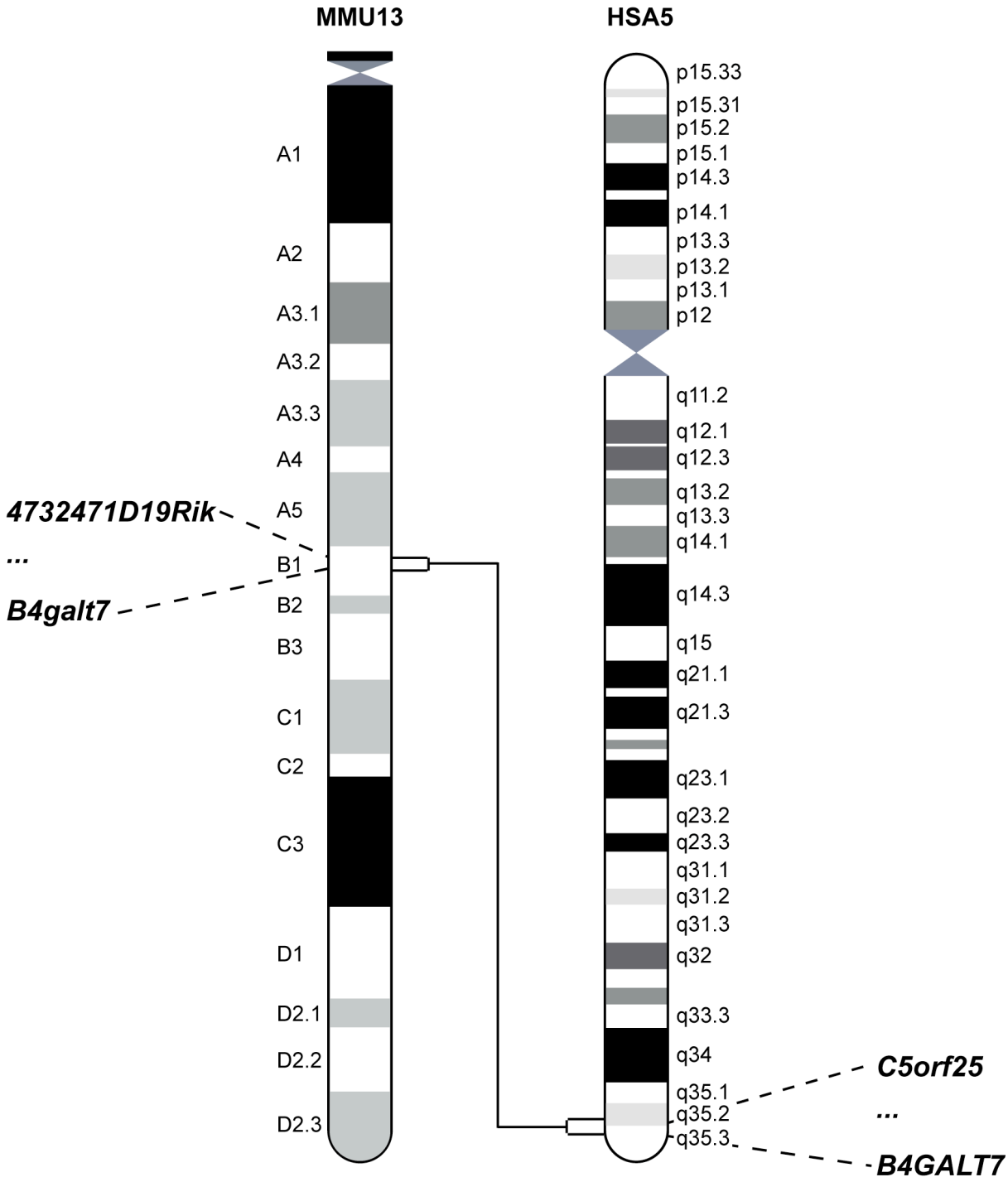

Figure 4.3. Schematic representation of the q35.2–q35.3 interval on HSA5 and the syntenic region in the B1 band on MMU13. Proximal and distal genes that map to the human 5q35.2–q35.3 region (NCBI build GRChr37) and the B1 band on MMU13 (NCBI build m37) have been listed.

Table 4.1. Summary of the genes mapped within the *4732471D19Rik–B4galt7* region (NCBI build m37). Functions/processes were gene-based on ontology classifications as listed by Mouse Genome Informatics (http://www.informatics.jax.org/).

Gene	Official name	Ensembl ID (v61)	Function/Processes (GO terms)
4732471D19Rik	RIKEN cDNA 4732471D19 gene	ENSMUSG00000043183	Unknown
4833439L19Rik	RIKEN cDNA 4833439L19 gene	ENSMUSG00000025871	Unknown
Arl10	ADP-ribosylation factor-like 10	ENSMUSG00000025870	GTP and nucleotide binding activity
Nop16	NOP16 nucleolar protein homolog (yeast)	ENSMUSG00000025869	Unknown
Higd2a	HIG1 domain family, member 2A	ENSMUSG00000025868	Unknown
Cltb	Clathrin, light polypeptide (Lcb)	ENSMUSG00000045547	Intracellular protein transport, vesicle-mediated transport, and peptide binding activity
Faf2	Fas associated factor family member 2	ENSMUSG00000025873	Response to unfolded proteins, and ubiquitin and ubiquitin protein ligase binding
Rnf44	Ring finger protein 44	ENSMUSG00000034928	Metal ion binding activity
Cdhr2	Cadherin-related family member 2	ENSMUSG00000034918	Cell adhesion, and negative regulation of cell growth
Gprin1	G protein-regulated inducer of neurite outgrowth 1	ENSMUSG00000069227	Neuron projection development, and phospoprotein binding activity
Sncb	Synuclein, beta	ENSMUSG00000034891	Dopamine metabolic processes, negative regulation of neuron apoptosis, and calcium ion binding activity
Eif4e1b	Eukaryotic translation initiation factor 4E family member 1B	ENSMUSG00000074895	Translational, and RNA binding and translation initiation factor activity
Tspan17	Tetraspanin 17	ENSMUSG00000025875	Unknown
Unc5a	Unc-5 homolog A (*C.elegans*)	ENSMUSG00000025876	Apoptosis, axon guidance, multicellular organismal development, and netrin receptor activity
Hk3	Hexokinase 3	ENSMUSG00000025877	Carbohydrate metabolic processes, phosphorylation, glycolysis, kinase and transferase activity, and ATP, kinase, hormone, enzyme and hexokinase binding activity
Uimc1	Ubiquitin interaction motif containing 1	ENSMUSG00000025878	Chromatin modifications, DNA repair, DNA-dependent, negative regulation of transcription, histone H2A K63-linked deubiquitination, and histone and K63-linked polyubiquitin binding activity
Zfp346	Zinc finger protein 346	ENSMUSG00000021481	Apoptosis, and metal ion and double-stranded RNA binding activity
Fgfr4	Fibroblast growth factor receptor 4	ENSMUSG00000005320	Alveolar secondary septum development, fibroblast growth factor receptor signalling pathway, protein phosphorylation, organ induction, lung development, and ATP, kinase and transferase binding activity
Nsd1	Nuclear receptor-binding SET-domain protein 1	ENSMUSG00000021488	Gastrulation with mouth forming second, histone H3-K36 methylation, and androgen receptor, zinc ion, transcription cofactor and chromatin binding activity

Table 4.1 continued. Summary of the genes mapped within the *4732471D19Rik–B4galt7* region (NCBI build m37).

Gene	Official name	Ensembl ID (v61)	Function/Processes (GO terms)
Rab24	RAB24, member RAS oncogene family	ENSMUSG00000034789	Autophagy, protein transport, small GTPase mediated signal transduction, and GTP and nucleotide binding activity
Mxd3	Max dimerization protein 3	ENSMUSG00000021485	Negative regulation of transcription, DNA-dependent regulation of transcription, and protein binding activity
Prelid1	PRELI domain containing 1	ENSMUSG00000021486	Unknown
Lman2	Lectin, mannose-binding 2	ENSMUSG00000021484	Protein transport, and metal ion and sugar binding
Rgs14	Regulator of G-protein signalling 14	ENSMUSG00000052087	Cell cycle and division, spindle organization, cell division, and GDP dissociation inhibitor, microtubule binding activity, and GTPase activator activity, and positive regulation of GTPase activity
Slc34a1	Solute carrier family 34 (sodium phosphate), member 1	ENSMUSG00000021490	Ion transport, bone remodelling, and sodium-dependent phosphate transmembrane transporter activity
Pfn3	Profilin 3	ENSMUSG00000044444	Actin cytoskeleton organization, and actin binding and ATP:ADP antiporter activity
F12	Coagulation factor XII (Hageman factor)	ENSMUSG00000021492	Plasma kallikrein-kinin cascade, Factor XII activation, positive regulation of plasminogen activation, protein autoprocessing, positive regulation of blood coagulation, zymogen activation, protein maturation by peptide bond cleavage, response to misfolded protein, and positive regulation of fibrinolysis
Grk6	G protein-coupled receptor kinase 6	ENSMUSG00000074886	Desensitization of G-protein coupled receptor protein signalling pathway, protein phosphorylation, and ATP binding, and transferase and protein serine/threonine kinase activity
Prr7	Proline rich 7 (synaptic)	ENSMUSG00000034686	Unknown
Dbn1	Drebrin 1	ENSMUSG00000034675	Actin filament organization, signal transduction, multicellular organismal development, nervous system development, cell differentiation, and actin and profiling binding activity
Pdlim7	PDZ and LIM domain 7	ENSMUSG00000021493	Actin cytoskeleton organization, ossification, multicellular organismal development, cell differentiation, and metal ion binding activity
Dok3	Docking protein 3	ENSMUSG00000035711	Ras protein signal transduction, and insulin receptor binding activity
Ddx41	DEAD (Asp-Glu-Ala-Asp) box polypeptide 41	ENSMUSG00000021494	mRNA processing, RNA splicing, and metal ion binding, ATP binding and helicase activity
Fam193b	Family with sequence similarity 193, member B	ENSMUSG00000021495	Unknown
Tmed9	Transmembrane emp24 protein transport domain containing 9	ENSMUSG00000058569	Transport
B4galt7	Xylosylprotein beta1,4-galactosyltransferase, polypeptide 7 (galactosyltransferase I)	ENSMUSG00000021504	Carbohydrate and proteoglycan metabolic process, glycosaminoglycan biosynthetic process, fibril organization, negative regulation of fibroblast proliferation, and metal ion binding, transferase and transferring glycosyl groups activity

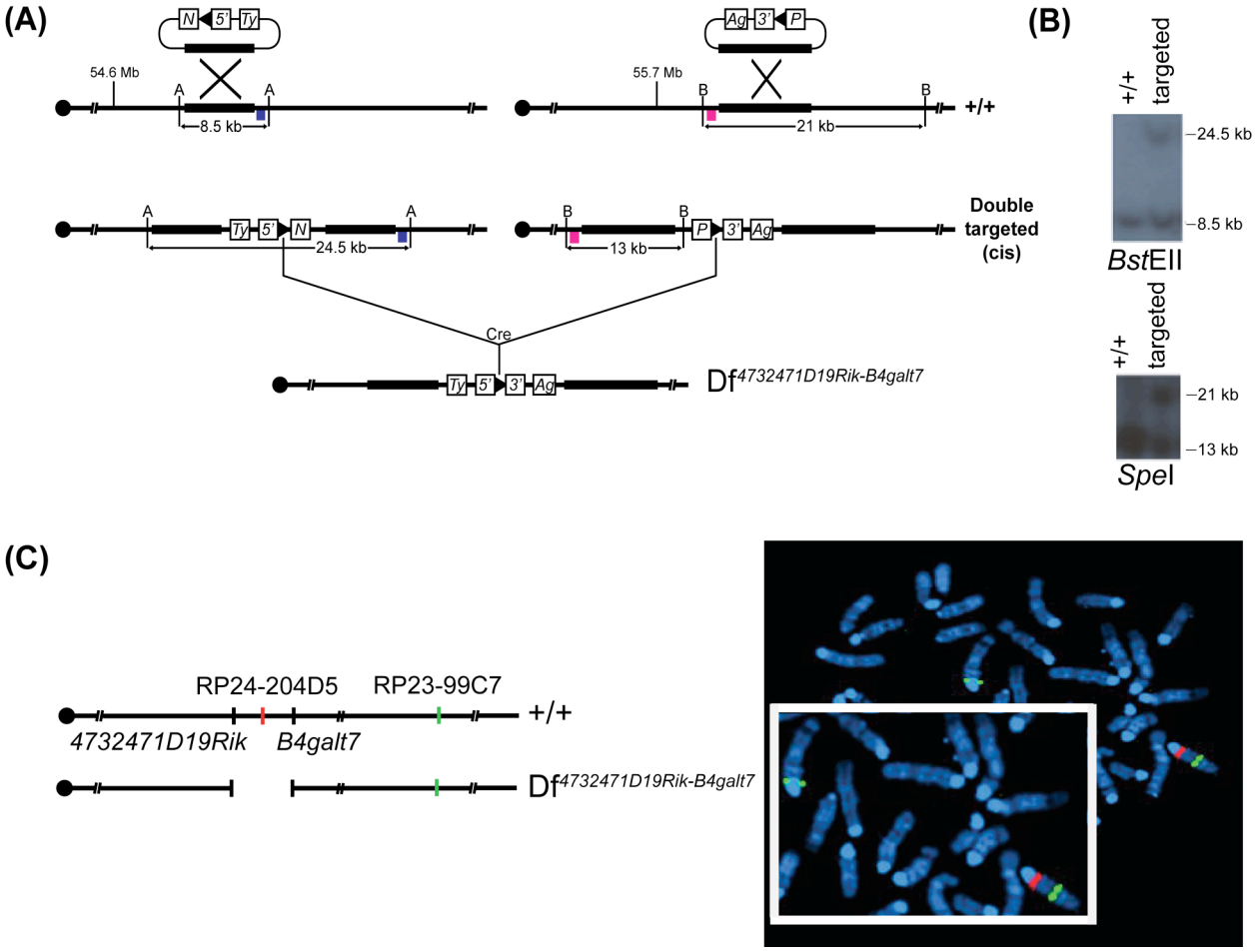

Figure 4.4. Generation of a 1.5 Mb deletion between the *4732471D91Rik* and *B4galt7* loci using Cre/*loxP*-mediated chromosomal engineering. (A) Strategy to generate the chromosomal rearrangement (using NCBI build GRChr37). The targeting vectors containing a *loxP* site (arrowhead), a selectable antibiotic resistance gene (*N* or *P*), a coat colour marker (*Ty* or *Ag*) and part of the *Hprt* gene (5′ or 3′) were integrated successively in the *Lipi* locus and the *Usp25* locus. The coloured boxes (blue and pink) indicate the location of the probes (5’ and 3’, respectively) used for Southern blotting. A, *Bst*EII; B, *Spe*I; 5’, 5’*Hprt*; 3’, 3’*Hprt*; *N*, neomycin-resistance gene; *P*, puromycin-resistance gene; *Ty*, tyrosinase minigene, *Ag*, K-14 agouti gene. **(B)** The targeting events were checked by Southern analysis showing an additional *Bst*EII fragment of 24.5 kb compared with the wildtype allele (8.5 kb) for the *4732471D19Rik* locus and an additional *Spe*I fragment of 13 kb compared with the wildtype allele (21 kb) for the *B4galt7* locus. **(C)** Interphase FISH analysis with BAC probes that map in the region of the deletion (red) and outside (green). Chromosomes from the ES cells double-targeted in *cis* ($Df^{4732471D19Rik\text{-}B4galt7}$) showed two green and only one red signal due to the deletion of the *4732471D19Rik–B4galt7* region, while chromosomes from the wildtype ES cells showed two green and two red signals.

4.3.2 PHENOTYPIC ANALYSIS OF MONOSOMIC MICE

The heterozygous (monosomic) *Df*$^{4732471D19Rik\text{-}B4galt7}$ mice were viable, fertile and did not show any gross abnormalities. In particular, no facial abnormalities or hypotonia were observed.

4.3.2.1 ANALYSIS OF BODY WEIGHT AND BODY COMPOSITION OF MONOSOMIC MICE REVEALS THAT LOSS OF ONE COPY OF THE *4732471D19Rik–B4galt7* REGION DOES NOT RESULT IN OVERGROWTH

To determine whether overgrowth in SoS patients could be observed in monosomic *Df*$^{4732471D19Rik\text{-}B4galt7}$ mice, the average weights of 8 monosomic mice and 8 wildtype littermates (controls) fed on a normal-fat diet were compared at 8 weeks of age. No statistical differences were observed between the analysed groups (**Figure 4.5**), suggesting that the overgrowth phenotype is either not recapitulated in mice (as overgrowth was also not observed in heterozygous *Nsd1* mice (Vani Rayasam 2003)) or recovered by that stage of development (as it has been shown that the overgrowth decreases with age and becomes much less apparent in SoS children after the first few years of life (Cole 1994; Tatton-Brown 2007) or perhaps overgrowth could only be observed during the gestation period, as has been shown in the *Cdkn1c* mouse model of Beckwith-Wiedemann syndrome (Tunster 2011)). In order to examine if the overgrowth is displayed in our monosomic mice during an earlier stage of pre- and/or postnatal development, I am currently in the process of analysing embryos at E15.5 and E18.5, and newborn pups at 1 day of age (P1).

8 monosomic and 8 control mice fed on a normal-fat diet were also subjected to the dual-energy X-ray absorptiometry (DEXA) analysis when they were 8 weeks of age in order to check whether any differences in body composition could be observed. However, no statistical differences in estimated total tissue mass, lean mass or fat mass were found between the two groups (**Figure 4.6A-C**). Also, no differences in bone mineral density were identified between monosomic mice and wildtype littermates (**Figure 4.7**).

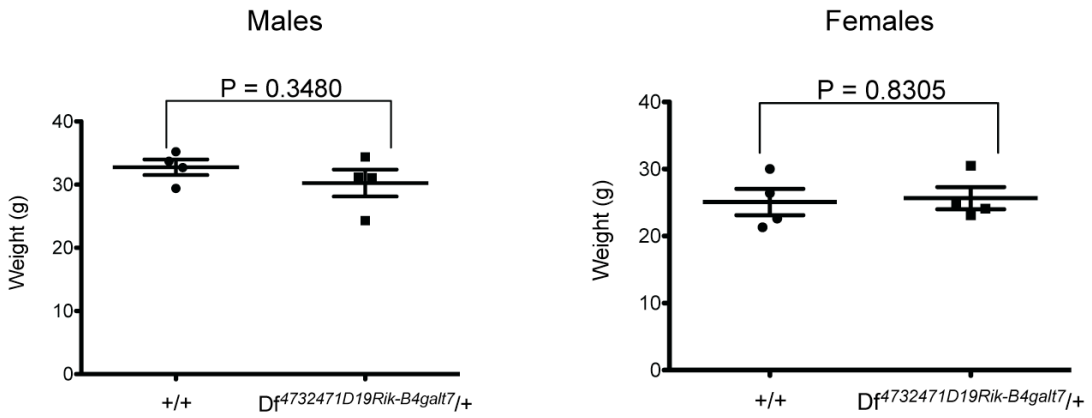

Figure 4.5. Body weight analysis of 8-week old control (+/+) and monosomic (*Df*$^{4732471D19Rik\text{-}B4galt7}$/+) littermates fed a normal-fat diet. Body weight results showing body weight in male and female littermates. Data was statistically analysed using the two-tailed Student's *t*-test. Four mice of each sex were analysed per genotype.

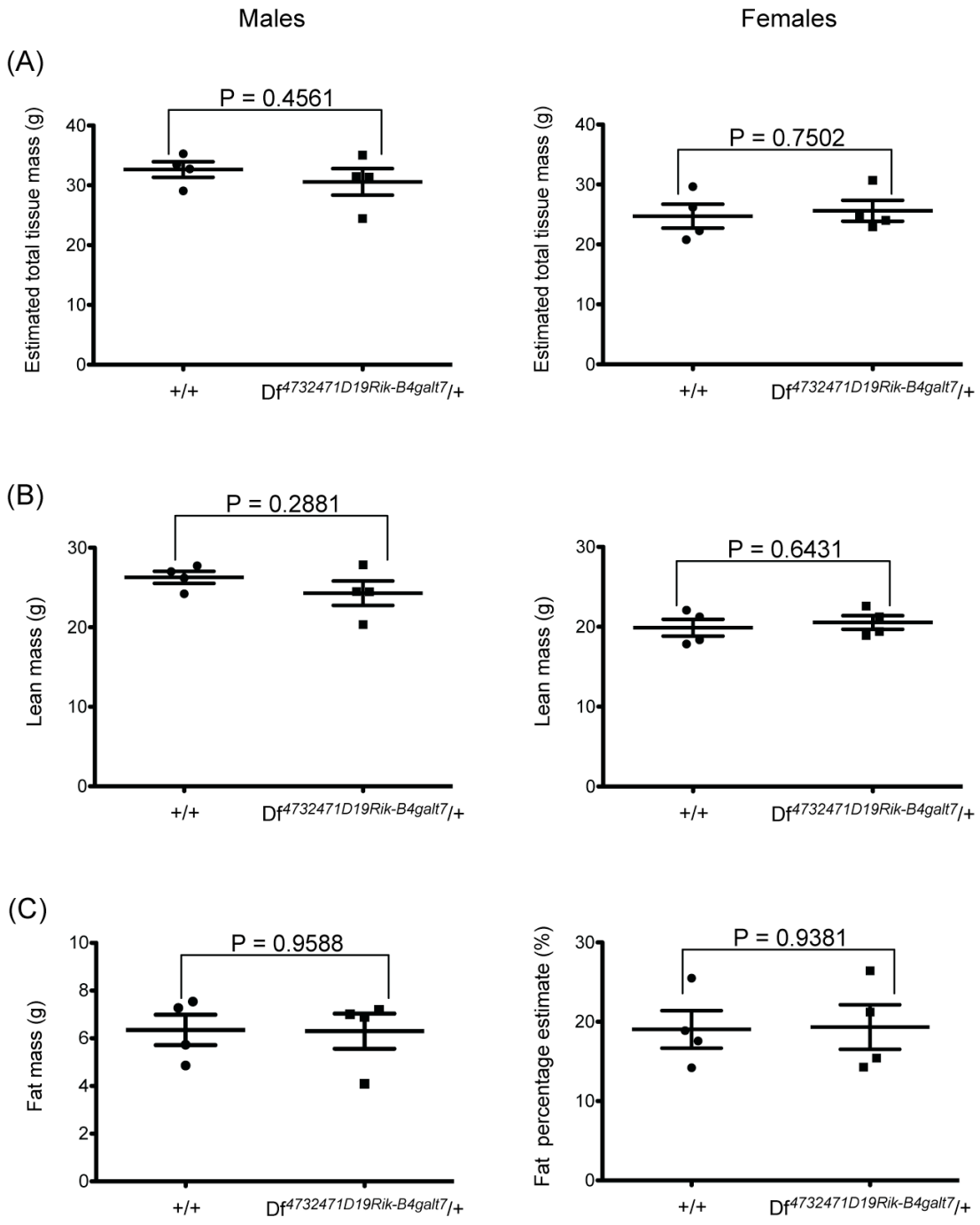

Figure 4.6. DEXA analysis of 8-week old control (+/+) and monosomic (*Df*$^{4732471D19Rik\text{-}B4galt7}$/+) littermates fed a normal-fat diet. (A) DEXA results showing estimated total tissue mass in male and female littermates. **(B)** DEXA results showing lean mass in male and female littermates. **(C)** DEXA results showing fat mass in male and female littermates. Data was statistically analysed using the two-tailed Student's *t*-test. Four mice of each sex were analysed per genotype.

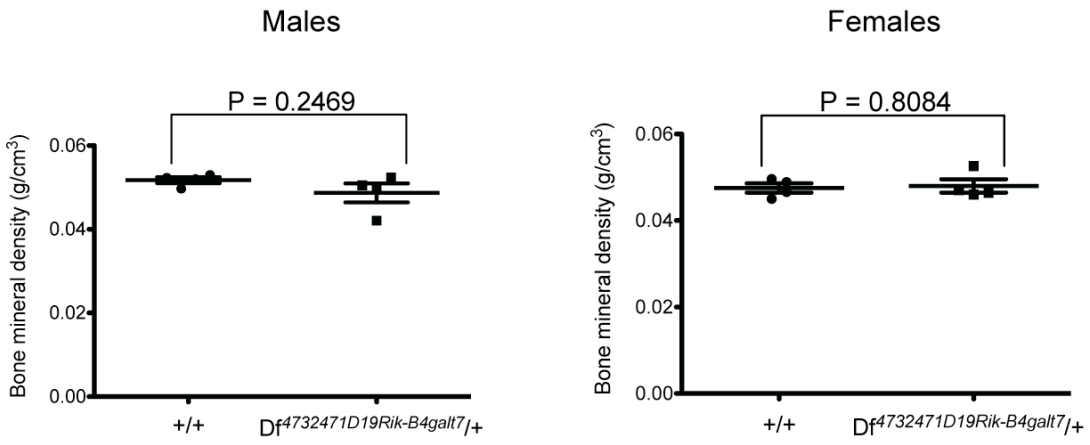

Figure 4.7. DEXA analysis of 8-week old control (+/+) and monosomic (*Df*$^{4732471D19Rik\text{-}B4galt7}$/+) littermates fed a normal-fat diet. DEXA results showing bone mineral density in male and female littermates. Data was statistically analysed using the two-tailed Student's *t*-test. Four mice of each sex were analysed per genotype.

4.3.3.2 HISTOPATHOLOGICAL ANALYSIS OF DIFFERENT TISSUES OF MONOSOMIC MICE REVEALS THAT LOSS OF ONE COPY OF THE *4732471D19Rik–B4galt7* REGION AFFECTS THE URINARY/RENAL SYSTEM

To determine whether any histopathological changes, in particular those associated with features found in SoS patients (such as advanced bone age, cardiovascular and/or urinary/renal abnormalities), could be observed in the monosomic *Df*$^{4732471D19Rik\text{-}B4galt7}$ mice, different tissues were collected from 8 monosomic mice and 8 wildtype littermates (controls) fed on a normal-fat diet when they were 8 weeks of age. The examination of a variety of tissues stained with haematoxylin and eosin (H&E) did not reveal the presence of any anatomical abnormalities of any organs, except kidneys where the dilation of the pelvicalyceal system was detected in monosomic mice (**Figure 4.8, 4.9A, 4.9B, 4.10A, 4.10B, 4.11**). Thus monosomic mice seem to recapitulate hydronephrosis, which is one of the renal abnormalities observed in patients with SoS. Interestingly, sex differences in the degree of severity of the dilation of the pelvicalyceal system were identified between monosomic males and females, as monosomic males showed mild/moderate (**Figure 4.9A and 4.9B**) to moderate/severe (**Figure 4.10A and 4.10B**) dilation of the pelvicalyceal

system while monosomic females showed only very mild dilation (**Figure 4.11**). At present, I am ageing a cohort of monosomic and control mice to investigate if observed kidney abnormalities will progress with age, and thus to determine whether dilation of the pelvicalyceal system is more pronounced in older monosomic mice. Also, we are planning to collect blood samples from those aged mice in order to analyse clinical chemistry parameters related to urinary/renal system functioning, including creatinine, urea and electrolyte.

Interestingly, histopathological analysis showed the presence of cancer (poorly differentiated sarcoma of the uterus) (**Figure 4.12**) in one of monosomic mice (1/8 mice; 12.5%). Although, the size of the analysed cohort was small, perhaps it is worth investigating whether the overall risk of tumourgenesis, currently estimated at less than 3% in individuals with Sotos syndrome (Tatton-Brown 2007), might be underestimated.

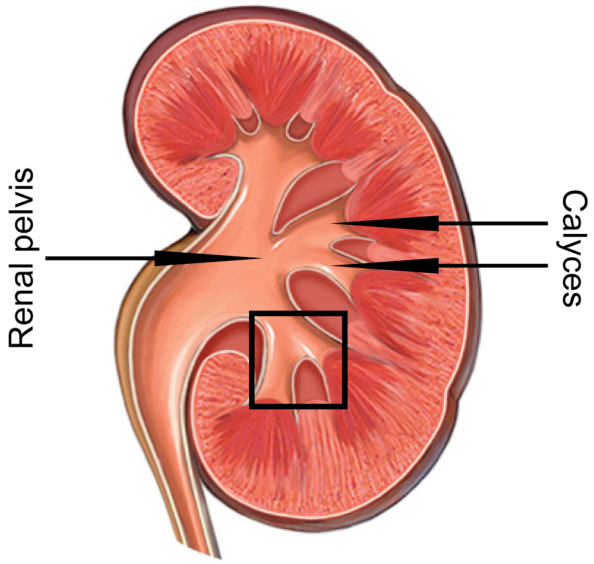

Figure 4.8. A schematic representation of a kidney. A black rectangle indicates a renal region where a dilation of the pelvicalyceal system was observed in monosomic (*Df*$^{4732471D19Rik\text{-}B4galt7}$/+) mice in all analysed kidney sections. Figure modified from www.empowher.com/files/ebsco/images/BI00031.jpg.

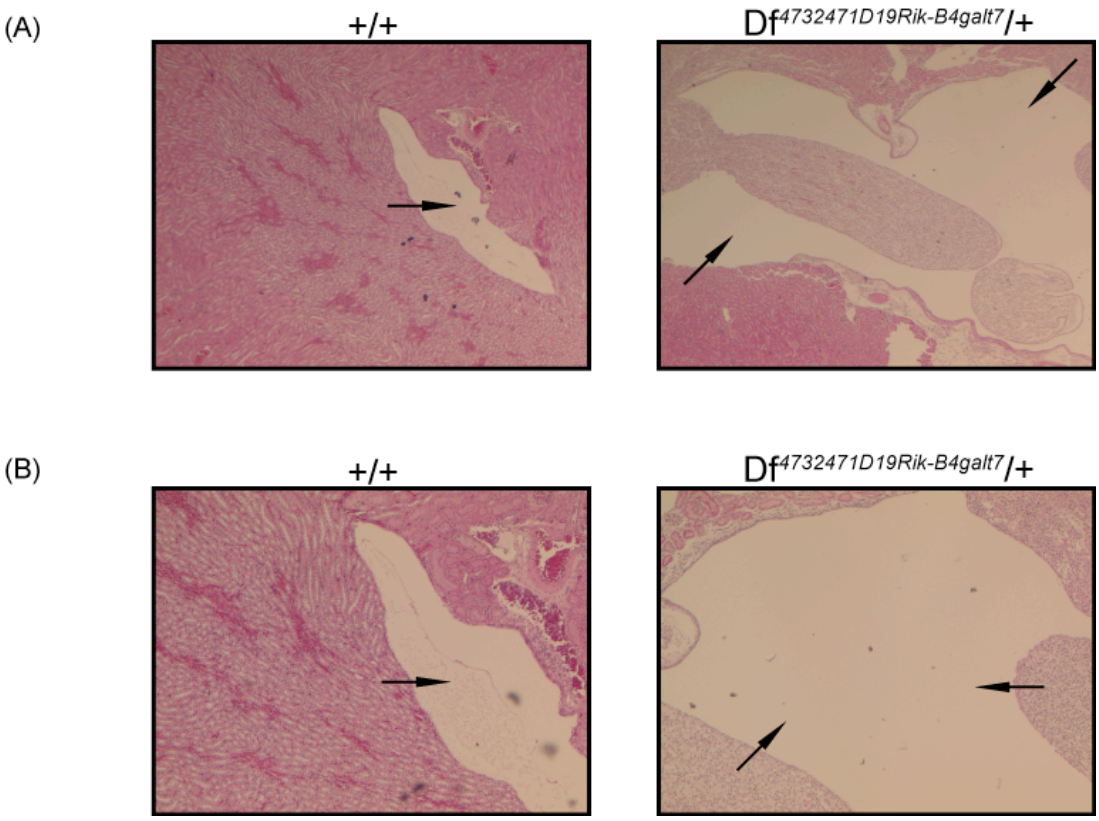

Figure 4.9. Histopathological analysis of kidneys collected from 8-week old control (+/+) and monosomic (*Df*$^{4732471D19Rik\text{-}B4galt7}$/+) male littermates fed on a normal-fat diet. (A), (B) Haematoxylin and eosin-stained kidney sections from control (+/+) and monosomic (*Df*$^{4732471D19Rik\text{-}B4galt7}$/+) males. Kidneys of monosomic (*Df*$^{4732471D19Rik\text{-}B4galt7}$/+) males showed mild/moderate dilation of the pelvicalyceal system (visible as empty spaces; indicated by arrows) compared to control (+/+) males (visible as compact structure of the pelvicalyceal system; indicated by arrows). Images are representative and taken at x25 magnification (A) and x50 magnification (B).

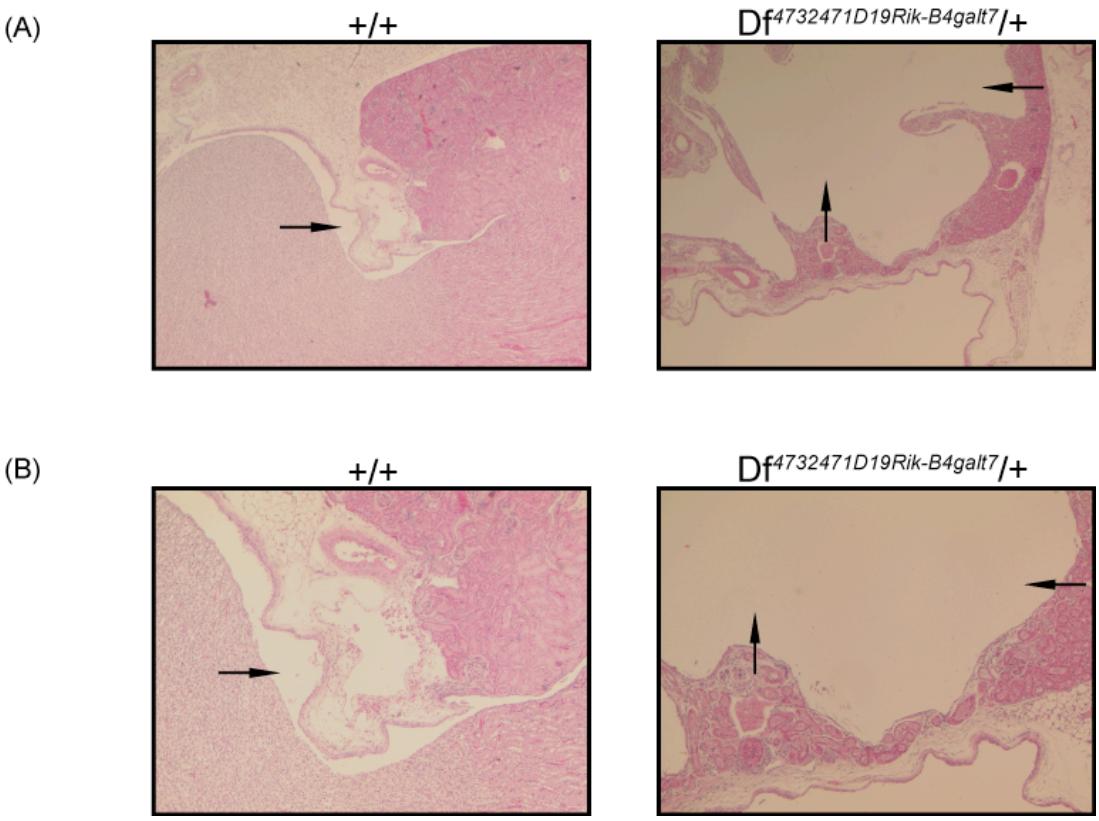

Figure 4.10. Histopathological analysis of kidneys collected from 8-week old control (+/+) and monosomic (*Df*$^{4732471D19Rik\text{-}B4galt7}$/+) male littermates fed on a normal-fat diet. (A), (B) Haematoxylin and eosin-stained kidney sections from control (+/+) and monosomic (*Df*$^{4732471D19Rik\text{-}B4galt7}$/+) males. Kidneys of monosomic (*Df*$^{4732471D19Rik\text{-}B4galt7}$/+) males showed moderate/severe dilation of the pelvicalyceal system (visible as empty spaces; indicated by arrows) compared to control (+/+) males (visible asc compact structure of the pelvicalyceal system; indicated by arrows). Images are representative and taken at x25 magnification (A) and x50 magnification (B).

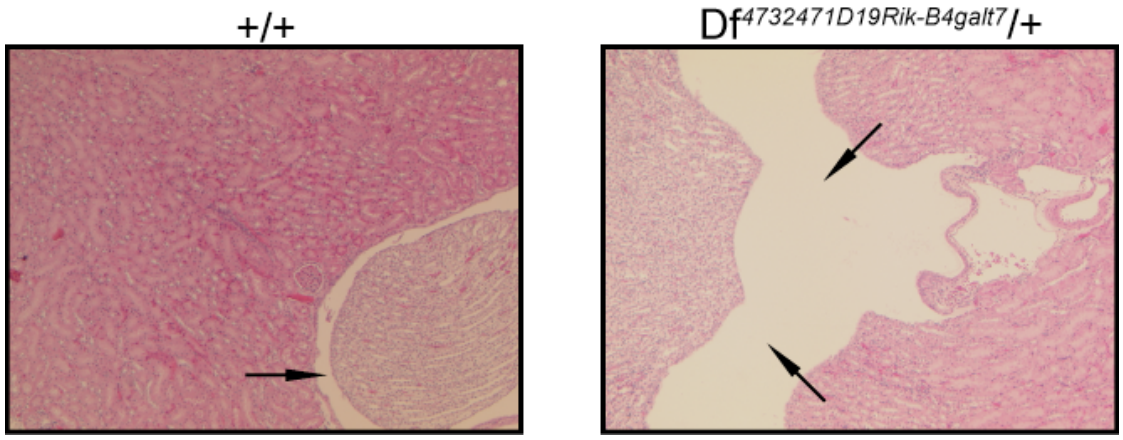

Figure 4.11. Histopathological analysis of kidneys collected from 8-week old control (+/+) and monosomic (*Df*$^{4732471D19Rik\text{-}B4galt7}$/+) female littermates fed on a normal-fat diet. Haematoxylin and eosin-stained kidney sections from control (+/+) and monosomic (*Df*$^{4732471D19Rik\text{-}B4galt7}$/+) females. Kidneys of monosomic (*Df*$^{4732471D19Rik\text{-}B4galt7}$/+) females showed very mild/mild dilation of the pelvicalyceal system (visible as empty spaces; indicated by arrows) compared to control (+/+) females (visible as compact structure of the pelvicalyceal system; indicated by arrows). Images are representative and taken at x50 magnification.

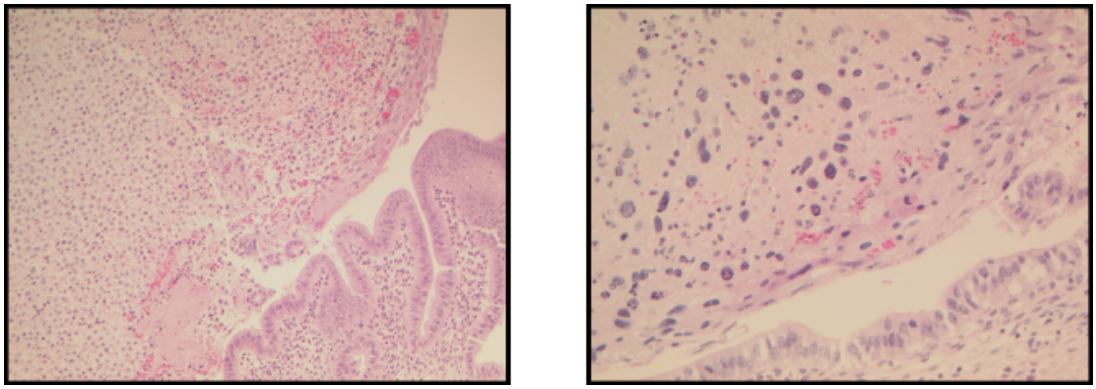

Figure 4.12. Analysis of uterus collected from a *Df*$^{4732471D19Rik\text{-}B4galt7}$/+ mouse. (A) Histopathological analysis of haematoxylin and eosin stained uterus section showing sarcoma. Images are representative and taken at x25 magnification (left) and x200 magnification (right).

4.4 DISCUSSION

We generated and phenotypically characterized the monosomic *Df*$^{4732471D19Rik\text{-}B4galt7}$ mouse, a model for the 1.5 Mb *4732471D19Rik–B4galt7* region. The deleted region is syntenic to the telomeric part of the human chromosome 5, and contains 36 genes that are conserved between human and mouse.

Gross phenotypical analysis of heterozygous (monosomic) *Df*$^{4732471D19Rik\text{-}B4galt7}$ mice found them to be viable, fertile and indistinguishable from controls (wildtype littermates). Specifically, no postnatal overgrowth was observed in 8-week old monosomic *Df*$^{4732471D19Rik\text{-}B4galt7}$ mice, and so these mice do not seem to recapitulate a cardinal feature observed in both SoS patients carrying intragenic *NSD1* mutations and 5q35 microdeletions. Interestingly, there was also no postnatal overgrowth identified in heterozygous *Nsd1* mice (Vani Rayasam 2003), suggesting that the overgrowth phenotype might not be exhibited in mice. However, it is possible that overgrowth might be present during earlier stages of postnatal development, as it has been shown that the overgrowth decreases with age and becomes much less apparent in SoS children after the first few years of life (Cole 1994; Tatton-Brown 2007). It is also possible that the overgrowth might only be displayed during prenatal development, as in a mouse model of Beckwith-Wiedemann syndrome, in which the overgrowth was only observed during embryonic development (Tunster 2011). For example, overgrowth in a mouse model of Beckwith-Wiedemann syndrome was only observed during embryonic development (Tunster 2011). So only through careful analysis of the growth dynamics of monosomic *Df*$^{4732471D19Rik\text{-}B4galt7}$ mice during gestation and early postnatal period might we see the presence of overgrowth, particularly as the overgrowth decreases with age and becomes much less apparent in SoS children after the first few years of life (Cole 1994; Tatton-Brown 2007) and SoS patients with the 5q35 microdeletions tend to show less pronounced overgrowth compared to SoS patients carrying intragenic *NSD1* mutations (Tatton-Brown 2005).

Histopathological examination of different tissues collected from

monosomic *Df*$^{4732471D19Rik\text{-}B4galt7}$ mice revealed no signs of advanced bone age or cardiovascular anomalies. However, it is worth mentioning that advanced bone age and/or cardiovascular abnormalities are observed only in a subset of SoS patients carrying 5q35 microdeletions (Nagai 2003; Tatton-Brown 2005; Saugier-Veber 2007). Histopathological examination did, however, reveal kidney abnormalities in monosomic *Df*$^{4732471D19Rik\text{-}B4galt7}$ mice compared to wildtype littermates. More specifically, both monosomic males and females showed dilation of the pelvicalyceal system at 8 weeks of age. It has been documented that some SoS patients carrying 5q35 microdeletions are identified with anomalies in the urinary/renal system, including vesicoureteric reflux, hydronephrosis, and small kidneys (Kurotaki 2003; Nagai 2003; Tatton-Brown 2005; Saugier-Veber 2007). Hydronephrosis is characterized by dilation of the renal pelvis calyces, and thus monosomic *Df*$^{4732471D19Rik\text{-}B4galt7}$ mice fully recapitulate this anomaly of the urinary/renal system. Nevertheless, further analysis of older monosomic mice is required to establish if this dilation progresses with age, to investigate whether monosomic mice show increased risk of urinary tract infections as a consequence of blocking flow of urine, and to check whether destruction of nephrons eventually leads to renal failure. To date, interestingly, none of the mouse models carrying knock-out mutations in genes mapped within the 5q35 interval have shown a dominant phenotype, thus suggesting that *Sncb*, *Unc5a*, *Nsd1*, *Fgfr4*, *Mxd3*, *Rgs14*, *Slc34a1*, *F12*, *Grk6*, *Dbn1* and *Dok3* are not dosage-sensitive, and therefore are not likely to be responsible for the renal abnormality observed in monosomic *Df*$^{4732471D19Rik\text{-}B4galt7}$ mice (Mouse Genome Informatics - http://www.informatics.jax.org).

To conclude, we show here that haploinsufficiency of a gene (or genes) in the *4732471D19Rik–B4galt7* region of MMU13, syntenic to human 5q35.2–q35.3, results in the dilation of the renal pelvicalyceal system in monosomic mice. These partially monosomic mice have contributed new insights into the existence of a dosage-sensitive gene (or genes) involved in SoS-associated anomaly of the urinary/renal system and will play a critical role in unraveling of the molecular mechanisms underlying this phenotype. On the other hand, haploinsufficiency of genes in the *4732471D19Rik–B4galt7* region

does not seem to model the other clinical phenotypes, including overgrowth, advanced bone age, hypotonia, facial and cardiovascular abnormalities, which are commonly seen in SoS patients. However, the contribution of genes mapped within the deletion to the development of intellectual disability observed in SoS individuals cannot be excluded, as such a phenotype has not yet been analysed in monosomic *Df*$^{4732471D19Rik\text{-}B4galt7}$ mice.