OPEN ACCESS

Original research

The Phenotypic variability of 16p11.2 distal BP2–BP3 deletion in a transgenerational family and in neurodevelopmentally ascertained samples

Marc Woodbury-Smith ,[1,2] Lia D'Abate,[1,3] Dimitri J Stavropoulos,[1,4] Jennifer Howe,[1] Irene Drmic,[1,5] Ny Hoang,[3,6] Mehdi Zarrei,[1] Brett Trost ,[1] Alana Iaboni,[7] Evdokia Anagnostou,[7] Stephen W Scherer[1,3,8]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/jmg-2022-108818).

For numbered affiliations see end of article.

Correspondence to
Dr Marc Woodbury-Smith, Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK; marc.woodbury-smith@newcastle.ac.uk

MW-S and LD'A contributed equally.

Received 21 July 2022
Accepted 3 May 2023
Published Online First 8 June 2023

ABSTRACT

Background We present genomic and phenotypic findings of a transgenerational family consisting of three male offspring, each with a maternally inherited distal 220 kb deletion at locus 16p11.2 (BP2–BP3). Genomic analysis of all family members was prompted by a diagnosis of autism spectrum disorder (ASD) in the eldest child, who also presented with a low body mass index.

Methods All male offspring underwent extensive neuropsychiatric evaluation. Both parents were also assessed for social functioning and cognition. The family underwent whole-genome sequencing. Further data curation was undertaken from samples ascertained for neurodevelopmental disorders and congenital abnormalities.

Results On medical examination, both the second and third-born male offspring presented with obesity. The second-born male offspring met research diagnostic criteria for ASD at 8 years of age and presented with mild attention deficits. The third-born male offspring was only noted as having motor deficits and received a diagnosis of developmental coordination disorder. Other than the 16p11.2 distal deletion, no additional contributing variants of clinical significance were observed. The mother was clinically evaluated and noted as having a broader autism phenotype.

Conclusion In this family, the phenotypes observed are most likely caused by the 16p11.2 distal deletion. The lack of other overt pathogenic mutations identified by genomic sequencing reinforces the variable expressivity that should be heeded in a clinical setting. Importantly, distal 16p11.2 deletions can present with a highly variable phenotype even within a single family. Our additional data curation provides further evidence on the variable clinical presentation among those with pathogenetic 16p11.2 (BP2–BP3) mutations.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The contribution of the proximal 16p11.2 deletion (BP4–BP5) to autism spectrum disorder (ASD) is well established, but while the distal deletion (BP2–BP3) is broadly linked with obesity, less is known about neurodevelopmental and psychiatric sequelae. Given the frequency of such deletions, a further understanding of their neuropsychiatric and other clinical manifestations is needed.

WHAT THIS STUDY ADDS

⇒ Here, we summarise in detail a family with the transgenerational segregation of a distal 16p11.2 deletion. The findings, along with our wider curation of 16p11.2 BP2–BP3 CNVs, provide evidence for its contribution to ASD with incomplete penetrance and variable expressivity.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The findings in this study demonstrate the contribution of the 16p11.2 distal deletion to ASD development and further reinforces the fact that ASD-relevant CNVs often present with variable phenotypes even within a family and that this should be considered when counselling a family in a clinical genetics context.

INTRODUCTION

Autism spectrum disorder (ASD) is a common neurodevelopmental disorder hallmarked by symptoms in two core diagnostic domains: social communication and restrictive and repetitive interests and behaviours.[1 2] Twin and family studies were the first to earmark the familial nature of the disorder, with estimates as high as 90% in monozygotic twins and 10% in dizygotic twins.[3–5] Large-scale genome-wide linkage studies to identify genetic loci shared between affected family members have been conducted, but the results were largely negative.[6 7] It has since become apparent that there are 100 or more ASD-implicated loci affected by rare (<0.1% of control populations) or inherited point mutations or smaller deletions and insertions, as well as larger CNV or structural variants, each associated with <1% of ASD cases.[8–10] Consequently, a multitude of variants are seen to contribute to the autism spectrum phenotype, even among affected individuals from the same nuclear family.[11 12]

This notwithstanding, some mutations/loci are seen more frequently, and sometimes are observed to segregate with phenotype in multiple affected families. In such instances, it may be possible to examine the relationship between gene and phenotype more precisely, as, to some extent, the

To cite: Woodbury-Smith M, D'Abate L, Stavropoulos DJ, *et al*. *J Med Genet* 2023;**60**:1153–1160.

influence of background genetic variation and environment are attenuated. Notable examples of recurrent variants are 15q11-13[13] duplications and 1p36[14] and 16p11.2[15] proximal deletions.

The 16p11.2 locus, which encompasses a series of segmental duplications which mediate the formation of multiple, recurrent CNVs, is one of the most common genetic aetiologies identified in ASD (0.5%–1%).[16–18] Approximately 20%–25% of individuals with the ~600 kb 16p11.2 proximal deletion (BP4–BP5) have a diagnosis of autism.[19] Additional phenotypes associated with the proximal deletion are obesity, usually fully penetrant by adulthood (75%), deficits in motor coordination (60%) and language, speech and phonological impairment (70%).[19] The less common recurrent CNVs in this region are the adjacent 16p11.2 distal deletion/duplication (BP2–BP3), both of which also exhibit reduced penetrance in relation to cognition.[20 21] Similar to its proximal counterpart, the distal deletion syndrome is highly associated with obesity,[22 23] increased head circumference (HC) and similar rates of ASD (26%)[24] as observed in 16p11.2 proximal deletion carriers.[19] Similarly, a high prevalence of developmental delay (DD)[23 25] and increased risk of schizophrenia have been described.[26]

Of the nine protein-coding and three non-coding genes within the distal deletion, neuronal expression of adapter protein Src homology 2 B adapter protein 1 (*SH2B1*) has been shown to positively regulate leptin and insulin signalling through the Janus kinase pathway, making this gene a strong candidate mediating the obesity phenotype.[27 28] Transmembrane adapter protein linker for activation of T cells (*LAT*) is a candidate gene responsible for the neurodevelopmental phenotype, with overexpression leading to microcephaly in zebrafish and brain volumetric changes in mice.[29]

In this paper, we profile a family of five who was referred to the clinical genetics service at the Hospital for Sick Children (SickKids, Toronto, Ontario, Canada) and the Holland Bloorview Kids Rehabilitation Hospital (Toronto, Ontario, Canada) as part of the POND Network (https://pond-network.ca/) because of a diagnosis of ASD in the first-born male offspring (figure 1). Both microarray and whole-genome sequencing (WGS) were

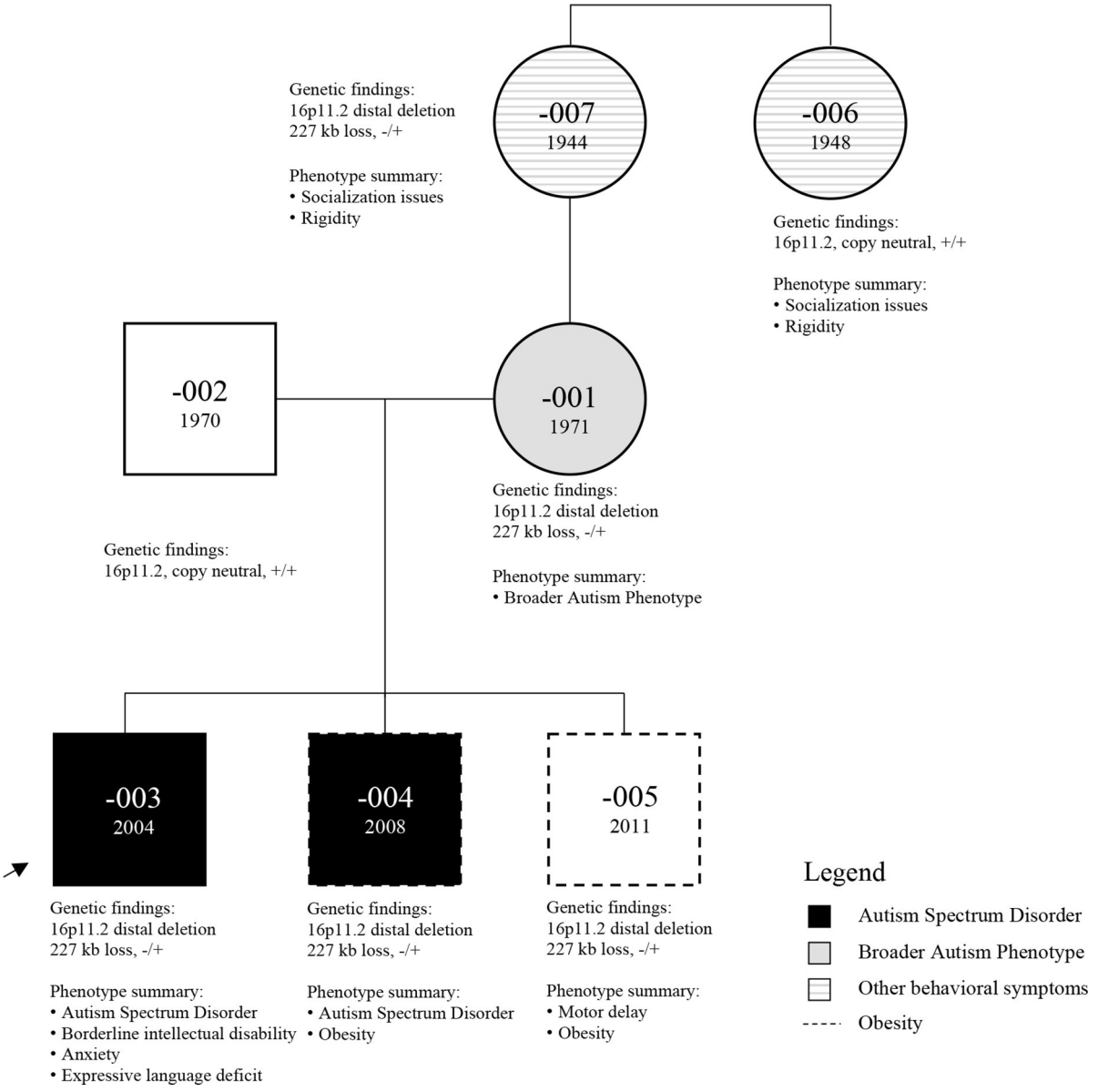

Figure 1 Pedigree of family 1-0616. Arrow indicates proband.

performed on the family, and a canonical, maternally inherited 227 kb 16p11.2 distal deletion was identified in all male offspring. All family members subsequently underwent medical and psychiatric assessments. Further data curation was undertaken from samples ascertained for neurodevelopmental disorders and congenital abnormalities.

METHODS

The family profiled in this study (ID: 1-0616) was first referred to the clinic for suspected DD of the first-born male offspring (-003, proband) at 20 months of age. At 4.5 years of age (August 2008), he received a clinical diagnosis of ASD, after being referred to a developmental paediatrician during the prior year. All five family members subsequently enrolled in SickKids' ASD research studies, which included WGS analysis as part of the Autism Speaks MSSNG project.[30–32]

WGS data are available through a data access committee at the MSSNG site (https://research.mss.ng). The microarray data are accessible in dbGaP via accession number phs001876.v1.p1. Full details of clinical, behavioural and cognitive assessments undertaken are available in the online supplemental note. For the purpose of this paper, the family also provided specific written consent for their data to be shared in the scientific literature in the form of this case report.

Microarray and WGS data were analysed as previously described.[32–34] For additional information, please consult the online supplemental note. The genomic coordinates presented in this paper are based on the February 2009 Human Genome Build (GRCh37/hg19).

RESULTS

Clinical profile of family 1-0616

The male proband was born at 38 weeks' gestation following a medically uneventful pregnancy and with birth weight of 6 lb 11 oz. The mother and father were aged 32 and 33, respectively, at the time of conception. No medical concerns were noted at birth. He was first referred to a paediatrician because of DD in language and motor milestones at 20 months, which included concerns about reduced motor tone and lack of babbling. He was then diagnosed with ASD at 4.5 years by a developmental paediatrician. As part of the ASD research study, he subsequently underwent a series of investigations at age 9 years. His full-scale IQ (FSIQ) was 78 (Stanford-Binet test: verbal IQ=91, non-verbal IQ=68), and adaptive behaviour was consistent with this (Vineland Adaptive Behaviour Scale, VABS-II: composite: 82, communication: 94, daily living: 78, socialisation: 80). Despite an average verbal IQ of 91, performance on the Oral and Written Language Scales (OWLS)-II revealed overall impairment in language (total language standard score: 62) and in expression (standard score: 50). Standard score in receptive language revealed borderline impairment (standard score: 78). Both Autism Diagnostic Interview - Revised (ADI-R, age 12 years) and Autism Diagnostic Observation Schedule (ADOS-G) were also performed. The ADOS (age 8 years) and ADI-R (age 12 years) were consistent with a diagnosis of autism across all domains (online supplemental table 4, Summary genomic and phenotypical findings for family 1-0616). At age 9 years, no concerns were expressed by the family regarding psychiatric comorbidity, and at this same age, weight was below average (weight: 31 kg, less than the first percentile), and he was noted to be tall for his age (height: 149 cm, 99th percentile). His HC was normal (age 9 years, HC: 53 cm, 64th percentile) (table 1 and online supplemental table 4)

The proband's younger siblings (male siblings -004 and -005) were subsequently enrolled in the Canadian Infant Siblings Study for ASD,[33] and both underwent detailed neurodevelopmental and clinical assessment. The older of the two siblings (-004) was born at 38 weeks' gestation following an uneventful pregnancy and labour with a birth weight of 6 lb 14 oz. At the time of assessment at the age of 5.5 years, he was noted to be obese (body mass index (BMI) >99th percentile) and to have difficulties with socialisation and communication with his peers. However, both Nipissing Developmental Screen[35] and NutriSTEP screening[36] were normal. A formal ASD evaluation by way of ADOS-2 was consistent with a research diagnosis of ASD, although a formal clinical diagnosis was never given. Unlike the proband, -004 did not have intellectual vulnerabilities (Wechsler Intelligence Scale for Children, WISC-V: Full Scale IQ, FSIQ: 92, Verbal IQ, VIQ: 106 and Non-verbal IQ, NVIQ: 85), and communication assessments and adaptive functioning were within normal limits

Table 1 Summary of phenotypical and genomic findings in family 1-0616

Family 1-0616	-001	-002	-003	-004	-005
Family member	Mother	Father	Proband	Second-born	Third-born
Sex	Female (XX)	Male (XY)	Male (XY)	Male (XY)	Male (XY)
Clinically relevant genetic findings‡	16p11.2 distal deletion, 227 kb, del(mat)/+	16p11.2 copy neutral+/+	16p11.2 distal deletion, 227 kb, del(mat)/+	16p11.2 distal deletion, 227 kb, del(mat)/+	16p11.2 distal deletion, 227 kb, del(mat)/+
Medical history:					
Obesity	–	–	–	+	+
Epilepsy	–	–	–	–	–
Other	–	–	–	–	–
ASD	–	–	+	+	–
BAP	+	–	NA	NA	NA
IQ*	109	122	78	92	117
Adaptive deficits†	NA	NA	82	88	NA
Anxiety	Moderate	NA	Severe	Severe	Mild
Other neurodevelopmental	Communication deficits	NA	NA	Attentional deficits	Motor deficits

*Full-scale IQ determined by multiple tests; see online supplemental table 4.
†Adaptive Behaviour Composite as measured by VABS-II.
‡CNV findings initially made by microarray (Affymetrix CytoScan HD) and subsequently called from whole-genome sequencing data (Illumina HiSeqX); see online supplemental table 5.
ASD, autism spectrum disorder; BAP, broader autism phenotype; NA, not applicable/phenotype not assessed.

(online supplemental table 4). Despite these relative patterns of strength, a number of externalising behaviours were reported by parents and teachers, both aggressive and oppositional/defiant in nature, as well as internalising problems such as anxiety and mood symptoms. However, none were consistent with a formal diagnosis of a psychiatric disorder. This sibling was followed up at the Hospital for Sick Children due to his obesity. At age 11 years, his BMI remained high (height 156.8 cm, 97th percentile; weight 66.3 kg, BMI 26.97, >99th percentile, Z=2.75).

Unlike his older siblings, the youngest brother (-005) was not reported to have any social and communication vulnerabilities, but did receive a diagnosis of developmental coordination disorder and dyslexia at age 6 years and was obese. He too was born following an uneventful pregnancy and delivery at 38 weeks. During his first year, gross motor vulnerabilities were noted and by formal assessment identified in the very low range in contrast to relatively intact fine motor development. By 24 months, however, both gross and fine motor developments were recorded as being delayed. He completed the Wechsler Preschool and Primary Scale of Intelligence, WPPSI-III at 4 years and 5 months of age as part of his enrolment in the Canadian Infant Sibling Study and scored above average in all domains (>113; see online supplemental table 4). Screening using the SRS and SCQ were within normal limits, and ADOS was not consistent with a diagnosis of ASD. However, vulnerabilities in the communicative domain were suggested by CCC-2 scores falling in the 14th percentile. No other developmental concerns were noted at this time. At 7 years, this sibling was also noted to be clinically overweight (height 132.5 cm, 98th percentile, weight 45.1 kg, BMI 25.69 kg/m^2, >99th percentile, Z=3.83).

Both the mother (-001) and father (-002) completed the Wechsler Abbreviated Scale of Intelligence (WASI-II) with scores consistent with average (mother, FSIQ: 103) and high average skills (father, FSIQ: 122). Results on the Broader Autism Phenotype Questionnaire (BAP-Q) and expert clinical judgement deemed the mother to have a broader autism phenotype (BAP). The father was not formally assessed for BAP, although his total BAP-Q score and subscale scores did indicate vulnerabilities.

Genetic findings

The proband and his two younger siblings all possessed a distal deletion at 16p11.2 (NC_000016.9:g.28824503_29051191del) detected by both microarray and WGS (figure 1 and online supplemental table 5). This 227 kb variant impacted the genes *ATXN2L*, *TUFM*, *SH2B1*, *ATP2A1*, *ATP2A1-AS*, *RABEP2*, *CD19*, *NFATC21P*, *SPNS1* and *LAT* and two microRNAs (MIR4517 and MIR4721) and is consistent with the 16p11.2 distal deletion region.[22] The gene *SH2B1* has been proposed to be associated with obesity.[22 23] This CNV was inherited from the mother, who, in turn, had inherited it from her mother (subject −007, figure 1). We were unable to gather phenotypic information on -007, although an anecdotal description by -001 suggests she may have had some social and communication vulnerabilities. The maternal aunt also was reported to have social vulnerabilities but did not possess the deletion.

While the 16p11.2 distal deletion in this family is the only genetic finding clinically deemed to contribute to the expression of the observed phenotypes, there were additional rare variants of unknown significance (VUSs) identified in the study subjects, given their rare, de novo status and/or possible, although tenuous, association with neurodevelopmental disorder (NDD) or neuropsychiatric phenotypes (see online supplemental table 5 for all rare CNVs and small variants derived from WGS). The third-born son possessed a de novo 192 kb duplication CNV at locus 5p15.33 involving four genes classified as a VUS. The proband and his siblings also inherited a rare 107 kb duplication at Xp22.31, impacting *KAL1*, a gene associated with Kallman syndrome, from their mother and maternal grandmother. Finally, the proband and one of his siblings (-004) inherited a 637 kb duplication at 2q21.2 from their father. This CNV impacted the *ANKRD30BL* and *GPR39* genes. Copy number detection via WGS did not detect any additional clinically relevant CNVs that were missed by microarray.

In the case of single-nucleotide variants (SNVs) and small insertions and deletions, the proband and mother possessed a stop-gain mutation in exon 19 of the calcium-channel *CACNA2D4* (NM_172364.5:c.1882C>T, p.(Arg628*)), one in a family of similar genes associated with NDDs and psychiatric disorders. Subjects -003 and -005 each had two rare de novo SNVs, none of which would be classified as pathogenic in the context of NDDs or ASD (online supplemental table 5). Polygenic risk scores (online supplemental figure 1) were calculated for all three male offspring as part of a larger study assessing the predictive utility of CNV detection in the diagnosis of ASD and neurodevelopmental vulnerabilities in a prospectively recruited sibling cohort.[33] The Pragmatic Rating Scale (PRS) for both male offspring with ASD (-003=0.24 and -004=0.84), reported as standardised values, was close to the mean for ASD cases (0.028, SD±1.024). The PRS of the youngest son (-005=−0.034) was also close to the mean of that of non-ASD sibs (−0.048, SD±0.952). While the general directionality of the scores aligned with ASD status, they did not correlate with clinical severity.

Curation of 16p11.2, BP2–BP3 CNVs in children ascertained for neurodevelopmental disorders and congenital anomalies

We also investigated the prevalence of 16p11.2 BP2–BP3 CNVs among families in the internal CNV database from The Centre for Applied Genomics, the Simons Simplex Collection (SSC) and the MSSNG database (table 2, extended form online supplemental tables 6–8) (n=10 113 ASD cases, 8372 male and 1741 female). A total of five individuals with ASD had a 16p11.2 distal deletion (one male and four female) (5/10 113=0.05%). We identified an affected female (1-0173-004) of European descent who had a paternally inherited deletion at this locus. The deletion was initially identified via microarray (NC_000016.9:g.28823927_29043875del). She was subsequently sequenced and included in the MSSNG database. A clinical assessment at 9.6 years of age led to an ASD diagnosis (ADOS calibrated severity score: 8, ADI algorithm total=43). Additionally, she had an FSIQ of >1 SD below the mean (Leiter-R full IQ=79). She was additionally noted to have communication deficits, as observed through the OWLS, with an oral composite score of 67 (>2 SD below the mean). At the same age, physical development metrics obtained from the medical clinic noted that her weight was average (29 kg, 12th percentile), accounting for age and sex. HC was 54 cm (93rd percentile). A second female (MSSNG00104-003) of East Asian ancestry with ASD had a 16p11.2 distal deletion (NC_000016.9:g.28747322_29062921del), which was de novo. No other medical or psychiatric history was available for this subject. A third female (AU4188302) of European ancestry with a maternally inherited deletion (NC_000016.9:g.28678522_29054721del) received an ASD diagnosis at 10.5 years of age. No additional history was available for this subject. There was an additional female participant with ASD who possessed a 16p11.2 distal deletion (-A136, NC_000016.9:g.28753451_29043875del). The origin of this deletion is unknown. Lastly,

Table 2 Cases and controls with 16p11.2 BP2–BP-3 distal deletion and duplications

Sample	Sex	Chr.	Start.hg19	End.hg19	Size (bp)	CNV	Genes impacted (n)	Inheritance
Cases								
1-0173-004	F	chr16	28 823 927	29 032 293	208 366	Loss	12	Paternal
- A136	F	chr16	28 753 451	29 043 875	290 424	Loss	13	NA
AU4188302	F	chr16	28 678 522	29 054 721	376 199	Loss	17	Maternal
MSSNG00104-003	F	chr16	28 747 322	29 062 921	315 599	Loss	13	De novo
SSC07716	M	chr16	28 683 322	29 055 121	371 799	Loss	14	De novo
1-0193-003	M	chr16	28 812 322	29 078 321	269 999	Gain	12	Maternal
1-0782-004	M	chr16	28 722 322	29 080 321	357 999	Gain	17	NA
7-0049-003	F	chr16	28 464 322	29 550 321	1 085 999	Gain	42	Paternal
Controls								
110036021749	M	chr16	28 814 098	29 083 503	269 405	Loss	12	NA

F, female; M, male; NA, information not available.

an autistic male proband from the SSC (SSC07716) with a de novo deletion (NC_000016.9: g.28683322_29055121del) was found.

In our queries, we also found three ASD cases with 16p11.2 distal duplications at this locus (two male and one female) (3/10 113=0.03%). One male proband (1-0193-003) of European descent had a duplication spanning 12 genes (NC_000016.9:g. 28812322 _29078321dup). The variant was maternally inherited. The parent did not report any psychiatric or medical diagnoses. The male was diagnosed with ASD at 12 years of age (ADOS calibrated severity metric=8). He had an average Leiter FSIQ of 97. In addition to ASD, the male proband was noted to have comorbid attention deficit hyperactivity disorder (ADHD) and clinically significant adaptive behaviour deficits evident through VABS-II subscale standard scores (socialisation=66, communication=48, daily living skills=39; 0.6–2.4 SD below the mean). Oral expression ability was slightly below average as assessed by the OWLS-II (oral expression standard score: 69). A second autistic male offspring (1-0784-004) was found to carry a 358 kb duplication (NC_000016.9:g.28722322_29080 321dup). He has an older autistic sibling who does not have the variant. He was diagnosed with ASD at 38 months of age. The inheritance status of this CNV could not be ascertained. Lastly, a female offspring (7-0049-003) was identified with a 1.1 Mb paternally inherited duplication (NC_000016.9:g.28464322_ 29550321dup), corresponding to BP1-4 at the 16p11.2 locus. She was diagnosed with ASD at 11 years of age, with comorbid ADHD (Swan Rating Scale 2010: ADHD inattentive subscale=5, oppositional defiant disorder subscale=5, ADHD hyperactive subscale=4) and adaptive behaviour deficits (Adaptive Behavior Assessment System (ABAS-II): general adaptive composite=56, >2 SD below the mean).

We further identified CNVs of >100 bp, which could reliably be detected from short-read WGS data,[37] and small variants in all genes within this interval (table 3, extended form online supplemental table 9) in ASD subjects. Deletions of <5 kb were noted in the following genes: *ATP2A1*, *NFATC2IP* and *TUFM*. Two rare loss of function (LoF) mutations in different probands were found in *ATP2A1*, one impacting exons 7–9 (NC_000016.9:g.28898057_28901860del) and the other exons 13–14 (NC_000016.9:g. 28907687_28911255del). An ASD proband and unaffected sibling carried a maternally inherited rare LoF mutation in *NFATC2IP* impacting exons 3–4. Lastly, a 177 bp deletion of exon 5 of the *TUFM* gene (NC_000016.9:g. 28855875_28856051del) was found in an ASD proband.

We examined the number of patients with deletions and duplications overlapping the 16p11.2 BP2–BP3 region (<2 Mb in size) identified in the Genome Diagnostics Clinical Laboratory at the Hospital for Sick Children (Toronto, Canada), which has 31 932 patients referred for DD, ASD and/or congenital anomalies. These diagnostic labels represent the reason for referral rather than confirmed diagnoses assigned to each child. Of note, a total of 28 (0.09%) distal deletions (16 male offspring and 12 female offspring) were identified (online supplemental table 10) in subjects with ages ranging from 9 months to 30 years. The most common reason for referral was the presence of DD (13/28=46.4%). Among the 16p11.2 distal deletions, five (17.9%) cited ASD as a reason for referral, three of whom also were described with DD. Additionally, four (14.3%) subjects were reported to have obesity. A number of individuals were also described as having dysmorphic features and other growth abnormalities such as clubfoot or cleft lip/palate on the clinical requisition. A total of 21 (0.07%) distal duplications (16 male offspring and 5 female offspring) overlapping this locus were found in children with ages ranging from 1 year to 12 years (online supplemental table 11). Two duplications were maternally inherited; two were paternally inherited; and one was de novo. The inheritance status of the remaining variants could not be ascertained. Of the 21 cases referred for genetic testing, 4 (19%) presented with ASD, 3 had ADHD (14.3%) and 15 (71%)

Table 3 Smaller deletions in genes within the 16p11.2 BP2–BP3 interval

Sample	Chr.	Start (hg19)	End (hg19)	Exons impacted	Size (bp)	Gene	CNV	Inheritance
2-0119-003	chr16	28 898 057	28 901 860	7–9	3804	*ATP2A1*	Loss	NA
1024	chr16	28 907 687	28 911 255	13–14	3569	*ATP2A1*	Loss	Maternal
SSC11925	chr16	28 969 291	28 972 486	3–4	3196	*NFATC2IP*	Loss	Maternal
SSC12583	chr16	28 969 291	28 972 486	3–4	3196	*NFATC2IP*	Loss	Maternal
SSC11987*	chr16	28 969 291	28 972 486	3–4	3196	*NFATC2IP*	Loss	Maternal
3-0707-000	chr16	28 855 875	28 856 051	5	177	*TUFM*	Loss	Maternal

*Unaffected sibling.

had a DD diagnosis. Unlike individuals with the distal deletion, two (9.5%) were underweight (less than the third percentile) and none reported obesity. This finding supports results from previous studies which demonstrate the mirrored phenotypes associated with 16p11.2 distal CNVs, with the deletion being primarily linked to elevated BMI.[20 21 38 39]

DISCUSSION

In this study, we describe a family in which a deletion impacting the 16p11.2 distal locus is transmitted across three generations. Despite sharing an identical CNV, the phenotype varies quite substantially between family members carrying it. Most significantly, there was no shared diagnosis between all three siblings, and only the two younger siblings were diagnosed with obesity and the proband and middle child diagnosed with ASD, although the middle child's ASD diagnosis was a research and not a clinical diagnosis. Of course, it is not unusual for ASD symptoms to manifest in middle childhood and beyond, and the same is true of obesity. Despite the 16p11.2 distal deletion being associated with early-onset obesity, both the mother and first-born son were not obese, with the latter in the first percentile for weight. Consequently, this individual may yet develop symptoms, as he was last seen in the clinic at 9 years of age. The third-born male offspring may equally qualify for an ASD diagnosis later in childhood. Indeed, the assessment of communication did earmark the presence of some early vulnerabilities. Notably, the three siblings' mother also met the criteria for the BAP and had no medical history, including the absence of obesity. There is much to support the possibility, therefore, that in this family the CNV has high penetrance yet variable expressivity from individual to individual.

Also of interest is the pattern of other vulnerabilities. For example, although the proband had relative deficits in non-verbal IQ, all other family members had a fairly 'flat' intellectual profile in the average range. No other genetic abnormalities were identified that would offer an explanation for the proband's non-verbal deficits, and so we believe it is most likely related to the 16p11.2 distal deletion. It would appear, therefore, that this may be a more variable aspect of phenotype, given that more family members had ASD/BAP than low IQ. Moreover, given the fact that the middle sibling (-004) had research-diagnosed ASD in association with normal IQ, it seems reasonable to conclude, for this family at least, that ASD is not simply an epiphenomenon and secondary to impaired intellectual function. Similarly, comorbidity for externalising behaviour was seen in the middle sibling but not the proband, again indicating that this is not simply the result of either impaired intellectual function or ASD. It is also important to note that the third-born child (-005) received intensive behavioural intervention from toddlerhood, specifically in the form of an assistant in the classroom, and help with school-related tasks at home. This stands in contrast to his older siblings. We postulate that this vigilance at least with respect to ASD may have contributed to improved developmental trajectory.

Our curation of data from the genome diagnostics lab at SickKids (online supplemental tables 8 and 9) indicates that neurodevelopmental phenotypes are common in the presence of the 16p11.2 distal deletion. Large-scale phenotype studies assessing pathogenic CNV frequencies in the general population estimate prevalence of 16p11.2 distal deletions between 0.014% and 0.02% and reciprocal duplications at 0.016%–0.038%, not unlike the association of distal deletions and, to a lesser extent, duplications, to neuropsychiatric phenotypes.[21 39 40] The reported penetrance for cognitive performance are 23% and 11% for the distal deletion and distal duplication, respectively.[21]

No pattern in phenotype overlap is evident from these clinical data, although there are biases inherent in data curated from different clinical sources. Specifically, cases referred though mental health services will be enriched for neuropsychiatric information, whereas paediatric services are more likely to comprise a broader range of medical characteristics. That being said, our assessment of this locus earmarks the importance of potential early detection using CNVs for ASD as biomarkers, given the strong likelihood of subsequent developmental vulnerabilities, as well as the potential for impact of behavioural interventions.[41]

Our curation from our internal CNV database also identified three autistic individuals with distal 16p11.2 duplications. Fernandez and colleagues previously described the clinical features of three children with 16p11.2 duplications. Although co-ordinates are not provided, they too observed a variable phenotype including ASD and DD.[15] There is in fact very little literature on neuropsychiatric phenotypes among patients with distal 16p11.2 duplications. Given out findings, this does highlight the need to also study duplications in this region in more detail.

The CNV in this family does impact *SH2B1*, which is implicated in obesity. However, although both siblings were clinically obese, neither the proband nor the mother had any history of obesity at the time of assessment. At present, the dietary habits of the individual members of the family are similar. The pattern of obesity observed in this family, therefore, may be indicative of variable expressivity for obesity, as neither of the other genetic abnormalities identified in this family are associated with growth or obesity. There may, of course, be other more common variants that impact risk through interaction with *SH2B1*.

It is unclear which genes confer risk of neurodevelopmental diagnoses. The deletion in this family overlaps a number of genes that are known to be brain expressed, particularly in the cerebellum, and so there are several candidates. For example, *ATXN2L* is widely expressed in tissue, including the central nervous system, but its ataxin type 2 protein product, although a member of a neurodegenerative family of proteins, is of unknown function.[42] The other impacted genes are largely also of undetermined function, although two are immune related (*CD19*[43] and *LAT*).[44] *LAT* was previously show to mediate brain size changes in a murine knockout model, although it is the only gene in this interval to be implicated in this phenotype. We also considered other CNVs in the literature impacting this locus to determine whether there is an area of minimal overlap, but the small number of cases identified and similar CNV size precluded this. Given that our study has demonstrated that this locus has a reasonably high penetrance for ASD, it is important to pursue further investigation of these genes as this may shed unique insight on the pathogenesis of ASD, given that none of the genes fall into existing functional domains, such as synaptic, that are known to be important in the disorder's pathogenesis.[45]

Author affiliations
[1]The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Ontario, Canada
[2]Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK
[3]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada
[4]Genome Diagnostics, Department of Pediatric Laboratory Medicine, Hospital for Sick Children, Toronto, ON, Canada
[5]Ron Joyce Children's Health Centre, Autism Spectrum Disorder (ASD) Program and Child and Youth Mental Health Program, McMaster Autism Research Team, McMaster University, Hamilton, Hamilton, Ontario, Canada
[6]Department of Genetic Counselling, The Hospital for Sick Children, Toronto, Ontario, Canada

[7]Autism Research Centre, Holland Bloorview Kids Rehabilitation Centre, Toronto, Ontario, Canada
[8]McLaughlin Centre, University of Toronto, Toronto, Ontario, Canada

Acknowledgements The authors wish to first thank the family for permitting the sharing of their medical data and for their general commitment and support of autism research activities at The Hospital for Sick Children. All genotyping and sequencing data were generated by The Centre for Applied Genomics. The authors further wish to acknowledge the resources of Autism Speaks, MSSNG (www.mss.ng) and the dedication of the involved families, without whom none of this would be possible, and the generosity of the donors who supported this programme. Access to the data contained in MSSNG can be obtained by completing data access agreements at https://research.mss.ng. We are grateful to all of the families at the participating Simons Simplex Collection site, as well as the principal investigators (A Beaudet, R Bernier, J Constantino, E Cook, E Fombonne, D Geschwind, R Goin-Kochel, E Hanson, D Grice, A Klin, D Ledbetter, C Lord, C Martin, D Martin, R Maxim, J Miles, O Ousley, K Pelphrey, B Peterson, J Piggot, C Saulnier, M State, W Stone, J Sutcliffe, C Walsh, Z Warren and E Wijsman). We appreciate obtaining access to genetic and phenotypical data on SFARI Base. Approved researchers can obtain the SSC population dataset described in this study by applying at https://base.sfari.org.

Contributors MW-S, LD'A, SWS and JH conceived the study. MW-S and LD'A wrote the manuscript and revisions according to input from all authors. EA, ID, NH and AI collected phenotypical data. LD'A, DJS, MZ and BT undertook genome analyses and data curation.MW-S is the guarantor.

Funding This study was also supported by grants from Genome Canada, Canada Foundation for Innovation (CFI), Government of Ontario, Canadian Institutes of Health Research, SickKids Foundation, Ontario Brain Institute, Autism Speaks and University of Toronto McLaughlin Centre. SWS holds the Northbridge Chair in Pediatric Research, a joint Hospital-University Chair between the University of Toronto, The Hospital for Sick Children and the SickKids Foundation. Additional dataset acknowledgements funding support for the Study of Addiction: Genetics and Environment (SAGE) was provided through the NIH Genes, Environment and Health Initiative (GEI) (U01 HG004422). SAGE is one of the genome-wide association studies funded as part of the Gene Environment Association Studies (GENEVA) under GEI. Assistance with phenotype harmonisation and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01 HG004446). Assistance with data cleaning was provided by the National Center for Biotechnology Information. Support for collection of datasets and samples was provided by the Collaborative Study on the Genetics of Alcoholism (U10 AA008401), the Collaborative Genetic Study of Nicotine Dependence (COGEND, P01 CA089392) and the Family Study of Cocaine Dependence (R01 DA013423). Funding support for genotyping, which was performed at the Johns Hopkins University Center for Inherited Disease Research (CIDR), was provided by the NIH GEI (U01HG004438), the National Institute on Alcohol Abuse and Alcoholism, the National Institute on Drug Abuse and the NIH contract 'High throughput genotyping for studying the genetic contributions to human disease' (HHSN268200782096C). The datasets used for the analyses described in this paper were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/projects/gap/cgibin/study.cgi?study_id=phs000092.v1.p1 through dbGaP accession number phs000092.v1.p. The authors acknowledge the contribution of data from Genetic Architecture of Smoking and Smoking Cessation accessed through dbGAP (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000404.v1.p1, accession number phs000404.v1.p1). Funding support for genotyping, which was performed at the CIDR, was provided by 1 X01 HG005274-01. CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University (contract number HHSN268200782096C). Assistance with genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01 HG004446). Funding support for collection of datasets and samples was provided by the COGEND (P01 CA089392) and the University of Wisconsin Transdisciplinary Tobacco Use Research Center (P50 DA019706, P50 CA084724). The NEI Refractive Error Collaboration (NEIREC) Database found at https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000303.v1.p1 through dbGaP accession number phs000303.v1.p1 was also used for analyses. Funding support for NEIREC was provided by the National Eye Institute. We thank NEIREC participants and the NEIREC Research Group for their valuable contribution to this research. Funding support for the 'CIDR Visceral Adiposity Study' was provided through the Division of Aging Biology and the Division of Geriatrics and Clinical Gerontology, NIA. The CIDR Visceral Adiposity Study includes a genome-wide association study funded as part of the Division of Aging Biology and the Division of Geriatrics and Clinical Gerontology, NIA. Assistance with phenotype harmonisation and genotype cleaning, as well as with general study coordination, was provided by Heath ABC Study Investigators. This dataset was obtained from dbGaP at https://www.ncbi.nlm.nih.gov/projects/gap/cgibin/study.cgi?study_id=phs000169.v1.p1 through accession number phs000169.v1.p1.

Competing interests SWS is on the Scientific Advisory Committee of Population Bio, serves as a highly cited academic advisor for the King Abdulaziz University, and intellectual property from aspects of his research held at the Hospital for Sick Children are licensed to Athena Diagnostics and Population Bio. DJS contributed to initial development of Phenotips and holds equity in the company.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by Hospital for Sick Children Research Ethics Boar (0019980189). The participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the paper or uploaded as supplementary information.

ORCID iDs
Marc Woodbury-Smith http://orcid.org/0000-0002-0593-8917
Brett Trost http://orcid.org/0000-0003-4863-7273

REFERENCES

1 Anagnostou E, Zwaigenbaum L, Szatmari P, *et al*. Autism spectrum disorder: advances in evidence-based practice. *CMAJ* 2014;186:509–19.
2 Zwaigenbaum L, Penner M. Autism spectrum disorder: advances in diagnosis and evaluation. *BMJ* 2018;361:k1674.
3 Hallmayer J, Cleveland S, Torres A, *et al*. Genetic heritability and shared environmental factors among twin pairs with autism. *Arch Gen Psychiatry* 2011;68:1095–102.
4 Colvert E, Tick B, McEwen F, *et al*. Heritability of autism spectrum disorder in a UK population-based twin sample. *JAMA Psychiatry* 2015;72:415–23.
5 Tick B, Bolton P, Happé F, *et al*. Heritability of autism spectrum disorders: a meta-analysis of twin studies. *J Child Psychol Psychiatry* 2016;57:585–95.
6 Liu X-Q, Paterson AD, Szatmari P, *et al*. Genome-wide linkage analyses of quantitative and categorical autism subphenotypes. *Biol Psychiatry* 2008;64:561–70.
7 Weiss LA, Arking DE, Daly MJ, *et al*. A genome-wide linkage and association scan reveals novel loci for autism. *Nature* 2009;461:802–8.
8 Yuen RKC, Merico D, Cao H, *et al*. Genome-wide characteristics of de novo mutations in autism. *NPJ Genom Med* 2016;1:160271–1602710.
9 Satterstrom FK, Kosmicki JA, Wang J, *et al*. Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell* 2020;180:568–84.
10 Schaaf CP, Betancur C, Yuen RKC, *et al*. A framework for an evidence-based gene list relevant to autism spectrum disorder. *Nat Rev Genet* 2020;21:367–76.
11 Woodbury-Smith M, Nicolson R, Zarrei M, *et al*. Variable phenotype expression in a family segregating microdeletions of the NRXN1 and MBD5 autism spectrum disorder susceptibility genes. *NPJ Genom Med* 2017;2:17.
12 Woodbury-Smith M, Zarrei M, Wei J, *et al*. Segregating patterns of copy number variations in extended autism spectrum disorder (ASD) pedigrees. *Am J Med Genet B Neuropsychiatr Genet* 2020;183:268–76.
13 Scoles HA, Urraca N, Chadwick SW, *et al*. Increased copy number for methylated maternal 15q duplications leads to changes in gene and protein expression in human cortical samples. *Mol Autism* 2011;2:19.
14 Heilstedt HA, Ballif BC, Howard LA, *et al*. Population data suggest that deletions of 1p36 are a relatively common chromosome abnormality. *Clin Genet* 2003;64:310–6.
15 Fernandez BA, Roberts W, Chung B, *et al*. Phenotypic spectrum associated with de novo and inherited deletions and duplications at 16p11.2 in individuals ascertained for diagnosis of autism spectrum disorder. *J Med Genet* 2010;47:195–203.
16 Kumar RA, KaraMohamed S, Sudi J, *et al*. Recurrent 16p11.2 microdeletions in autism. *Hum Mol Genet* 2008;17:628–38.
17 Marshall CR, Noor A, Vincent JB, *et al*. Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 2008;82:477–88.
18 Weiss LA, Shen Y, Korn JM, *et al*. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 2008;358:667–75.
19 Chung WK, Roberts TP, Sherr EH, *et al*. 16P11.2 deletion syndrome. *Curr Opin Genet Dev* 2021;68:49–56.

20 Jacquemont S, Reymond A, Zufferey F, *et al*. Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature* 2011;478:97–102.
21 Kendall KM, Bracher-Smith M, Fitzpatrick H, *et al*. Cognitive performance and functional outcomes of carriers of pathogenic copy number variants: analysis of the UK Biobank. *Br J Psychiatry* 2019;214:297–304.
22 Bachmann-Gagescu R, Mefford HC, Cowan C, *et al*. Recurrent 200-kb deletions of 16p11.2 that include the SH2B1 gene are associated with developmental delay and obesity. *Genet Med* 2010;12:641–7.
23 Bochukova EG, Huang N, Keogh J, *et al*. Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature* 2010;463:666–70.
24 Loviglio MN, Leleu M, Männik K, *et al*. Chromosomal contacts connect loci associated with autism, BMI and head circumference phenotypes. *Mol Psychiatry* 2017;22:836–49.
25 Barge-Schaapveld DQCM, Maas SM, Polstra A, *et al*. The atypical 16p11.2 deletion: a not so atypical microdeletion syndrome? *Am J Med Genet A* 2011;155A:1066–72.
26 Guha S, Rees E, Darvasi A, *et al*. Implication of a rare deletion at distal 16p11.2 in schizophrenia. *JAMA Psychiatry* 2013;70:253–60.
27 Giuranna J, Volckmar A-L, Heinen A, *et al*. The effect of SH2B1 variants on expression of leptin- and insulin-induced pathways in murine hypothalamus. *Obes Facts* 2018;11:93–108.
28 Jiang L, Su H, Wu X, *et al*. Leptin receptor-expressing neuron SH2B1 supports sympathetic nervous system and protects against obesity and metabolic disease. *Nat Commun* 2020;11.
29 Loviglio MN, Arbogast T, Jønch AE, *et al*. The immune signaling adaptor LAT contributes to the neuroanatomical phenotype of 16p11.2 BP2-BP3 CNVs. *Am J Hum Genet* 2017;101:564–77.
30 Yuen RKC, Thiruvahindrapuram B, Merico D, *et al*. Whole-genome sequencing of quartet families with autism spectrum disorder. *Nat Med* 2015;21:185–91.
31 Jiang Y, Yuen RKC, Jin X, *et al*. Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. *Am J Hum Genet* 2013;93:249–63.
32 C Yuen RK, Merico D, Bookman M, *et al*. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nat Neurosci* 2017;20:602–11.
33 D'Abate L, Walker S, Yuen RKC, *et al*. Predictive impact of rare genomic copy number variations in siblings of individuals with autism spectrum disorders. *Nat Commun* 2019;10:5519.
34 Zarrei M, Burton CL, Engchuan W, *et al*. A large data resource of genomic copy number variation across neurodevelopmental disorders. *NPJ Genom Med* 2019;4:26.
35 Limbos MM, Joyce DP, Roberts GJ. Nipissing district developmental screen: patterns of use by physicians in Ontario. *Can Fam Physician* 2010;56:e66–72.
36 Randall Simpson JA, Keller HH, Rysdale LA, *et al*. Nutrition Screening Tool for Every Preschooler (NutriSTEP): validation and test-retest reliability of a parent-administered questionnaire assessing nutrition risk of preschoolers. *Eur J Clin Nutr* 2008;62:770–80.
37 Trost B, Walker S, Wang Z, *et al*. A comprehensive workflow for read depth-based identification of copy-number variation from whole-genome sequence data. *Am J Hum Genet* 2018;102:142–55.
38 Crawford K, Bracher-Smith M, Owen D, *et al*. Medical consequences of pathogenic CNVs in adults: analysis of the UK Biobank. *J Med Genet* 2019;56:131–8.
39 Birnbaum R, Mahjani B, Loos RJF, *et al*. Clinical characterization of copy number variants associated with neurodevelopmental disorders in a large-scale multiancestry Biobank. *JAMA Psychiatry* 2022;79:250–9.
40 Stefansson H, Meyer-Lindenberg A, Steinberg S, *et al*. CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature* 2014;505:361–6.
41 Vorstman J, Scherer SW. What a finding of gene copy number variation can add to the diagnosis of developmental neuropsychiatric disorders. *Curr Opin Genet Dev* 2021;68:18–25.
42 Auburger G, Sen N-E, Meierhofer D, *et al*. Efficient prevention of neurodegenerative diseases by depletion of starvation response factor ataxin-2. *Trends Neurosci* 2017;40:507–16.
43 Wang K, Wei G, Liu D. CD19: a biomarker for B cell development, lymphoma diagnosis and therapy. *Exp Hematol Oncol* 2012;1:36.
44 Balagopalan L, Kortum RL, Coussens NP, *et al*. The linker for activation of T cells (LAT) signaling hub: from signaling complexes to microclusters. *J Biol Chem* 2015;290:26422–9.
45 Stamouli S, Anderlid B-M, Willfors C, *et al*. Copy number variation analysis of 100 twin pairs enriched for neurodevelopmental disorders. *Twin Res Hum Genet* 2018;21:1–11.