Segregating patterns of copy number variations in extended autism spectrum disorder (ASD) pedigrees

Marc Woodbury-Smith1,2 | Mehdi Zarrei2 | John Wei2 | Bhooma Thiruvahindrapuram2 | Irene O'Connor3 | Andrew D. Paterson2,4 | Ryan K. C. Yuen2 | Jila Dastan5 | Dimitri J. Stavropoulos2,5 | Jennifer L. Howe2 | Ann Thompson3 | Morgan Parlier6 | Bridget Fernandez7 | Joseph Piven6 | Evdokia Anagnostou8 | Stephen W. Scherer2,9 | Veronica J. Vieland10 | Peter Szatmari11

1Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, UK
2The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Ontario, Canada
3Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, Ontario, Canada
4Division of Epidemiology and Biostatistics, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada
5Department of Paediatric Laboratory Medicine, Molecular Genetics Laboratory, The Hospital for Sick Children, Toronto, Ontario, Canada
6Carolina Institute for Developmental Disabilities, University of North Carolina at Chapel Hill, School of Medicine, Chapel Hill, North Carolina
7Provincial Medical Genetics Program, Health Sciences Center, St. John's, Newfoundland, Canada
8Holland Bloorview Hospital, Toronto, Ontario, Canada
9McLaughlin Centre and Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada
10Battelle Center for Mathematical Medicine, The Research Institute at Nationwide Children's Hospital, Columbus, Ohio
11Centre for Addiction and Mental Health, The Hospital for Sick Children & University of Toronto, Toronto, Ontario, Canada

Abstract
Autism spectrum disorder (ASD) is a relatively common childhood onset neurodevelopmental disorder with a complex genetic etiology. While progress has been made in identifying the de novo mutational landscape of ASD, the genetic factors that underpin the ASD's tendency to run in families are not well understood. In this study, nine extended pedigrees each with three or more individuals with ASD, and others with a lesser autism phenotype, were phenotyped and genotyped in an attempt to identify heritable copy number variants (CNVs). Although these families have previously generated linkage signals, no rare CNV segregated with these signals in any family. A small number of clinically relevant CNVs were identified. Only one CNV was identified that segregated with ASD phenotype; namely, a duplication overlapping DLGAP2 in three male offspring each with an ASD diagnosis. This gene encodes a synaptic scaffolding protein, part of a group of proteins known to be pathologically implicated in ASD. On the whole, however, the heritable nature of ASD in the families studied remains poorly understood.
1 | INTRODUCTION

Major progress in elucidating the genetic causes of the complex neuropsychiatric disorders, including autism spectrum disorder (ASD), has been made in recent years. ASD is one of the commonest early childhood onset developmental disorders, occurring with a population prevalence of more than 1% (Centers for Disease Control & Prevention, 2012). It is characterized by core socio-communicative and behavioral symptoms, and associated with major psychiatric and medical disorders (Anagnostou et al., 2014). Variants of different allele frequencies are now known to play an etiological role, with most progress having been made in the identification of rare, de novo, copy number variants (CNVs) and single nucleotide variants (Woodbury-Smith & Scherer, 2018). As a result of ascertainment and study design, many studies focus on only nuclear families, including singletons, discordant sibpairs, and affected sibpairs, and so the wider segregation of variants has not been studied. What is particularly striking about ASD, however, is the fact that ASD tends to segregate in families, with first and second-degree members of families with ASD probands often diagnosed with ASD or classified with a lesser variant termed the broader autism phenotype (BAP; Sasson et al., 2013). Heritability estimates as high as 90% support a major role for genetic factors rather than common environment in explaining this familial nature of ASD and BAP (Tick, Bolton, Happe, Rutter, & Rijsdijk, 2016). There is also evidence that genetic predisposition may extend to other neuropsychiatric disorders in families (De Rubeis & Buxbaum, 2015).

The fact that ASD often shows familial segregation has been investigated in a series of genome-wide linkage studies (Abrahams & Geschwind, 2008; Allen-Brady et al., 2010; Szatmari, 2007; Woodbury-Smith et al., 2018). These have identified a number of genome-wide significant signals for ASD and related phenotypes, earmarking genomic loci that could potentially harbor ASD-associated variants. Some of these loci overlap genes of potential significance to ASD, while other loci overlap regions linked and/or associated with other neurodevelopmental disorders. Only two loci, at 7q35 (Abrahams & Geschwind, 2008) and 20p13 (Werling, Lowe, Luo, Cantor, & Geschwind, 2014), have been replicated at genome-wide significance.

More recent studies have also investigated the segregation of CNVs across families (Woodbury-Smith et al., 2015). We reported previously on the segregation of CNVs in 19 US and Canadian extended pedigrees (Woodbury-Smith et al., 2015). Such families are likely to be enriched for heritable forms of the disorder, and provide the opportunity to track variants more widely (Wijmstra, 2012). Although we did not identify any rare variants that were shared widely among family members with ASD and/or BAP, there were examples of CNVs that segregated with ASD phenotype in a subset of family members. None coincided with identified linkage signals in these same families (Piven et al., 2013), pointing to a potential source of complexity in ASD’s heritable genetic architecture. There is also evidence that CNVs of potential etiological significance may segregate with other psychiatric disorders in such families (Sato et al., 2012), consistent with the evidence of shared genetic risk, and earmarking loci that cause brain vulnerability rather than disorder per se.

In this current study, we focused on an examination of CNVs in nine newly recruited Canadian extended pedigrees. Having already undertaken a genome-wide linkage study that included these same families (Woodbury-Smith et al., 2018), we were interested in whether any rare ASD and/or BAP segregating CNVs overlapped with any of these identified linkage signals. Additionally, and based on our previous study, we were also interested in identifying rare CNVs that were shared among ASD/BAP family members.

2 | METHODS

2.1 | Recruitment

We recruited extended pedigrees with at least three ASD cases spread across at least two nuclear families. All families were either known to the authors through previous studies or identified through advertising. Initial screening was by way of a telephone based assessment in conjunction with the Autism Family History Interview (Piven et al., 1990). All families who met the inclusion criteria were then followed for further clinical assessment. To minimize etiologic heterogeneity, families were excluded from the study if there was evidence of the following co-occurring medical conditions, thought to be etiologically related to autism, in one of the index probands: tuberous sclerosis, neurofibromatosis, phenylketonuria, Fragile X syndrome, or significant CNS injury, although no such families were identified. All individuals were of northern European ancestry. Data collection took place under Institutional Review Board’s approval, and the research was conducted in accordance with the World Medical Association Declaration of Helsinki. Written informed consent was obtained from subjects or their proxy decision maker after the study had been fully explained.

2.2 | Clinical diagnosis

ASD diagnosis was confirmed by expert clinical judgment incorporating information from the Autism Diagnostic Interview Revised (Lord, Rutter, & Le Couteur, 1994) and Autism Diagnostic Observation Schedule – Revised (Lord et al., 2000), which were administered by trained and reliable clinicians. All participants classified as ASD met DSM-IV criteria for either Autistic Disorder, Asperger syndrome or Pervasive Developmental Disorder Not Otherwise Specified according to the criteria in Risi et al. (2006). Although DSM-IV criteria were used at the time of assessment, all diagnosed participants also meet the DSM 5 criteria for ASD.
Among non-ASD family members, assessment for BAP was undertaken. The BAP Questionnaire was used for diagnosis of BAP in individuals greater than 15 years of age (Hurley, Losh, Parlier, Reznick, & Piven, 2007). The measure was completed by the participant about him/herself (the self-version) and by someone close to the participant about him/her (parent or spouse, the informant version) to obtain an average score (between the self and informant scores). Whenever available, the average scores were utilized. A BAP diagnosis was assigned if an individual met gender-specific criteria in any domain.

Additional clinical evaluation was undertaken using the Wechsler Abbreviated Scale of Intelligence (Pearson Clinical), the Vineland Adaptive Behavior Scale (Pearson Clinical), the Repetitive Behavior Scale (Lam & Aman, 2007), and the Oral and Written Language Scales (Pearson Clinical). Clinical information pertaining to diagnosis of other neurodevelopmental disorders, including epilepsy, along with known neonatal complications was collected using a structured parent completed questionnaire.

### 2.3 Genotyping

A total of 170 family members from the 9 extended pedigrees were genotyped on one of four different platforms (a small number of individuals were genotyped on more than one platform): Illumina HumanCoreExome (N = 166), Illumina Omnie2.5M or 1.0M (N = 18), Affymetrix Genomewide Human SNP array 6.0 (N = 15) or Affymetrix CytoScan HD (N = 18). We called CNVs as previously described (Oskoui et al., 2015; Pinto et al., 2011; Woodbury-Smith et al., 2015; Zarrei et al., 2018). Four different CNV calling algorithms were used to annotate high-confidence CNVs. These included Affymetrix Chromosome Analysis Suite (ChAS), iPattern (Pinto et al., 2011), BioDiscovery Nexus (Darvishi, 2010), and Partek Genomics Suite (Grayson & Aune, 2011). A stringent set of variants was defined for further analyses. For the Affymetrix chip, this set included variants detected by one or both of ChAS or iPattern, and if detected by only one of these, then also by one of Nexus or Partek. For stringent calls on the X chromosome, we required calling by both ChAS and iPattern. For Illumina generated genotypes, stringent calls included variants detected by iPattern, and one of PennCNV (Wang et al., 2007) or QuantiSNP (Colella et al., 2007). Stringent calls on the X chromosome required calling by all three algorithms.

We analyzed CNVs covered by at least five consecutive probes and those with a minimum length of 15 kb. CNVs were filtered to prioritize rare variants that occurred with a frequency of <0.1% in control samples (N = 10,851; refer to Zarrei et al., 2018, for list of control samples). For the purpose of filtering, CNVs with >50% reciprocal overlap were deemed identical. We also removed all CNVs that had >70% overlap with a known segmental duplication. We further restricted our list to those with more than 75% overlap with copy number stable regions, according to the stringent CNV map of the human genome (Zarrei, MacDonald, Merico, & Scherer, 2015). All clinically relevant CNVs described in the index family have been validated using the SYBR green-based quantitative PCR method, TaqMan Copy Number Assays and/or visualization of WGS generated bam files (see below) using the Integrated Genomics Browser (IGV; Nicol, Helt, Blanchard, Raja, & Loraine, 2009). The genomic coordinates presented in this article are based on the February 2009 Human Genome Build (GRCh37/hg19).

### 2.4 CNV prioritization

We first examined whether any of the rare CNVs overlapped linkage signals in these same families. Linkage signals were defined as those with a PPL ≥ 0.1, whose margins were at PPL ≥ 0.05. We then prioritized deleterious and duplications ≥ 3 Mb in size. Next, variants were filtered to extract only those that overlapped one or more coding exons, and then the following CNVs were extracted: (a) variants overlapping a known ASD gene (Yuen et al., 2017) and (b) variants congruent with known genomic syndromes as those curated in DECIPHER (Firth et al., 2009) or ClinGen (Rehm et al., 2015). Only those DECIPHER/ClinGen loci that earmarked developmental brain syndromes were retained in our final list. Further, variants were clinically classified as "likely benign," "of unknown significance" (VOUS), and "likely pathogenic" according to the American College of Medical Genetics guidelines (Kearney et al., 2011). CNVs in cases with whole genome data from MSSNG project (Yuen et al., 2017) were also visualized in IGV to further investigate structure of duplications and to refine breakpoints of variants.

### 2.5 Whole genome sequencing

Twenty six individuals also underwent whole genome sequencing as part of the MSSNG initiative (Yuen et al., 2017). Variant prioritizations were followed as previously described (Yuen et al., 2017). Briefly, identified variants were annotated, and likely deleterious mutations prioritized to capture those that were rare (MAF ≤ 1%), LoF ( nonsense, splicing site, frameshift), and damaging de novo missense mutations (damaging as evidenced by four of following algorithms: SIFT ≤ 0.05, Polyphén ≥ 0.95, CADD ≥ 15, Mutation Assessor score ≥ 2, PhyloP ≥ 2.4). MAF was based upon filtering according to frequencies in the following control samples: 1000 Genomes Project (Genomes Project Consortium et al., 2015), NHLBI-ESP (Fu et al., 2013), ExAC (v.0.3.1) (Lek et al., 2016), and Complete Genomics internal control data. Variants described in this article have been validated by Sanger sequencing.

### 3 RESULTS

A total of nine extended pedigrees were recruited. The total sample includes 318 family members, among whom 170 were genotyped on one or more platforms and 132 had phenotype data (122 with both genotype and phenotype data). The sample comprised 46 individuals diagnosed with ASD, 22 with BAP, and 68 with neither ASD nor BAP. The clinical characteristics of these pedigrees are described in Supplementary Table S1. Identified CNVs are detailed in Supplementary Table S2.

---

*Woodbury-Smith et al.*
3.1 | CNVs segregating with linkage signals within and across families

We first examined linkage signals generated by all nine pedigrees combined. A number of CNVs overlapped these signals, but none were inherited. As such, there were no segregating rare CNVs that would theoretically explain any of the signals observed. The same was true when examining pedigree specific linkage signals. Only two families showed peaks above a 10% PPL threshold (Ped 4, ASD, X chromosome; Ped N1, BAP, chromosomes 9 and 20; Supplementary Figure S1), but in neither case was there evidence of CNVs aligning with these signals.

3.2 | Inherited CNVs segregating with ASD/BAP phenotypes

A small number of rare CNVs overlapped known ASD genes and/or recurrent known genomic syndromes (Table 1). One father of an ASD male child, who himself met the criteria for BAP, had a 2.5 Mb clinically pathogenic duplication at 22q11.21 (Wenger et al., 2016), impacting many genes, several, including COMT and SLC25A1, implicated in ASD and other neurodevelopmental disorders. The male offspring of this individual has a diagnosis of ASD and did not inherit this mutation.

At five other loci, CNVs of unknown clinical significance were identified. This included one duplication shared by an ASD male and his unaffected female sibling impacting the ASD-implicated gene NEDD9. This CNV tandemly duplicates three shorter isoforms of NEDD9, but intragenic for the two larger isoforms of this gene. This gene belongs to a family of molecules that mediate protein–protein interactions in signal transduction pathways. This same ASD male also had a tandem duplication that overlapped the cortical dysplasia implicated gene TUBB2A, associated with complex cortical dysplasia and other brain malformations (Romaniello et al., 2018). In Family 3 (Table 1), three members had a hemizygous deletion impacting the ASD-implicated gene AGMO. One ASD female from this same family had a large de novo, 1.6 Mb CNV duplication at Xp22.31, impacting seven genes including STS.

The most striking finding was evidence of a paternally inherited rare tandem duplication impacting the ASD-implicated gene DLGAP2 in three ASD-affected male offspring (Figure 1, individuals 003, 004, and 006). This CNV partially duplicates DLGAP2 and ARHGEF10 and fully impacts CLN8. None of the three male offspring is known to have had neonatal complications or congenital abnormalities, and none has ever had a seizure. For 003, 004, and 006 full-scale IQs are 108, 106, and 96, and Vineland daily living skills are 117, 91, and 97, respectively. These scores all lie in the average range (Table 2). Two of these individuals, 004 and 006, are diagnosed with ADHD, and both have impaired social functioning as indexed by their VABS social-domain scores lying in the impaired range (i.e., <70). The third individual, 003, however, does not have ADHD and his social functioning is 107, lying in the average range.

### Table 1

<table>
<thead>
<tr>
<th>Family</th>
<th>Type</th>
<th>Location (coordinates, hg19)</th>
<th>Size (bp)</th>
<th>Genes</th>
<th>Inheritance</th>
<th>Patency</th>
<th>Genotype</th>
<th>Pattern of familial segregation (ID, sex, diagnosis)</th>
<th>Insertion pathogenicity</th>
<th>Copy number variants (CNVs) of interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ped 4</td>
<td>Dup</td>
<td>6p24.2 (11,182,995–11,360,163)</td>
<td>179,739</td>
<td>NEDD9</td>
<td>unk</td>
<td>VUS</td>
<td>2-0938-003, M, ASD</td>
<td>2-0938-004, F, unaff</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ped 4</td>
<td>Dup</td>
<td>6p25.2 (2,813,733–3,197,897)</td>
<td>384,165</td>
<td>TUBB2A+14 genes</td>
<td>unk</td>
<td>VUS</td>
<td>2-0938-003, M, ASD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ped N1</td>
<td>Del</td>
<td>7p21.2 (15,404,211–15,425,180)</td>
<td>21,418</td>
<td>AGMO</td>
<td>pat</td>
<td>Likely benign</td>
<td>3-0129-101, M, non-ASD; 3-0129-001, M, ASD; 3-0129-011, M, non-ASD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ped 19</td>
<td>Dup</td>
<td>8p23.3 (1,499,963–1,854,917)</td>
<td>354,073</td>
<td>DLGAP2</td>
<td>pat</td>
<td>VUS</td>
<td>2-1005-002, M, BAP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ped 19</td>
<td>Del</td>
<td>16q24.1–q23.3 (83,912,536–85,774,917)</td>
<td>1,899,729</td>
<td>27 genes inc. NECAB2, TAF1C, ATP2C2, FAM92B</td>
<td>unk</td>
<td>VUS</td>
<td>1-0627-007, M, ASD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ped 4</td>
<td>Dup</td>
<td>22q11.21 (18,901,001–21,460,220)</td>
<td>2,559,217</td>
<td>Many loci inc. COMT, SLC25A1, TBX1, CLCN3, CLN8, GNB1L</td>
<td>unk</td>
<td>Pathogenic</td>
<td>2-1005-002, M, BAP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ped N1</td>
<td>Dup</td>
<td>Xq22.3 (1,499,963–1,854,917)</td>
<td>1,685,927</td>
<td>7 genes inc. STS</td>
<td>unk</td>
<td>VUS</td>
<td>3-0395-000, M, ASD</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Genes in bold are ASD-implicated (Yuen et al., 2017). Structure of the rest of duplications are not resolvable because they are flanked by large segmental duplications.

Abbreviations: Del, deletion; DN, de novo; Dup, duplication; pat, paternally inherited; unk, unknown; VUS, variant of uncertain significance.
range. DLGAP2 is translated into a membrane bound protein that is brain expressed and believed to play a role in synapse organization and signaling in neuronal cells.

Individual 007, a male sibling with ASD from this same family has a large 1.9 Mb de novo deletion on 16q23.3–24.1, impacting 27 genes, several of which are ASD-implicated. His early development is notable for motor floppiness, although the Vineland daily living skills domain is consistent with level of functioning in the average range (93) and no comorbidity is described. An IQ score is not available for this individual.

Whole genome sequencing in this family revealed that four of the five siblings also have a maternally inherited heterozygous nonsense variant of MYO1A (chr 12: 57,441,459 G>A, p.R93X). The significance of this is uncertain in light of its frequency in the population (0.28%, https://gnomad.broadinstitute.org/variant/12-57441459-G-A). Two of these siblings also harbor the DLGAP2 mutation, while the other sibling has the 16q23.3–16q24.1 deletion (Figure 1; Table 2). These four with the MYO1A variant individuals all have impairment in social functioning as evidenced by their Vineland results. Of note, neither the MYO1A nor the DLGAP2 mutations generated a linkage signal in this family. No other clinically relevant was identified in these individuals.

4 | DISCUSSION

Our examination of large, multigenerational pedigrees, each with at least three members with ASD, has found no evidence of widely segregating CNVs in these families. While a small number of CNVs potentially implicated in ASD were identified, these did not explain the heritable nature of the disorder in these families. Importantly, these CNVs did not overlap any of the linkage signals previously identified in these families (Woodbury-Smith et al., 2018). As such, other heritable genetic risk remains to be identified. Moreover, the exact nature of how heritable risk and identified CNVs interact with each other and additional factors needs to be more fully evaluated. The possibility of an interaction between two or more genes in raising susceptibility to ASD has previously been argued, but such a mechanism remains to be demonstrated.

One family with four male offspring all diagnosed with ASD was characterized by the presence of shared mutations, none of which neatly segregated with phenotype. One tandem duplication impacting DLGAP2 was identified in three male offspring who inherited this CNV from their father, who is of unknown phenotype. Two of the three males with this CNV have been diagnosed with ADHD, and for two social functioning as measured by the Vineland Adaptive Behavior Schedule, is in the impaired range. This is a strong candidate for ASD as DLGAP2 is strongly brain expressed (https://gtxexportal.org/home/gene/DLGAP2) and is identified as important in synapse development (Rasmussen, Rasmussen, & Silahtaroglu, 2017). Indeed, as a synaptic scaffolding protein, it interacts structurally at the domain level with SHANK1-3 (Rasmussen et al., 2017; Takeuchi et al., 1997) themselves implicated in ASD (Sato et al., 2012). Mouse models have shown that knockout of DLGAP2 impacts social behavior, as well synaptic structural integrity (Jiang-Xie et al., 2014). The potential role of DLGAP2 in ASD is unclear, however, as one study failed to identify FIGURE 1 Pedigree diagram for Ped 19 [Color figure can be viewed at wileyonlinelibrary.com]
statistically significant difference in frequency of rare missense and nonsense mutations in ASD cases versus controls, although the sample size was small: all cases were Asian, having been recruited in Taiwan (Chien et al., 2013). However, one other study has identified ASD and other neurodevelopmental phenotypes in nine non-overlapping families with \textit{DLGAP2} duplications (Poquet et al., 2017).

Four individuals with ASD in a single pedigree also had the same stop-gain point mutation in \textit{MYO1A}. Although IQ information was unavailable for one individual, the other three all had intellectual ability in the average range. Despite this, however, their social functioning was in the impaired to borderline impaired range. Of additional relevance, the two offspring with ADHD both had \textit{DLGAP2} and \textit{MYO1A} mutations. However, other aspects of phenotype, including language skills and medical history, did not clearly differentiate between individuals in this family with different mutations. The one individual with both a \textit{MYO1A} mutation and large \textit{16q23.3–q24.1} deletion was floppy at birth but otherwise unremarkable in terms of perinatal complications and medical history.

We have previously highlighted the complexity of making genotype-phenotype correlations in ASD, particularly in view of the phenotypic heterogeneity and infrequency of individual mutations (Woodbury-Smith et al., 2017). This is compounded by the variable penetrance of these mutations. Although we are not able to confidently make causal statements, overcoming this formidable obstacle will require the accumulation of phenotype and genotype information afforded by in-depth studies such as ours. Such variable penetrance is widely recognized in ASD and other neurodevelopmental disorders (Bateman et al., 2018; Le, Williams, Alaimo, & Elsea, 2019; Ropers & Wienker, 2015), even at "high risk" neurodevelopmental loci. For example, the well-described \textit{2q37} deletion syndrome is associated with a wide range of neurodevelopmental and medical phenotypes, and it is only by considering size and position of CNV along with downstream expression of underlying genes and their interacting partners, that the relationship of genotype to phenotype is more fully appreciated (Le et al., 2019). Moreover, similar to ours, other studies have also found variable penetrance for identical mutations in family members of identified probands (Bateman et al., 2018).

We did identify a number of other CNVs in ASD and/or BAP cases that have been previously implicated in ASD (\textit{COMT}, \textit{SLC25A1}, \textit{NEDD9}, and \textit{AGMO}) and cortical dysplasia (\textit{TUBB2A}; Yuen et al., 2017). The likely variable penetrance of these mutations, however, is evidenced by their occurrence in both BAP cases as well as typically developing siblings of ASD probands. Further work, including functional studies, is therefore required to improve our understanding of their penetrance in relation to ASD. Indeed, it is notable that several of the identified CNVs are interpreted clinically as of unknown significance. This level of uncertainty presents major problems for genetic counselors tasked with explaining causality and risk to families. A further CNV impacting \textit{STS} is associated with X-linked ichthyosis, itself described in association with neurodevelopmental phenotypes including ASD and ADHD (Kent et al., 2008).

### TABLE 2  Phenotypes of Ped 19 (see Figure 1 for pedigree)

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>001</th>
<th>002</th>
<th>003</th>
<th>004</th>
<th>005</th>
<th>006</th>
<th>007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>F</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>ASD/BAP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genetics</td>
<td>355 kb dup 8p23.3; \textit{MYO1A} p. R93X +/−</td>
<td>355 kb dup 8p23.3; \textit{MYO1A} p. R93X +/−</td>
<td>355 kb dup 8p23.3; \textit{MYO1A} p. R93X +/−</td>
<td>355 kb dup 8p23.3; \textit{MYO1A} p. R93X +/−</td>
<td>1.9 Mb dn del 16q23.3–q24.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IQ: FSIQ</td>
<td></td>
<td></td>
<td>108</td>
<td>104</td>
<td>96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRI</td>
<td></td>
<td></td>
<td>112</td>
<td>101</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VCI</td>
<td></td>
<td></td>
<td>103</td>
<td>102</td>
<td>94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adaptive behavior: SSS</td>
<td></td>
<td></td>
<td>107</td>
<td>57</td>
<td>68</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>CSS</td>
<td></td>
<td></td>
<td>98</td>
<td>81</td>
<td>91</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>DSS</td>
<td></td>
<td></td>
<td>117</td>
<td>93</td>
<td>97</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>RBS: Total</td>
<td></td>
<td></td>
<td>43</td>
<td>51</td>
<td>56</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>SCQ</td>
<td></td>
<td></td>
<td>20</td>
<td>17</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OWLS II</td>
<td></td>
<td></td>
<td>106</td>
<td>104</td>
<td>109</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epilepsy</td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>ADHD</td>
<td></td>
<td></td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Congenital abnormalities</td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Neonatal complications</td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Hypotonia</td>
</tr>
</tbody>
</table>

Abbreviations: -, data unavailable; Adaptive behavior (Vineland) SSS, social scaled score; CSS, communication scaled score; DSS, daily living scaled score; OWLS II, oral and written language scales; PRI, perceptual reasoning index; RBS, repetitive behavior score; SCQ, social communication questionnaire; VCI, verbal comprehension index; (Wechsler) FSIQ, full-scale IQ.
Extended pedigrees are a valuable resource for the identification of heritable genetic causes of ASD. Many of the genetic variants identified so far as ASD implicated are de novo and therefore will not explain ASD’s heritability of ~90% (Tick et al., 2016). Families such as the ones studied here often contain members with milder autism phenotypes, termed BAP. We have shown previously that these pedigrees can individually generate reasonable sized linkage signals, with strong evidence that these signals are consistent with segregating uncommon variants.

Moreover, such families offer the opportunity to tease out the relative roles of environmental risk factors that may be shared among affected members. While we do not expect variants to be shared across different families, it is likely that within individual families some shared variants will explain the propensity for clustering of ASD and other mental disorders even in the presence of intrafamilial genetic heterogeneity. Therefore, although the findings are modest, families such as those studied here are a unique resource for future studies tackling the complex genetic etiology of ASD.

ACKNOWLEDGMENTS

The authors thank the families described in this article for supporting this publication. We also thank all other families described herein for their contribution to our ongoing studies of the genetics of ASD. We thank The Centre for Applied Genomics (TCAG), which is funded by Genome Canada and the Ontario Genomics Institute, Canada Foundation for Innovation (CFI), and the Ontario Research Fund of the Government of Ontario. M. W.-S. was supported by a Clinical Investigatorship Award from the Canadian Institutes of Health Research (CIHR) Institute of Genetics during this study; S. W. S. holds the GlaxoSmithKline-CIHR Chair in Genome Sciences at the University of Toronto and The Hospital for Sick Children. Funding for the project was also provided by the Wellcome Trust, Autism Speaks and CIHR. We acknowledge the support of the Province of Ontario Neurodevelopmental Disorders (POND) Network. This work was funded in part by CIHR operating grants #79499 and #89777 and NIH grants MH076028, HD003110 (J. P.), and MH086117 (V. J. V.).

CONFLICT OF INTEREST

S. W. S. is on the Scientific Advisory Committees of Population Bio and Deep Genomics, and intellectual property from his research held at the Hospital for Sick Children is licensed to Athena Diagnostics, and separately Lineagen. These relationships did not influence data interpretation or presentation during this study, but are still being disclosed for potential future considerations. M. W.-S. served as national co-ordinator (UK) for Servier, a pharmaceutical company, in connection with two planned ASD clinical trials. This relationship did not influence data interpretation or presentation. Other authors declared no potential conflict of interests.

AUTHOR CONTRIBUTIONS

P.S., V.V., A.D.P., and S.W.S. conceived the project. M.W.-S. and M.Z. analyzed the data with bioinformatics support from J.W. and B.T. E.A. provided access to family data. I.O.C. and A.T. collected and managed the database of clinical phenotypes respectively. D.J.S., J.D., and B.F. provided clinical genetics input for the interpretation of data. J.L.H. oversaw all laboratory based aspects of the study. M.P. and J.P. provided the US-based data for the linkage analyses, and were part of the research team. M.W.-S. and M.Z. wrote the draft manuscript, with subsequent revisions based on involvement from all listed authors.

ORCID

Marc Woodbury-Smith https://orcid.org/0000-0002-0593-8917
Ryan K. C. Yuen https://orcid.org/0000-0001-7273-4968
Stephen W. Scherer https://orcid.org/0000-0002-8326-1999

REFERENCES


Volume 8, Issue 1, 2020 – 7

Molecular Psychiatry

Page 7


SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section at the end of this article.