A Genetics-First Approach to Dissecting the Heterogeneity of Autism: Phenotypic Comparison of Autism Risk Copy Number Variants

Samuel J.R.A. Chawner, Ph.D., Joanne L. Doherty, Ph.D., Richard J. L. Anney, Ph.D., Kevin M. Antshel, Ph.D., Carrie E. Bearden, Ph.D., Raphael Bernier, Ph.D., Wendy K. Chung, M.D., Ph.D., Caitlin C. Clements, M.A., Sarah R. Curran, Ph.D., M.R.C.Psych., Goran Cuturilo, M.D., Ph.D., Ania M. Fiksinski, M.Sc., Louise Gallagher, M.D., Ph.D., Robin P. Goin-Kochel, Ph.D., Raquel E. Gur, M.D., Ph.D., Ellen Hanson, Ph.D., Sebastien Jacquemont, M.D., Wendy R. Kates, Ph.D., Leila Kushan, M.S., Anne M. Maillard, Ph.D., Donna M. McDonald-McGinn, M.S., L.C.G.C., Marina Mihaljevic, M.D., Ph.D., Judith S. Miller, Ph.D., Hayley Moss, B.Sc., Milica Pejovic-Milovancevic, M.D., Ph.D., Robert T. Schultz, Ph.D., LeeAnne Green-Snyder, Ph.D., Jacob A. Vorstman, M.D., Ph.D., Tara L. Wenger, M.D., Ph.D., IMAGINE-ID Consortium, Jeremy Hall, Ph.D., F.R.C.Psych., Michael J. Owen, Ph.D., F.R.C.Psych., Marianne B.M. van den Bree, Ph.D.

Objective: Certain copy number variants (CNVs) greatly increase the risk of autism. The authors conducted a genetics-first study to investigate whether heterogeneity in the clinical presentation of autism is underpinned by specific genotype-phenotype relationships.

Methods: This international study included 547 individuals (mean age, 12.3 years [SD=4.2], 54% male) who were ascertained on the basis of having a genetic diagnosis of a rare CNV associated with high risk of autism (82 16p11.2 deletion carriers, 50 16p11.2 duplication carriers, 370 22q11.2 deletion carriers, and 45 22q11.2 duplication carriers), as well as 2,027 individuals (mean age, 9.1 years [SD=4.9], 86% male) with autism of heterogeneous etiology. Assessments included the Autism Diagnostic Interview—Revised and IQ testing.

Results: The four genetic variant groups differed in autism symptom severity, autism subdomain profile, and IQ profile.

However, substantial variability was observed in phenotypic outcome in individual genetic variant groups (74%–97% of the variance, depending on the trait), whereas variability between groups was low (1%–21%, depending on the trait). CNV carriers who met autism criteria were compared with individuals with heterogeneous autism, and a range of profile differences were identified. When clinical cutoff scores were applied, 54% of individuals with one of the four CNVs who did not meet full autism diagnostic criteria had elevated levels of autistic traits.

Conclusions: Many CNV carriers do not meet full diagnostic criteria for autism but nevertheless meet clinical cutoffs for autistic traits. Although profile differences between variants were observed, there is considerable variability in clinical symptoms in the same variant.

Am J Psychiatry 2021; 178:77–86; doi: 10.1176/appi.ajp.2020.20010015

Autism is a behaviorally defined condition characterized by deficits in social interaction and communication, as well as the presence of restricted, repetitive behaviors and interests (1). There is considerable heterogeneity in the clinical presentation of autism in terms of symptom profile, cognitive function, and developmental trajectories (2–5). Studies of large genotyped cohorts of autistic individuals and typically developing control subjects have identified several chromosomal copy number variants (CNVs) (deletions and duplications >1 kilobase [kb] [6]) as increasing the risk of autism (7–12) and have been demonstrated in clinical settings to have

predictive value (13). Although individually rare, collectively pathogenic CNVs are identified in 15% of patients with neurodevelopmental disability (14). A number of researchers have suggested that the time is ripe for a reverse strategy based on a genetics-first rather than phenotype-first approach in order to better understand the clinical heterogeneity of autism (15–17).

Deletions and duplications at the 16p11.2 locus (600 kb, breakpoints 4 and 5 [BP4–BP5] critical region 29.6–30.2 Mb, build hg19) and the 22q11.2 locus (3 Mb, breakpoints A and D, critical region 19.0–21.5 Mb, build hg19) have been identified

as increasing the risk of autism, both in phenotype-first studies showing that these variants occur with greater frequency in cohorts of individuals with autism compared with control subjects (7–9) and in genetics-first studies showing that patients diagnosed with 16p11.2 and 22q11.2 CNVs in medical genetics clinics have an elevated frequency of autism diagnosis (18–25) compared with the frequency of 1% in the general population (26, 27). It is important to determine whether these variants lead to the same autism phenotype or whether the presentation differs by genotype. The former would indicate that genomic risk for autism has common phenotypic effects, while the latter would suggest that genetic heterogeneity underpins clinical heterogeneity. In the autism field, there is a strong notion that the condition is dissociable by genetics (28, 29), with some researchers using the term "autisms" (30). Early evidence indicates that the 22q11.2 deletion and duplication may have unique autism profiles (25, 31); however, the profiles of the two groups have not been directly compared in the same study, and hence the differences reported could be due to methodological inconsistencies. For the 16p11.2 locus, it has been reported that duplication carriers with autism have lower IQ compared with deletion carriers with autism (18); however, the autism profiles of the two groups have not been compared. It is also important to investigate the extent to which the autism profile of these variants differs from individuals without these variants who have autism (referred to hereafter as heterogeneous autism).

Comprehensive clinical phenotyping of individuals with genetic variants associated with autism requires large integrated networks of researchers and clinicians using the same clinical instruments. Here, we bring together patient data from several international genetics-first consortia of individuals with rare chromosomal conditions associated with high risk of autism. Individuals with deletions and duplications that span critical regions at the 22q11.2 and 16p11.2 loci were ascertained clinically via medical genetics clinics and patient organizations. We aimed to characterize and contrast the phenotypes of different autism genetic variants in terms of autism prevalence, symptom severity, symptom domain profile, subdomain profile, and IQ and to investigate whether CNV carriers with autism differ in phenotype from individuals with autism of heterogeneous origins.

METHODS

Participants

Genetics-first cohorts. We identified several clinical research sites and consortia that had independently established genetics-first cohorts and utilized the Autism Diagnostic Interview–Revised (ADI-R) (32) to assess autism, thus allowing data to be easily combined. Data on 566 clinically ascertained CNV carriers were available, but 19 case subjects were removed because of insufficient genotypic information (N=18) and cohort overlap (N=1). This resulted in 547 CNV carriers (mean age, 12.3 years [SD=4.2], 54% male). Eighty-

two case subjects had 16p11.2 deletion, 50 had 16p11.2 duplication, 370 had 22q11.2 deletion, and 45 had 22q11.2 duplication. Data on these case subjects were obtained from the ECHO study (Experiences of People With Copy Number Variants), the IMAGINE-ID study (Intellectual Disability and Mental Health: Assessing Genomic Impact on Neurodevelopment), the neurodevelopmental CNV cohort at Belgrade University Children's Hospital, the International 22q11.2DS Brain and Behavior Consortium, the Center for Autism Research at Children's Hospital of Philadelphia, the 16p11.2 European consortium, and the Simons Variation in Individuals Project Consortium (for further details, see Table S1 in the online supplement). Demographic and genetic characteristics by continent are presented in Table 1. The characteristics of these studies have been described elsewhere (18, 24, 25, 33–37).

Carrier status for CNVs at the 16p11.2 locus (critical region 29.6–30.2 Mb; spanning breakpoints 4–5; build hg19) and the 22q11.2 locus (critical region 19.0–21.5 Mb; spanning at least low copy repeat regions A–B because pathogenicity of atypical variants outside the A–B region is uncertain; build hg19) was confirmed for all individuals through clinical chromosome microarrays, medical records, or confirmation in a research laboratory (for information on the full genotype, see Table S2 in the online supplement). Analysis included individuals ≥4 years old. The study was approved by the appropriate local ethics committees and institutional review boards. Before recruitment, written consent or assent was obtained from each participant and, where appropriate, his or her caregiver.

Heterogeneous autism cohort. Data on 2,053 autistic individuals, age ≥4 years, were accessed from the Autism Genome Project (38). These individuals were ascertained via autism diagnostic clinics. Of these 2,053 individuals, 26 had CNVs at the 16p11.2 locus or 22q11.2 locus. Seven individuals had 16p11.2 deletion, four had 16p11.2 duplication, four had 22q11.2 deletion, and 11 had 22q11.2 duplication. Given the small sample sizes, we did not compare these groups with the remainder of the Autism Genome Project cohorts. Additionally, previous studies have reported on the phenotype of CNV carriers in the Autism Genome Project cohort (39). These individuals were not included in the genomic condition groups given the different ascertainment strategies. The remaining 2,027 individuals represent a subgroup with autism for whom the underlying etiology is heterogeneous (Table 1). Following previous investigators (17), we refer to this subgroup as a heterogeneous autism cohort rather than an idiopathic autism cohort.

Autism Assessment

All individuals were assessed with the Autism Diagnostic Interview–Revised (ADI-R) (32) by a research reliable assessor (for further information on assessors and assessment sites, see Table S3 in the online supplement). The ADI-R is a semi-structured interview conducted with the primary

caregiver about a child's symptoms both currently and during early development. The total ADI-R score was used as an index of autism symptom severity (37). Autism domain scores for social interaction, communication, and restricted, repetitive, and stereotyped behaviors (RRBs) were extracted, as well as autism subdomain scores (for further details on ADI-R scores, see the online supplement). To meet autism criteria on the ADI-R, an individual had to meet the clinical cutoff score on each domain (cutoff score for social interaction, 10; communication, 8 [7 for non-verbal communication]; and RRB, 3), and there must also have been evidence of developmental abnormality before age 36 months.

Cognitive Assessment
Scores for full-scale IQ, verbal IQ, and performance IQ were derived from age and developmentally appropriate standardized IQ measures as described elsewhere (18, 24, 33, 35, 39).

Statistical Analysis
Autism prevalence in genetic variant groups. Autism prevalence was determined on the basis of the ADI-R diagnostic algorithm (40). A logit mixed model was performed to determine whether genetic variant group (22q11.2 deletion, 22q11.2 duplication, 16p11.2 deletion, and 16p11.2 duplication) was a predictor of autism diagnosis, while accounting for gender and age. Following previous international studies of the 16p11.2 duplication, we included site (Europe compared with United States) as a covariate (18, 36). Post hoc contrasts were conducted to establish autism prevalence differences between genetic variant groups with Tukey's adjustment for multiple comparisons. The percentage of individuals who did not meet autism criteria but did meet the clinical cutoff criteria in one or more domains was additionally calculated.

Autism profiles between genetic variant groups. To investigate possible differences in autism profiles between genetic variant groups, a series of analysis of covariance models was conducted with group as a predictor and the following phenotypic variables as outcome measures: ADI-R total score as an index of the clinical severity of autism symptoms, autism domain profile, autism subdomain profile, and IQ profile, accounting for gender, age, and site (for further details, see the online supplement). Tukey's method was used to conduct post hoc contrasts between genetic variant groups, producing p values adjusted for the number of contrasts. Eta-squared values were calculated to estimate the proportion of variance explained by genetic variant group (between-group differences). We also calculated the variance that was explained by variable expressivity in the four genetic variant groups (i.e., variance not explained by genetic variant group, age, gender, and site).

TABLE 1. Demographic characteristics of a cohort with a genetic diagnosis of a copy number variant associated with autism (N=547) and a cohort of heterogeneous autism (N=2,027)

Genetics-First Cohorts	Sample Size	Male		Age (years)		Autism	
		N	%	Mean	SD	N	%
16p11.2 deletion	82	43	52	9.6	3.7	35	43
Europe	12	7	58	10.3	4.4	7	58
United States	70	36	51	9.5	3.6	28	40
16p11.2 duplication	50	35	70	10.8	6.7	29	58
Europe	17	16	94	12.6	8.6	13	76
United States	33	19	58	9.9	5.4	16	48
22q11.2 deletion	370	182	49	13.4	3.4	85	23
Europe	215	102	47	13.7	2.4	50	23
United States	155	80	52	12.9	4.4	35	23
22q11.2 duplication	45	35	78	10.1	4.3	20	44
Europe	11	9	82	10.3	2.8	6	55
United States	34	26	76	10.0	4.7	14	41
Heterogeneous autism	2,027	1,753	86	9.1	4.9	2,027	100
Europe	848	731	86	8.6	4.8	848	100
United States	1,179	1,022	87	9.5	5.0	1179	100

Analyses of the ADI-R total score and domain and subdomain scores were repeated, including full-scale IQ as a covariate, to investigate whether differences in autism phenotype were driven by full-scale IQ.

To compare autism in the genetic variant groups with heterogeneous autism (i.e., individuals from the Autism Genome Project data set who did not have 16p11.2 or 22q11.2 CNVs; N=2,027), we conducted analyses excluding individuals in the genetic variant groups who did not meet ADI-R criteria for autism and compared the profiles with individuals with heterogeneous autism. This resulted in five groups: 16p11.2 deletion plus autism, 16p11.2 duplication plus autism, 22q11.2 deletion plus autism, 22q11.2 duplication plus autism, and heterogeneous autism (Table 1).

Multivariate analysis of covariance was conducted with group as a predictor and phenotypic scores as the outcomes, while accounting for gender, age, and site. Here too, analyses were run for autism symptom severity (ADI-R total score), autism domain profile, autism subdomain profile, and IQ profile. Post hoc contrasts to investigate the difference between the genetic variant plus autism groups and the heterogeneous autism group were conducted with Tukey's adjustment for multiple comparisons.

To investigate whether male-to-female ratios differed between the five groups, we used a logit model, with gender as a binary outcome and group as a predictor, while taking into account fixed effects of age and the random effect of site.

For both analyses, a Benjamini-Hochberg false discovery rate multiple testing correction value of 0.05 was applied to p values.

RESULTS

Autism Prevalence in Genetic Variant Groups
In our cohort of CNV carriers ascertained clinically via medical genetics clinics and patient organizations, 43% of individuals with 16p11.2 deletion, 58% of individuals with 16p11.2 duplication, 23% of individuals with 22q11.2 deletion,

TABLE 2. Comparison of genetic variant groups on IQ and autism measures[a]

Measure	16p11.2 Deletion			16p11.2 Duplication			22q11.2 Deletion			22q11.2 Duplication			Between-Group Variation		Within-Group Variation η^2 (%)
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	p^b	η^2 (%)	
Cognitive scores															
FSIQ	80	81.3	15.4	50	70.9	22.4	337	70.3	13.5	32	88.1	21.1	**<0.001**	11.7	82.2
VIQ	80	78.2	18.6	49	73.6	24.4	329	72.8	13.9	32	84.2	18.9	**<0.001**	3.7	91.7
PIQ	81	85.7	15.2	50	71.9	22.4	333	72.3	14.3	32	87.1	21.7	**<0.001**	11.8	84.8
ADI-R scores															
Total score	82	22.0	11.8	50	27.7	14.0	370	16.7	11.9	45	21.2	15.4	**<0.001**[c]	7.4	88.7
Social domain	82	12.9	7.5	50	15.8	8.0	370	10.1	7.2	45	12.4	9.5	**<0.001**[c]	5.4	91.3
Subdomains															
Social interaction	82	2.4	1.9	50	2.9	2.0	370	2.2	2.0	45	2.5	2.2	0.120	1.0	96.6
Peer relationships	82	4.2	2.3	50	4.9	2.6	370	3.5	2.4	45	3.8	2.8	**<0.001**[c]	3.1	94.2
Shared enjoyment	82	2.5	2.2	50	3.2	2.0	370	1.7	1.8	45	2.5	2.4	**<0.001**[c]	5.7	90.1
Socioemotional reciprocity	82	3.9	2.5	50	4.8	2.5	370	2.7	2.2	45	3.7	2.9	**<0.001**[c]	7.9	89.8
Communication domain	82	5.8	3.7	50	6.5	4.9	370	4.5	3.8	45	5.4	4.3	**<0.001**[c]	3.1	93.6
Subdomains															
Gestures	82	2.4	2.5	50	3.1	2.9	370	1.7	2.3	45	2.7	2.6	**<0.001**[c]	3.6	94.7
Imagination and imitation	82	3.5	1.8	50	3.4	2.3	370	2.8	2.0	45	2.7	2.1	**0.009**[c]	2.0	92.2
RRB domain	82	3.3	2.2	50	5.3	2.6	370	2.1	2.3	45	3.3	2.9	**<0.001**[c]	14.7	81.7
Subdomains															
Unusual interests	82	0.8	1.0	50	1.6	1.1	370	0.8	1.0	45	0.9	1.0	**<0.001**[c]	4.2	92.8
Routines and rituals	82	0.5	0.9	50	1.1	1.3	370	0.5	1.0	45	0.8	1.1	**<0.001**[c]	3.2	96.2
Motor mannerisms	82	0.9	0.9	50	1.2	0.9	370	0.2	0.5	45	0.8	0.9	**<0.001**[c]	20.9	77.8
Sensorimotor interests	82	1.2	0.8	50	1.5	0.7	370	0.5	0.7	45	0.8	0.9	**<0.001**[c]	18.8	74.0

[a] ADI-R=Autism Diagnostic Interview; FSIQ=full-scale IQ; PIQ=performance IQ; RRB=restricted, repetitive, and stereotyped behaviors; VIQ=verbal IQ. Separate multivariate analysis of covariance analyses were conducted for domain and subdomain scores to avoid including mathematically related scores in the same analysis. Total ADI-R score was analyzed separately using an analysis of covariance model. Age, gender, and site were included as covariates. Post hoc contrasts are presented in Table S7 in the supplement.
[b] Bold indicates that the p value was significant after Benjamini-Hochberg false discovery rate 0.05 correction.
[c] Significant after correcting for IQ (full results are presented in Table S3 in the supplement).

and 44% of individuals with 22q11.2 duplication met ADI-R criteria for autism (Table 1). Genetic variant group was a significant predictor of autism diagnosis (p<0.001). Post hoc contrasts revealed that autism prevalence in the 22q11.2 deletion carrier group (23%) was significantly lower compared with the 16p11.2 deletion (43%, p=0.004) and 16p11.2 duplication (58%, p<0.001) groups. The remaining genetic variant group differences were not significant.

In CNV carriers who did not meet criteria for a formal autism diagnosis, we examined the proportion who met clinical cutoff criteria for one or more domains on the ADI-R. Among the 378 of 547 (69%) individuals who did not meet criteria for autism, 205 (54%) were found to meet the clinical cutoff for at least one domain, indicating a significant domain-based impairment (N=38/47 [81%] of 16p11.2 deletion carriers, N=19/21 [90%] of 16p11.2 duplication carriers, N=135/285 [47%] of 22q11.2 deletion carriers, and N=13/25 [52%] of 22q11.2 duplication carriers). For each CNV, the proportion of individuals who met the clinical cutoff scores for each domain is presented in Table S4 and Figure S1 in the online supplement.
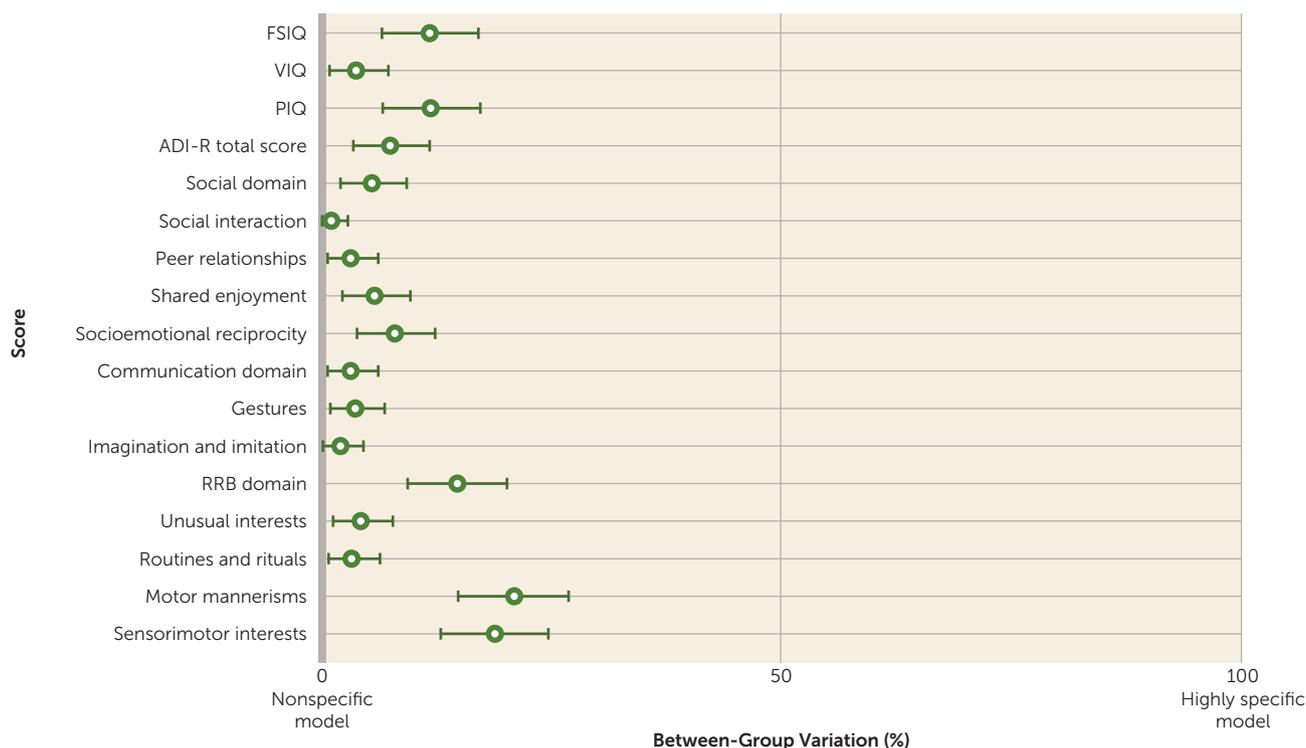
Autism Profiles Between Genetic Variant Groups
Genetic variant group predicted autism symptom severity (7% of the variance, p<0.001), autism domain profile (5% of the variance, p<0.001), and autism subdomain profile (1% of the variance, p<0.001). In terms of individual domain scores, genetic variant group predicted 5% of the social domain total score, 3% of the communication domain score, and 15% of the RRB domain (Table 2, Figure 1). For subdomain scores, the proportion of variance predicted by genetic variant group varied between 1% (social interaction) and 21% (motor mannerisms). In addition to motor mannerisms, the proportion of variance explained was also high for sensorimotor interests (19%). Genetic variant group predicted 12% of variance in full-scale IQ (p<0.001), 12% of variance in performance IQ (p<0.001), and 4% of variance in verbal IQ (p<0.001). Findings for autism symptom severity, domain scores, and subdomain scores remained significant after controlling for full-scale IQ, and the eta-squared values remained relatively unchanged (see Table S5 in the online supplement). Age accounted for 0%−3% of variance in phenotypic traits (see Table S6 in the online supplement).

After accounting for between-group variability, age, gender, and site, a large proportion of variability remained: 74%–97% within-group variability, depending on trait (Table 2). This is illustrated in Figure 1 and in Figure S2 in the online supplement, which show that although group differences were present, there was much more variability within

FIGURE 1. Score variability between genetic variant groups in a cohort with a genetic diagnosis of a copy number variant associated with autism[a]

[a] This plot illustrates the between–genetic variant group variation data (as presented in Table 2). Between-group eta-squared values are plotted on a scale of 0% to 100% of the variance. These values represent the proportion of variation in phenotypic outcome predicted by genetic variant group. A value close to 0% indicates a nonspecific model whereby different genotypes lead to similar phenotypic outcomes. A value close to 100% indicates a highly specific model whereby different genotypes lead to different and discrete phenotypic outcomes. Error bars indicate 95% confidence intervals. FSIQ=full-scale IQ; PIQ=performance IQ; RRB=restricted, repetitive, and stereotyped behaviors; VIQ=verbal IQ.

groups across traits. For IQ, we found greater variability for duplications than deletions for both the 16p11.2 locus (Levene's test, p=0.001) and the 22q11.2 locus (Levene's test, p<0.001). For autism symptom severity, we found greater variability in outcome for duplications than deletions for the 22q11.2 locus (Levene's test, p<0.001) but not for the 16p11.2 locus (Levene's test, p=0.071).

Post hoc Tukey contrasts between groups that were significant (p values adjusted for multiple contrasts) are summarized in Table S7 in the online supplement. To briefly summarize phenotypic profiles, 16p11.2 deletion carriers had relatively moderate autism symptom severity scores and moderate cognitive impairment (IQ=81.3), 16p11.2 duplication carriers had relatively greater autism symptom severity scores and greater cognitive impairment (IQ=70.9), 22q11.2 deletion carriers had relatively lower autism symptom severity scores but greater cognitive impairment (IQ=70.3), and 22q11.2 duplication carriers had relatively higher autism symptom severity scores but less cognitive impairment (IQ=88.1).

Mean scores for each phenotypic trait for each of the five groups (heterogeneous autism, 16p11.2 deletion plus autism, 16p11.2 duplication plus autism, 22q11.2 deletion plus autism, and 22q11.2 duplication plus autism) are presented in Table S8 in the online supplement. With the exception of verbal IQ

and the routines and rituals domain, all phenotypic traits and subdomains were found to differ between the five groups (see Table S8 in the online supplement). These findings remained significant after a Benjamini-Hochberg false discovery rate correction of 0.05 for multiple testing. Age accounted for 0% −5% of variance in phenotypic traits (see Table S9 in the online supplement). The profile of each genetic variant plus autism group is shown in Figure 2A, and the profile of each genetic variant plus autism group compared with the heterogeneous autism group is shown in Figure 2B. Aspects of the phenotypic profile showing significant contrasts between the heterogeneous autism group and the genetic variant groups are presented in Table S8 in the online supplement.

To briefly summarize phenotypic profile differences relative to the heterogeneous autism group, the 16p11.2 deletion plus autism group had relatively lower autism symptom severity but had a similar level of cognitive impairment, the 16p11.2 duplication plus autism group had greater performance IQ deficits but did not differ on any of the other phenotypic measures, the 22q11.2 deletion plus autism group had greater cognitive impairment but relatively lower autism symptom scores, and the 22q11.2 duplication plus autism group did not differ significantly from the heterogeneous autism group on any phenotypic measure.

FIGURE 2. Domain profiles of the genetic variant plus autism groups[a]
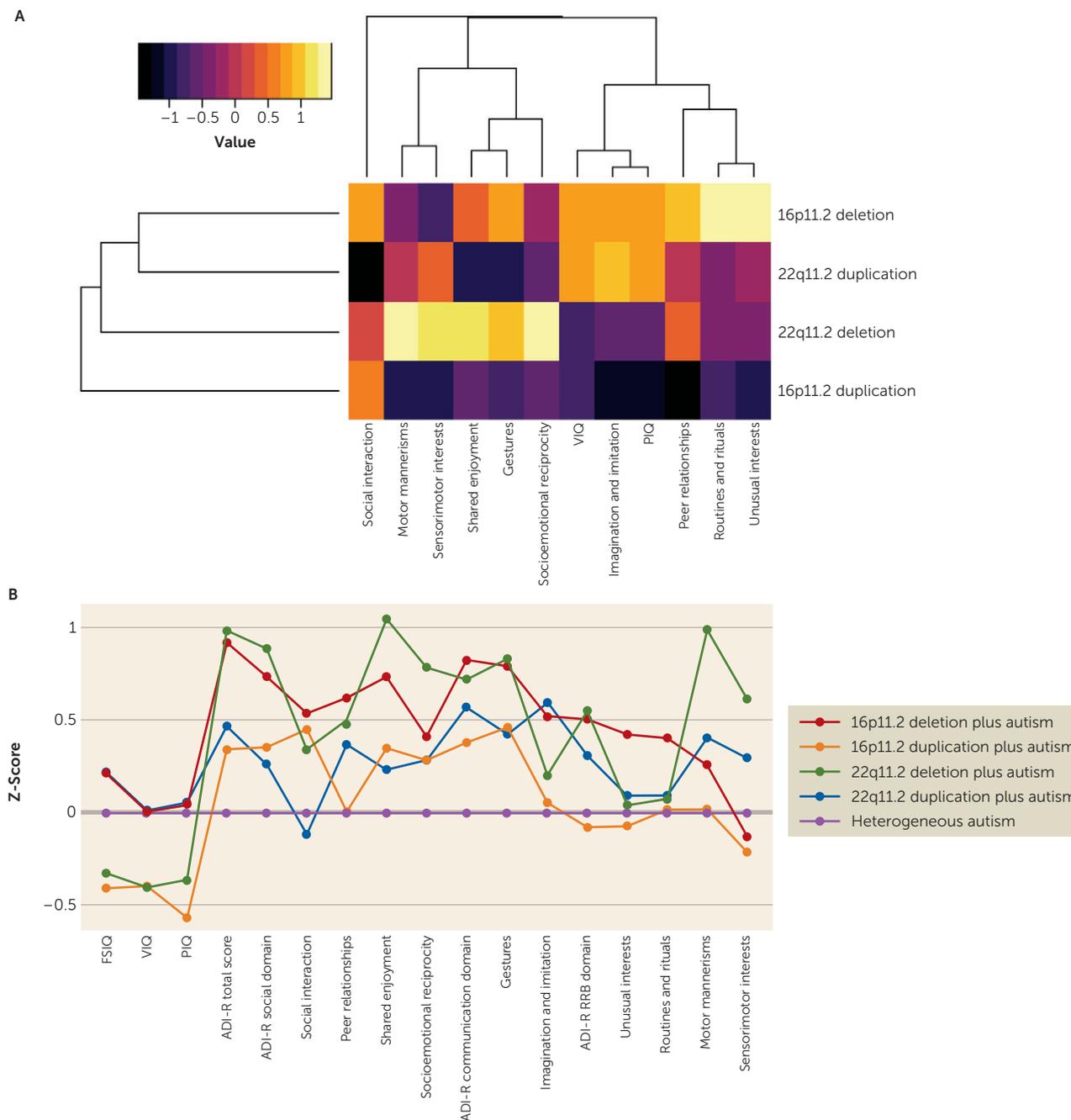
[a] Panel A illustrates how each genetic variant plus autism group differed; a heat map plot was generated by transforming IQ and Autism Diagnostic Interview–Revised (ADI-R) scores of the genetic variant plus autism groups to z-scores. Dendrograms showing the clustering of copy number variants (CNVs) and phenotypes were generated using Ward's method and Euclidian distance. Scores for each genetic variant plus autism group were standardized into z-scores relative to each other and were adjusted for gender, age, and site. The z-scores were constructed so that a negative score always denoted a poorer performance. Black denotes a relative deficit in that neuropsychiatric domain compared with other CNV carriers, and yellow represents a relative strength compared with other CNV carriers. Panel B illustrates the profiles of each genetic variant plus autism group relative to the heterogeneous autism group, whereby phenotypic scores were standardized to z-scores using the mean and standard deviation of the heterogeneous autism group as a reference (i.e., the difference in the individual's score and the mean score for the entire heterogeneous autism group was divided by the standard deviation for the heterogeneous autism group). The z-scores were adjusted for gender, age, and site. The z scores were constructed so that a negative score for a CNV carrier denoted a worse outcome. FSIQ=full-scale IQ; PIQ=performance IQ; VIQ=verbal IQ.

Gender

Male CNV carriers (all groups combined) were at increased risk of autism (odds ratio=2.3, p<0.001) compared with female CNV carriers. However, male-to-female ratios were lower in CNV carriers with autism (2.3:1) compared with ratios in the heterogeneous autism group (6.4:1) (p<0.001).

DISCUSSION

This study is the result of a collaboration between several international genetics-first consortia and the Autism Genome Project. The availability of a large sample of individuals with one of four autism CNVs allowed us to use a genetics-first approach, which meant that we were not constrained by ascertaining patients on the basis of autism diagnosis, allowing examination of the impact of genotype on autism symptom severity and domain profiles across the spectrum. The use of the widely accepted ADI-R clinical research instrument across all sites represents a methodological strength, enabling us to directly compare the autism profiles of 22q11.2 and 16p11.2 CNVs. Our findings indicate that although genetic variants associated with autism differ in several aspects of the autism phenotype, including autism symptom severity, symptom domain profile, and cognitive profile, only 1%−21% of the variance is explained by genetic variant group, depending on autism trait. In contrast, variation in each of the four genetic variant groups is much greater, explaining between 74%−97% of the variability, depending on autism measure. This highlights the fact that even in individuals with the same genetic variant, the autism profile is difficult to predict on the basis of CNV alone and that phenotypic profiles overlap, providing evidence against a highly specific model (41), whereby each genotype leads to a unique autism phenotype (see Figure S3 in the online supplement); instead, our findings support a partially specific model whereby autism profiles are distinct but overlapping.

The risk of autism differed by genetic variant group. In terms of autism prevalence, fewer 22q11.2 deletion carriers met criteria for autism (23%) compared with 22q11.2 duplication carriers (44%), 16p11.2 deletion carriers (43%), and 16p11.2 duplication carriers (58%). These figures represent autism prevalence in a clinically ascertained cohort of CNV carriers and should not be taken as the prevalence for CNV carriers in the wider population. Among CNV carriers with an autism diagnosis, we found that 22q11.2 deletion and 16p11.2 deletion carriers with an autism diagnosis had relatively less severe symptom profiles compared with the heterogeneous autism group. On the other hand, individuals with 16p11.2 duplication and 22q11.2 duplication with an autism diagnosis had a profile more consistent with the heterogeneous autism group. Our findings complement genome-wide CNV studies showing that the strength of association and penetrance for autism varies by genetic variant, in particular that the association of 22q11.2 deletion is weaker relative to the other three CNVs (8).

We found evidence that the four genetic variant groups were associated with differences in autism symptom severity, the three autism domains, nine of the 10 subdomains we studied, full-scale IQ, verbal IQ, and performance IQ. However, the proportion of variance explained by genetic variant group for each subdomain varied between 1% and 21%. It was only the social interaction subdomain that did not differ, indicating that this trait was a universal aspect of autism across the four genetic variant groups. The subdomains for which genetic variant group explained the greatest proportion of variance were motor aspects of the RRB domain, motor mannerisms (21%), and sensorimotor interests (19%), indicating that genetic variant group particularly distinguishes motor aspects of the autism phenotype.

Cognitive profile was also influenced by genetic variant group: 22q11.2 deletion and 16p11.2 duplication carriers had greater cognitive impairments in full-scale IQ, verbal IQ, and performance IQ compared with 22q11.2 duplication carriers and 16p11.2 deletion carriers. There was evidence at both the 22q11.2 locus and 16p11.2 locus that cognitive outcomes were more variable for duplication carriers than deletion carriers. This was previously reported for 16p11.2 duplication carriers (18), and our findings indicate that the same may be true for the 22q11.2 locus. Autism symptom severity of a genetic variant did not covary with the magnitude of cognitive deficit. The 22q11.2 duplication carriers had the highest mean IQ (88.1) of the CNV groups, yet had high symptom severity scores. The 22q11.2 deletion carriers had the greatest cognitive impairment, yet had lower risk of autism compared with individuals with the other genetic variants. Additionally, when we controlled for IQ, differences in autism domain and subdomain scores between CNVs remained relatively unchanged. These findings suggest that the mechanisms underlying autism and cognitive impairment are at least partially distinct among carriers of pathogenic CNVs.

However, although specific group differences exist, it is clear that phenotypic profiles overlap (see Figure S2 in the online supplement), and we found greater variability between individuals with the same CNV than between CNVs. Overall, our findings provide most support for a partially specific model, whereby autism profiles are distinct but highly overlapping, although the magnitude of these differences is closer to the nonspecific effect end of the scale, whereby all genotypes lead to similar autism phenotypes, than the highly specific effect end of the scale, whereby genotypes lead to discrete autism subtypes (Figure 1). These findings highlight the fact that it will be important for behavioral phenotyping research to move beyond a focus on average differences between variants and to investigate the genetic (including additional rare variants and polygenic risk, which we were not able to analyze in this study) and environmental factors that contribute to variation in clinical phenotypes. There is already evidence that family background is important to consider in a genetic counseling context. Parental IQ has been found to predict the IQ impairment in 16p11.2 and 22q11.2 deletion carriers (42–44).

There was a male preponderance for autism across all genetic variant groups, and gender significantly influenced domain and subdomain profiles. However, the male-to-female ratio among CNV carriers was approximately 2.3:1, which is considerably less pronounced than in the heterogeneous autism group (6.4:1). It may be that the genetic variants we studied have such a large effect on neurodevelopment that they partially override the protective effect of being female (9, 11). Age did influence phenotypic traits; however, the proportion of variance age explained in our analyses was low (≤5%).

By using a genetics-first approach, we identified a significant proportion of CNV carriers (54%) who did not meet autism criteria but did meet clinical cutoff criteria for diagnosis-related impairments. Furthermore, the profile of CNV carriers with autism does, to some extent, present differently from heterogeneous autism (Figure 2B). This has the potential implication that the clinical needs of patients with genomic conditions may be overlooked because they do not meet diagnostic criteria despite exhibiting a range of impairments across domains. Parents of children with CNVs at 16p11.2 or 22q11.2 who have taken part in our studies in the United Kingdom have anecdotally reported that their child's genetic diagnosis can be a barrier to receiving an autism diagnosis and support, with some service providers having stated that a child with a genetic diagnosis cannot also have a secondary diagnosis of autism despite DSM-5 specifying that autism can be diagnosed when "associated with a known medical or genetic condition or environmental factor" (1). It is important that clinicians are aware of the likelihood of autism associated with certain genetic variants to improve the chances that these children will receive an early diagnosis and gain access to clinical and educational support.

Further clinical implications arise from our finding that there are not highly specific genotype-phenotype relationships between individual CNVs and autism, at least for 16p11.2 and 22q11.2 deletion and duplication variants. This indicates that although CNVs are presymptomatically predictive of autism and therefore can inform opportunity for early clinical and educational support, individual genotypes are not specific in predicting symptom subtypes. Rather, our findings indicate an overlap in clinical phenotypes between these CNVs, suggesting that neurodevelopmental service provision for different CNVs could be grouped together. Our genetics-first approach reveals great variability in CNV groups, highlighting the fact that autism-associated variants are not deterministic for autism. It is important in genomic counseling that pathogenic CNVs are considered as one factor within a broader biopsychosocial context, rather than being the only causative factor for autism. Identification of genetic and environmental modifiers of phenotypes of autism CNVs has potential for informing clinical care and management.

Our study benefits from several features, including a large sample size by combining data from individuals with these rare genetic conditions from a number of international cohorts and synchronization of phenotyping measures across sites allowing for analysis extending beyond categorical diagnosis, allowing for autism domains and subdomains to be analyzed. However, there are potential limitations. First, ascertainment bias needs to be considered, because our study focused on individuals who received a clinical genetic diagnosis, and our findings therefore do not necessarily extend to individuals with these CNVs in the population who are affected below a clinical threshold and as a consequence are not referred for genetic testing. Because one of the main indications for genetic testing currently is often developmental delay (18, 25), our findings may not be representative for individuals with these CNVs with a more typical developmental pattern. However, despite these ascertainment considerations, not all CNV carriers in this study met autism criteria or had cognitive impairment, thus allowing us to study the impact of genotype across a broad spectrum of abilities. Another source of possible ascertainment bias is that referral reason for genetic testing may differ by genetic variant. For example, it has been reported that 22q11.2 deletion carriers are more likely to be referred as a result of physical abnormalities, such as heart defects, compared with 22q11.2 duplication carriers, who are more likely to be referred for developmental reasons (25). However, this may actually reflect true phenotypic differences, as shown in a recent population-based study that was able to identify individuals in the population undiagnosed with a 22q11.2 CNV, as well as individuals with a diagnosis through a clinic, and reported a higher frequency of congenital abnormalities in the deletion carriers (45). Before taking part in the study, individuals in our cohorts had a variety of diagnostic experiences; some had a preexisting autism diagnosis before the ADI-R assessment, while others had had no interaction with autism diagnostic services. This potentially introduced caregiver reporter bias, but this was partly mitigated by the semistructured nature of the ADI-R. That is, although the ADI-R interview is based on caregiver report, the scoring of a particular trait is based on concrete descriptions coded by a trained interviewer. We were not able to assess cross-site reliability of ADI-R administration because it was not preplanned that ADI-R data would be combined across several international sites; however, all assessors underwent ADI-R formal training and were research reliable. Finally, we were not able to control for ethnicity and socioeconomic and environmental factors, because these data were not available at all sites, or they were not internationally comparable. Future studies would benefit from greater alignment of measurement of environmental factors across international sites.

CONCLUSIONS

The genetics-first approach we employed represents a novel method for investigating genotype-phenotype relationships unconstrained by categorical diagnostic criteria. We found that the phenotypic profiles of 16p11.2 and 22q11.2 CNVs differ in terms of symptom severity, symptom profile, and

cognitive profile. However, although genetic variants have specific effects, within-variant variability is much greater than between-variant variability, thus indicating that the phenotypic consequences of genomic risk factors for autism fit a partially specific model rather than a highly specific model. It will be important that future studies of autism variants consider the genetic and environmental factors that contribute to clinical variability in variant carriers. An important message from our study is that individuals with genomic conditions are likely to present with clinically significant symptoms of autism but not meet diagnostic criteria. Clinical services need to adapt because otherwise individuals without a formal autism diagnosis are unlikely to access necessary clinical and educational support.

AUTHOR AND ARTICLE INFORMATION

MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, United Kingdom (Chawner, Doherty, Anney, Moss, Hall, Owen, van den Bree); Cardiff University Centre for Human Developmental Science, School of Psychology, Cardiff University, Cardiff, United Kingdom (Chawner); Department of Psychology, Syracuse University, Syracuse, N.Y. (Antshel); Department of Psychiatry and Behavioral Sciences and Psychology, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles (Bearden, Kushan); Departments of Psychiatry and Behavioral Sciences, University of Washington, Seattle (Bernier); Departments of Pediatrics and Medicine, Columbia University Irving Medical Center, New York (Chung); Center for Autism Research, Children's Hospital of Philadelphia (Clements, Miller, Schultz); Department of Psychology, University of Pennsylvania, Philadelphia (Clements); South West London and St. George's Mental Health National Health Service Foundation Trust, London (Curran); University Children's Hospital, Belgrade, Serbia; Faculty of Medicine, University of Belgrade, Serbia (Cuturilo); Department of Psychiatry, Brain Center, University Medical Center Utrecht, the Netherlands (Fiksinski, Vorstman); Clinical Genetics Research Program, Centre for Addiction and Mental Health, and the Dalglish Family 22q Clinic for 22q11.2 Deletion Syndrome, Toronto General Hospital, University Health Network, Toronto (Fiksinski); Department of Psychiatry, Trinity College Dublin, Ireland (Gallagher); Department of Pediatrics, Baylor College of Medicine, Houston (Goin-Kochel); Lifespan Brain Institute, Children's Hospital of Philadelphia and Perelman School of Medicine, University of Pennsylvania, Philadelphia (Gur); Department of Psychiatry, Neurodevelopment and Psychosis Section, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Gur); Department of Child and Adolescent Psychiatry, Children's Hospital of Philadelphia (Gur, Schultz); Developmental Medicine, Children's Hospital Boston/Harvard Medical School, Boston (Hanson); Department of Pediatrics, University of Montreal (Jacquemont, Maillard); Centre de recherche, Centre Hospitalier Universitaire Sainte Justine, Montreal (Jacquemont); Department of Psychiatry and Behavioral Sciences, State University of New York at Upstate Medical University, Syracuse (Kates); Service des Troubles du Spectre de l'Autisme et Apparentés, Lausanne University Hospital, Lausanne, Switzerland (Maillard); Division of Human Genetics and Molecular Biology, Children's Hospital of Philadelphia (McDonald-McGinn); 22q and You Center, Children's Hospital of Philadelphia (McDonald-McGinn); Department of Pediatrics, University of Pennsylvania, Philadelphia (McDonald-McGinn, Schultz); Clinic for Psychiatry, Clinical Center of Serbia, Faculty of Medicine, University of Belgrade, Serbia (Mihaljevic); Institute of Mental Health, Belgrade, Serbia, Faculty of Medicine, University of Belgrade, Serbia (Pejovic-Milovancevic); Simons Foundation, New York (Green-Snyder); Program in Genetics and Genome Biology, SickKids Research Institute, Toronto (Vorstman); Department of Psychiatry, Hospital for Sick Children, University of Toronto (Vorstman); Department of Pediatrics, Seattle Children's Hospital, Seattle (Wenger); and Neuroscience and Mental Health Research Institute, Cardiff University, Cardiff, United Kingdom (Hall, Owen, van den Bree).

Send correspondence to Prof. van den Bree (vandenbreemb@cardiff.ac.uk).

REFERENCES

1. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Washington, DC, American Psychiatric Association, 2013
2. Charman T, Pickles A, Simonoff E, et al: IQ in children with autism spectrum disorders: data from the Special Needs and Autism Project (SNAP). Psychol Med 2011; 41:619–627
3. Fountain C, Winter AS, Bearman PS: Six developmental trajectories characterize children with autism. Pediatrics 2012; 129: e1112–e1120
4. Kim SH, Macari S, Koller J, et al: Examining the phenotypic heterogeneity of early autism spectrum disorder: subtypes and short-term outcomes. J Child Psychol Psychiatry 2016; 57:93–102
5. Eaves LC, Ho HH, Eaves DM: Subtypes of autism by cluster analysis. J Autism Dev Disord 1994; 24:3–22
6. Lee C, Scherer SW: The clinical context of copy number variation in the human genome. Expert Rev Mol Med 2010; 12:e8

7. Pinto D, Pagnamenta AT, Klei L, et al: Functional impact of global rare copy number variation in autism spectrum disorders. Nature 2010; 466:368–372

8. Malhotra D, Sebat J: CNVs: harbingers of a rare variant revolution in psychiatric genetics. Cell 2012; 148:1223–1241

9. Pinto D, Delaby E, Merico D, et al: Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. Am J Hum Genet 2014; 94:677–694

10. Sanders SJ, Ercan-Sencicek AG, Hus V, et al: Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. Neuron 2011; 70:863–885

11. Levy D, Ronemus M, Yamrom B, et al: Rare de novo and transmitted copy-number variation in autistic spectrum disorders. Neuron 2011; 70:886–897

12. Sanders SJ, He X, Willsey AJ, et al: Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. Neuron 2015; 87:1215–1233

13. 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

14. Miller DT, Adam MP, Aradhya S, et al: Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. Am J Hum Genet 2010; 86:749–764

15. Vorstman JAS, Parr JR, Moreno-De-Luca D, et al: Autism genetics: opportunities and challenges for clinical translation. Nat Rev Genet 2017; 18:362–376

16. Lord C, Veenstra-VanderWeele J: Following the trail from genotype to phenotypes. JAMA Psychiatry 2016; 73:7–8

17. Bruining H, de Sonneville L, Swaab H, et al: Dissecting the clinical heterogeneity of autism spectrum disorders through defined genotypes. PLoS One 2010; 5:e10887

18. D'Angelo D, Lebon S, Chen Q, et al: Defining the effect of the 16p11.2 duplication on cognition, behavior, and medical comorbidities. JAMA Psychiatry 2016; 73:20–30

19. Hanson E, Bernier R, Porche K, et al: The cognitive and behavioral phenotype of the 16p11.2 deletion in a clinically ascertained population. Biol Psychiatry 2015; 77:785–793

20. Hippolyte L, Maillard AM, Rodriguez-Herreros B, et al: The number of genomic copies at the 16p11.2 locus modulates language, verbal memory, and inhibition. Biol Psychiatry 2016; 80:129–139

21. Richards C, Jones C, Groves L, et al: Prevalence of autism spectrum disorder phenomenology in genetic disorders: a systematic review and meta-analysis. Lancet Psychiatry 2015; 2:909–916

22. Vorstman JAS, Morcus MEJ, Duijff SN, et al: The 22q11.2 deletion in children: high rate of autistic disorders and early onset of psychotic symptoms. J Am Acad Child Adolesc Psychiatry 2006; 45:1104–1113

23. Antshel KM, Aneja A, Strunge L, et al: Autistic spectrum disorders in velo-cardio facial syndrome (22q11.2 deletion). J Autism Dev Disord 2007; 37:1776–1786

24. Schneider M, Debbané M, Bassett AS, et al: Psychiatric disorders from childhood to adulthood in 22q11.2 deletion syndrome: results from the International Consortium on Brain and Behavior in 22q11.2 Deletion Syndrome. Am J Psychiatry 2014; 171:627–639

25. Wenger TL, et al: 22q11.2 duplication syndrome: elevated rate of autism spectrum disorder and need for medical screening. Mol Autism 2016; 7:1

26. Baird G, Simonoff E, Pickles A, et al: Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). Lancet 2006; 368:210–215

27. Kogan MD, Blumberg SJ, Schieve LA, et al: Prevalence of parent-reported diagnosis of autism spectrum disorder among children in the US, 2007. Pediatrics 2009; 124:1395–1403

28. Ronald A, Happé F, Bolton P, et al: Genetic heterogeneity between the three components of the autism spectrum: a twin study. J Am Acad Child Adolesc Psychiatry 2006; 45:691–699

29. Bruining H, Eijkemans MJ, Kas MJ, et al: Behavioral signatures related to genetic disorders in autism. Mol Autism 2014; 5:11

30. Geschwind DH, Levitt P: Autism spectrum disorders: developmental disconnection syndromes. Curr Opin Neurobiol 2007; 17:103–111

31. Kates WR, Antshel KM, Fremont WP, et al: Comparing phenotypes in patients with idiopathic autism to patients with velocardiofacial syndrome (22q11 DS) with and without autism. Am J Med Genet A 2007; 143A:2642–2650

32. Rutter M, Le Couteur A, Lord C: Autism Diagnostic Interview, Revised (ADI-R). Torrance, Calif, Western Psychological Services, 2003

33. Niarchou M, Zammit S, van Goozen SH, et al: Psychopathology and cognition in children with 22q11.2 deletion syndrome. Br J Psychiatry 2014; 204:46–54

34. Chawner S, Owen MJ, Holmans P, et al: Genotype-phenotype associations in children with copy number variants associated with high neuropsychiatric risk in the UK (IMAGINE-ID): a case-control cohort study. Lancet Psychiatry 2019; 6:493–505

35. Gur RE, Bassett AS, McDonald-McGinn DM, et al: A neurogenetic model for the study of schizophrenia spectrum disorders: the International 22q11.2 Deletion Syndrome Brain Behavior Consortium. Mol Psychiatry 2017; 22:1664–1672

36. Niarchou M, Chawner SJRA, Doherty JL, et al: Psychiatric disorders in children with 16p11.2 deletion and duplication. Transl Psychiatry 2019; 9:8

37. Fiksinski AM, Breetvelt EJ, Duijff SN, et al: Autism spectrum and psychosis risk in the 22q11.2 deletion syndrome: findings from a prospective longitudinal study. Schizophr Res 2017; 188:59–62

38. Szatmari P, Paterson AD, Zwaigenbaum L, et al: Mapping autism risk loci using genetic linkage and chromosomal rearrangements. Nat Genet 2007; 39:319–328

39. Merikangas AK, Segurado R, Heron EA, et al: The phenotypic manifestations of rare genic CNVs in autism spectrum disorder. Mol Psychiatry 2015; 20:1366–1372

40. Lord C, Rutter M, Le Couteur A, et al: Autism Diagnostic Interview–Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord 1994; 24:659–685

41. Hodapp RM: Direct and indirect behavioral effects of different genetic disorders of mental retardation. Am J Ment Retard 1997; 102:67–79

42. Moreno-De-Luca A, Evans DW, Boomer KB, et al: The role of parental cognitive, behavioral, and motor profiles in clinical variability in individuals with chromosome 16p11.2 deletions. JAMA Psychiatry 2015; 72:119–126

43. Klaassen P, Duijff SN, Swanenburg de Veye H, et al: Explaining the variable penetrance of CNVs: parental intellectual level modulates the expression of intellectual impairment caused by the 22q11.2 deletion. Am J Med Genet Part B Neruopsych Genet 2016; 171:(6) 790–796

44. Olszewski AK, Radoeva PD, Fremont W, et al: Is child intelligence associated with parent and sibling intelligence in individuals with developmental disorders? an investigation in youth with 22q11.2 deletion (velo-cardio-facial) syndrome. Res Dev Disabil 2014; 35:3582–3590

45. Olsen L, Sparsø T, Weinsheimer SM, et al: Prevalence of rearrangements in the 22q11.2 region and population-based risk of neuropsychiatric and developmental disorders in a Danish population: a case-cohort study. Lancet Psychiatry 2018; 5:573–580