Identification of Developmental and Behavioral Markers Associated With Genetic Abnormalities in Autism Spectrum Disorder

Somer L. Bishop, Ph.D., Cristan Farmer, Ph.D., Vanessa Bal, Ph.D., Elise B. Robinson, Ph.D., A. Jeremy Willsey, Ph.D., Donna M. Werling, Ph.D., Karoline Alexandra Havdahl, Ph.D., Stephan J. Sanders, M.D., Ph.D., Audrey Thurm, Ph.D.

Objective: Aside from features associated with risk of neurogenetic syndromes in general (e.g., cognitive impairment), limited progress has been made in identifying phenotype-genotype relationships in autism spectrum disorder (ASD). The objective of this study was to extend work in the Simons Simplex Collection by comparing the phenotypic profiles of ASD probands with or without identified de novo loss of function mutations or copy number variants in high-confidence ASD-associated genes or loci.

Method: Analyses preemptively accounted for documented differences in sex and IQ in affected individuals with de novo mutations by matching probands with and without these genetic events on sex, IQ, and age before comparing them on multiple behavioral domains.

Results: Children with de novo mutations (N=112) had a greater likelihood of motor delay during early development (later age at walking), but they were less impaired on certain measures of ASD core symptoms (parent-rated social communication abnormalities and clinician-rated diagnostic certainty about ASD) in later childhood. These children also showed relative strengths in verbal and language abilities, including a smaller discrepancy between nonverbal and verbal IQ and a greater likelihood of having achieved fluent language (i.e., regular use of complex sentences).

Conclusions: Children with ASD with de novo mutations may exhibit a "muted" symptom profile with respect to social communication and language deficits relative to those with ASD with no identified genetic abnormalities. Such findings suggest that examining early milestone differences and standardized testing results may be helpful in etiologic efforts, and potentially in clinical differentiation of various subtypes of ASD, but only if developmental and demographic variables are properly accounted for first.

Am J Psychiatry 2017; 174:576–585; doi: 10.1176/appi.ajp.2017.16101115

Although the majority of children with autism spectrum disorder (ASD) do not have genetic abnormalities that are identifiable with currently available technology, a variety of single-gene disorders and chromosomal abnormalities have been associated with ASD and/or intellectual disability (1). In addition, among children with clinical diagnoses of ASD, those with dysmorphic features or complex medical problems (2, 3) are more likely to be identified as having strongly predisposing genetic risk factors (4). Together, these observations have led to a distinction between "syndromic" ASD, in which ASD is one of many diagnoses recognized as part of a neurogenetic syndrome, and the more common "idiopathic" ASD, in which ASD is presumed to occur as a result of unknown causes (3).

Recent advances in genomics technology, together with analyses of large-scale collections of ASD probands, have challenged the syndromic-idiopathic distinction. Microarray analysis and whole exome sequencing in large data sets like the Simons Simplex Collection (SSC) have identified numerous ASD-associated genetic loci in probands (5–9) and have clearly demonstrated an important role for highly penetrant de novo genetic mutations in individuals previously assumed to have idiopathic ASD and specifically selected for minimal syndromic features. These findings highlight the importance of changing methodological standards to require genetic testing prior to idiopathic classification, but they also leave open the question of whether individuals with identifiable genetic abnormalities are phenotypically distinguishable.

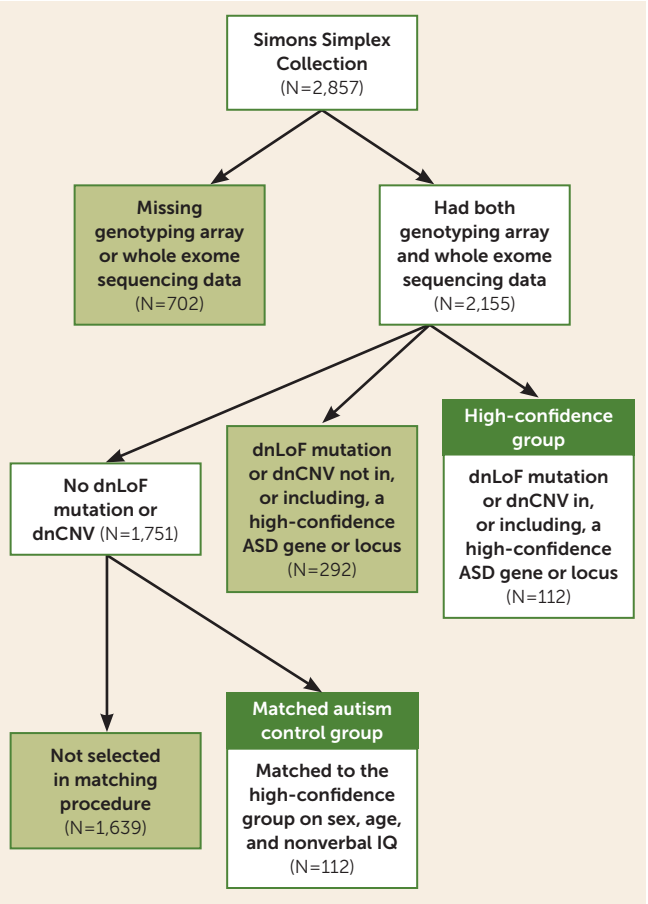
Multiple investigations have compared individuals with ASD-associated syndromes to those with presumed idiopathic ASD to understand how various neurobiological mechanisms might contribute to ASD behavioral phenotypes (10–12). Previous comparisons of individuals with ASD with or without

See related features: **CME course** *(p. 605) and* **AJP Audio** *(online)*

an associated syndrome (or a de novo mutation of potential pathogenic significance) have been limited by the difficulty of identifying appropriate controls with idiopathic ASD (10). Individuals with neurogenetic syndromes with ASD often have significantly lower cognitive abilities than those with either ASD or the neurogenetic syndrome, making it difficult to interpret direct comparisons on behavioral measures (13). Because ASD symptom measures are strongly influenced by IQ, comparing ASD severity across cognitive ability is particularly problematic (14). Thus, while associations have emerged between individual phenotypic variables (female sex, lower IQ, seizures, deviation in head circumference and body mass index) and the presence of de novo mutations in ASD loci (5, 8, 15), efforts to link genetic findings to ASD-related behavioral profiles (e.g., strengths, weaknesses, developmental features) have had limited success (16).

The present study extends work in the SSC by simultaneously considering both genetic and phenotypic data in comparing matched groups of probands with ASD with or without identified de novo loss of function (dnLoF) mutations or de novo copy number variants (dnCNVs) in ASD-associated genes or loci. Evaluation of these abnormalities was based on findings from relatively new statistical methods for defining the likelihood that a particular genetic locus is associated with ASD (8). In contrast to previous phenotype-genotype explorations of the SSC, our analytic strategy preemptively accounts for the documented IQ difference in affected individuals with de novo mutations (8), comparing them to age-, sex-, and nonverbal IQ-matched probands ("controls") from the SSC who did not have any of the genetic events described above. Although this is not the first exploration of the SSC phenotypic data, we believe it is the first to use appropriately matched ASD controls to gain insight into the phenotypic profiles of individuals with ASD with certain types of genetic abnormalities. In addition to group profiles, we provide individual-level phenotypic data in relation to each genetic abnormality specified, to facilitate ongoing efforts to explore genotype-phenotype relationships (17, 18).

FIGURE 1. Process for Including Participants From the Simons Simplex Collection in the High-Confidence and Matched Autism Control Groups[a]



[a] ASD=autism spectrum disorder; dnLoF=de novo loss of function; dnCNV=de novo copy number variant.

METHOD

Sample Collection

Phenotypic assessments and biological samples were collected from 12 university-based centers as part of the SSC. Probands with ASD were included if they were between 4 years and 17 years, 11 months of age, did not have any first-, second-, or third-degree relatives with ASD, and met criteria for autism, ASD, or Asperger syndrome based on the standard SSC assessment (see reference 19). Participants provided written informed consent (and assent, as appropriate) after receiving a complete description of the study.

Genetic Data and Participants

A recent comprehensive, integrated analysis of transmitted and de novo variation in ASD identified 65 ASD-associated genes and an additional six ASD-associated loci with high confidence (false discovery rate ≤0.1) (8). Most evidence for ASD association came from dnLoF mutations or dnCNVs. Based on the results of Illumina genotyping array and whole exome sequencing data to identify dnLoF mutations and dnCNVs, we divided the SSC probands into three groups: 1) 112 probands with at least one dnLoF mutation or dnCNV in, or including, a high-confidence ASD gene or locus (the high-confidence group); 2) 292 probands with a dnLoF mutation or dnCNV, but not in, or including, a high-confidence ASD gene or locus (the low-confidence group); and 3) 1,751 probands with no dnLoF mutations or dnCNVs in any gene or locus (the none group). An additional 702 probands were excluded from these groups because they did not have both genotyping array and whole exome sequencing data available, and therefore we could not be sure of their mutation status.

The main analyses were conducted between the high-confidence group and a subset of 112 cases from the none group, matched on nonverbal IQ, age, and sex. We refer to these cases as "matched autism controls." Figure 1 depicts the process by which participants were included in the

TABLE 1. Demographic Characteristics of Participants in the Simons Simplex Collection for Whom Genotyping Array and Whole Exome Sequencing Were Available[a]

Characteristic	None (N=1,751)		Low-Confidence Group (N=292)		High-Confidence Group (N=112)		Matched Autism Control Group (N=112)	
	N	%	N	%	N	%	N	%
Male	1,546	88	245	84	86	77	86	77
White	1,375	79	224	77	95	85	86	77
Hispanic	211	12	36	9	9	8	14	13
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age (months)	107.56	42.56	114.93	44.85	113.10	39.75	112.83	39.40
Nonverbal IQ	86.28	26.06	80.34	27.31	74.88	23.96	75.46	24.40

[a] The matched autism control group consists of participants from the "none" group (no de novo loss of function mutations or de novo copy number variants in any gene or locus), matched on sex, age (within 8 months), and nonverbal IQ (within 10 points) to the high-confidence group. As a result, the high-confidence–matched autism control pairs did not differ significantly in sex, age, or nonverbal IQ.

high-confidence group or the matched autism control group. The participants' demographic characteristics are summarized in Table 1. A list of the specific genetic abnormalities represented in the high-confidence group is available in Table S1 in the data supplement that accompanies the online edition of this article. We examined the high-confidence group as a whole, and we also identified seven dnLoF mutations or dnCNVs found in at least four participants; these have previously been reported separately, and they include 16q11.2 deletions and duplications, 15q11.2–13 duplications, 1q21.1 duplication, and 7q11.23 duplications (7), as well as DYRK1A LoF mutations (20) and CHD8 LoF mutations (16). Supplementary analyses compared individuals from the low-confidence group, who may later be identified as high confidence as further studies are completed, to a separate group of matched controls from the none group (see Table S2 and Figure S1 in the online data supplement).

Measures

Matched groups were compared on a number of phenotypic domains. Cognitive ability was indexed using nonverbal IQ and verbal IQ, which were derived from standardized tests administered according to the child's ability level. Standard scores from the daily living skills domain of the Vineland Adaptive Behavior Scales, 2nd Edition (Vineland II) (21) provide a measure of independent functioning that can be used alongside cognitive ability to index the presence and severity of intellectual disability. Motor skills were measured using item 5 from the Autism Diagnostic Interview–Revised (ADI-R) (22), which inquired about age at onset of independent walking, and the raw scores from the Purdue pegboard task. Language was measured using age at first words (item 9) and age at first phrases (item 10) from the ADI-R; the module of the Autism Diagnostic Observation Schedule (ADOS) (23) that provides a gross estimate of expressive language level (module 1, nonverbal/single words; module 2, flexible phrase speech; and modules 3 and 4, fluent speech/regular use of complex sentences); the standard score on the Peabody Picture Vocabulary Test, 4th Edition (24); and the standard score on the Vineland-II communication domain. We also report a language deficit variable, coded as present when the child's ADOS module was lower than what would be expected based on his or her nonverbal mental age. Social communication deficits and restricted and repetitive behaviors associated with ASD were measured using total scores from the social, communication, and repetitive behavior domains of the ADI-R and the domain-calibrated scores from the ADOS (25). The ADI-R domain scores are based on behaviors retrospectively reported by the parent to have occurred when the child was between ages 4 and 5 or ever in the past, whereas the ADOS is based on currently observed behaviors. Current level of overall ASD symptoms was assessed using total scores from the Social Responsiveness Scale (SRS) (26), the ADOS overall calibrated severity scores (27), and a clinician-rated measure of ASD diagnostic certainty (the minimum score was 6 in the presence of an ASD diagnosis, so SSC scores ranged from 6 to 15). Behavior problems not specific to ASD were measured using T scores for externalizing and internalizing problems from the Child Behavior Checklist (28), a parent-report questionnaire. Presence of seizures was assessed using combined information from the SSC medical history form and ADI-R item 85. Family history of major psychiatric problems was determined from the SSC medical history form, based on presence or absence of schizophrenia, bipolar disorder, or depressive disorder in a family member with a level of genetic relatedness at least on the order of first cousin (see reference 29).

Statistical Analysis

A randomized "nearest neighbor" approach was used to match probands with dnLoF mutations or dnCNVs in genes or loci with previously established ASD significance (high-confidence group) to probands with no such genetic events (none group) at a 1:1 ratio. Matching procedures were performed separately for male and female participants, using ranges of 10 nonverbal IQ points and 8 months of age. These ranges were selected as the narrowest ranges within which probands from the matched autism control group could be found for all probands from the high-confidence group. Matching procedures were performed using a macro in SAS (30). Case-control differences were evaluated using

a mixed model with a random effect of the case-control pair (to reflect the correlated nature of the matched data) and a fixed effect of group for continuous variables, or a conditional logistic regression for categorical variables. In both types of models, an interaction with nonverbal IQ was included to determine whether group differences were moderated by cognitive level. We present both uncorrected and false-discovery-rate-corrected (31) p values. False discovery rate was calculated separately for the case-control differences and the moderator analyses, both using the total number of comparisons. Analyses were conducted in SAS, version 9.3 (32).

RESULTS

As has been reported in previous phenotype-genotype explorations within the SSC (5, 8), probands with dnLoF mutations or dnCNVs had lower nonverbal IQ and were more likely to be female than those without (in the high-confidence group and the entire none group, nonverbal IQ was 75 and 86, respectively [$p<0.0001$], and the proportion female was 23% and 12%, respectively [$p=0.0002$]). The results of the matching procedures are shown in Table 1. Results of the paired comparisons are also shown in Table 2 and are illustrated relative to the full SSC sample in Figure 2. After correction for multiple comparisons, several differences between the matched groups were observed.

Children from the high-confidence group scored significantly lower (indicating fewer ASD symptoms) on the ADI-R social domain total than the matched autism control group ($p_{corrected}=0.01$), but the difference in ADI-R nonverbal communication total scores did not survive correction ($p_{corrected}=0.07$). Current ASD symptoms (ADOS social affect calibrated severity score) did not differ significantly between the high-confidence and matched autism control groups after correction ($p_{corrected}=0.07$), although the trend was for less severe symptoms in the high-confidence group. Clinicians were significantly less confident in the ASD diagnosis for probands in the high-confidence group ($p_{corrected}=0.001$).

Generally, the verbal cognitive and language abilities of the high-confidence group exceeded those of the matched autism control group (Table 2). Verbal IQ was higher ($p_{corrected}=0.02$) and more consistent with nonverbal IQ (nonverbal IQ-verbal IQ difference between groups, $p_{corrected}=0.01$) in the high-confidence group than in the matched autism control group, which had larger splits between nonverbal and verbal IQ. The mean split in the high-confidence group was nearly zero (mean=0.61, SD=16.46), compared with a mean of 7.40 (SD=16.10) in the matched autism control group (Cohen's d=0.41, 95% CI=0.14, 0.68). Probands in the high-confidence group also had significantly higher scores on the Peabody Picture Vocabulary Test than the matched autism control group ($p_{corrected}=0.01$) (a difference that was more pronounced at lower levels of nonverbal IQ; interaction $p_{corrected}=0.02$) and were more likely to receive modules 3 or 4 of the ADOS ($p_{corrected}=0.01$).

Probands in the high-confidence group reportedly started walking at a significantly later age than the matched autism control group ($p_{corrected}=0.001$). This difference depended on the nonverbal IQ level of the case-control pair, such that the magnitude of the difference in age at first walking was larger at lower IQ (interaction $p_{corrected}=0.02$). When nonverbal IQ was held constant at 30, the least squares mean estimate for age at first walking in the high-confidence group was 19.0 months, compared with 13.6 months in the matched autism control group; at a nonverbal IQ of 50, the mean estimates were 17.6 and 13.6 months, respectively; at a nonverbal IQ of 70, the mean estimates were 16.1 and 13.5 months, respectively; and at a nonverbal IQ of 90, the mean estimates were 14.7 and 13.5 months, respectively. No differences between groups were observed on the Purdue pegboard task, a measure of current fine motor skills.

Phenotypic profiles for subgroups of high-confidence probands with identified de novo mutations in the same locus (observed in at least four individuals in this sample) are presented in Figure 3. Although a few discernable profiles are apparent, readers are cautioned that within-group variability was high and sample sizes were small.

DISCUSSION

Findings from previous phenotype-genotype explorations within the SSC, and from other comparisons of syndromic and idiopathic ASD, indicate that children with ASD who have identifiable genetic abnormalities have lower IQ and higher rates of medical problems and dysmorphology than those without genetic abnormalities (4, 5, 8). Differences in behaviors that are related to ASD more specifically (rather than to neurodevelopmental disruption or intellectual disability more generally) have not typically emerged from large genotyped data sets, although this may be attributable to the fact that ASD symptom measures are strongly influenced by IQ (14). In order to further our understanding of whether and how children with ASD with either dnLoF mutations or dnCNVs in the SSC differ from comparable children with ASD without these abnormalities, we identified a group of sex-, age-, and nonverbal IQ-matched individuals to serve as controls. These matched groups were then compared across several phenotypic domains relevant to the characterization of individuals with ASD. Although the smaller male-to-female ratio in the high-confidence group compared with the none group was interesting and consistent with the literature on female sex conferring specific risk for de novo genetic abnormalities (5), the small number of female participants in the sample prohibited sex-based comparisons.

Results of the matched comparisons indicated that children with dnLoF mutations or dnCNVs in high-confidence ASD-associated genes or loci were *less* impaired on certain measures of ASD core symptoms (primarily social communication and diagnostic certainty) than their matched counterparts.

TABLE 2. Phenotypic Comparison of the High-Confidence Group and the Matched Autism Control Group[a]

						Matched Autism Control Group Versus High-Confidence Group[b]			Group-by-Nonverbal IQ Interaction (Nonverbal IQ as Moderator)[b]		
Measure	Pairs (N)	Matched Autism Control Group		High-Confidence Group		Test Statistic	p	FDR p	Test Statistic	p	FDR p
		Mean	SD	Mean	SD						
Age (months) (matching variable)	112	112.83	39.41	113.10	39.75						
Nonverbal IQ (matching variable)	112	75.46	24.00	74.88	23.96						
Verbal IQ	112	68.05	31.18	74.28	29.49	−2.81	0.01	0.02	1.16	0.28	0.75
Nonverbal IQ minus verbal IQ	112	7.40	16.10	0.61	16.46	3.12	0.002	0.01	1.16	0.28	0.75
Autism Diagnostic Interview–Revised											
Age at first words (months)	112	27.28	17.43	29.71	27.18	−0.80	0.43	0.55	3.69	0.06	0.39
Age at first phrases (months)	112	49.10	23.16	46.86	31.81	0.71	0.48	0.58	2.56	0.11	0.50
Age at onset of walking (months)	112	13.54	3.44	15.79	4.93	−4.28	<0.001	0.001	12.13	0.001	0.02
Vineland Adaptive Behavior Scales, 2nd Edition											
Communication, standard score	112	72.63	12.17	73.49	12.19	−0.86	0.39	0.55	0.46	0.50	0.92
Daily living skills, standard score	112	74.25	12.77	72.81	13.22	1.11	0.27	0.43	0.06	0.81	0.97
Socialization, standard score	112	69.83	11.51	69.30	12.20	0.44	0.66	0.71	0.04	0.84	0.97
Adaptive behavior composite, standard score	112	70.82	10.53	69.72	10.94	1.15	0.25	0.42	0.01	0.93	0.97
Peabody Picture Vocabulary Test, standard score	109	74.86	31.24	81.77	25.23	−3.32	0.001	0.01	10.93	0.001	0.02
Purdue pegboard task											
Both hands, raw score	70	6.14	3.41	6.21	3.33	−0.20	0.85	0.85	0.59	0.44	0.92
Dominant hand, raw score	70	8.81	2.78	9.04	3.10	−0.64	0.52	0.61	0.10	0.75	0.96
Nondominant, raw score	70	7.77	3.50	8.39	3.59	−1.38	0.17	0.34	0.11	0.74	0.96
Child Behavior Checklist											
Internalizing scale, total T score	111	58.88	9.15	60.46	8.62	−1.33	0.19	0.35	0.40	0.53	0.92
Externalizing scale, total T score	111	55.70	10.57	57.85	11.60	−1.44	0.15	0.32	0.15	0.70	0.96
Social Responsiveness Scale, total T score	111	81.34	9.93	80.34	11.13	0.71	0.48	0.58	0.01	0.92	0.97
Autism Diagnostic Interview–Revised											
Social domain, total score	112	22.43	5.35	20.15	5.43	3.43	0.001	0.01	0.18	0.67	0.96
Nonverbal communication, total score	112	10.01	3.47	9.09	3.34	2.25	0.03	0.07	1.11	0.29	0.75
Restricted and repetitive behaviors, total score	112	6.76	2.59	6.49	2.59	0.84	0.40	0.55	0.01	0.91	0.97
Autism Diagnostic Observation Schedule											
Total, calibrated severity score	112	7.79	1.53	7.29	1.82	1.26	0.21	0.37	0.54	0.46	0.92
Social affect, calibrated severity score	109	7.47	1.69	7.07	1.84	2.23	0.03	0.07	2.48	0.12	0.50
Restricted and repetitive behaviors, calibrated severity score	109	8.18	1.76	7.86	1.98	1.65	0.10	0.25	2.38	0.12	0.50
Overall ASD diagnostic certainty	112	13.83	1.98	12.55	2.60	4.25	<0.001	0.001	1.82	0.18	0.62
		N	%	N	%						
Autism Diagnostic Observation Schedule module[c]	112					8.63	0.003	0.01	0.20	0.65	0.96
1		31	28	19	17						
2		24	21	20	18						
3		53	47	71	63						
4		4	4	2	2						

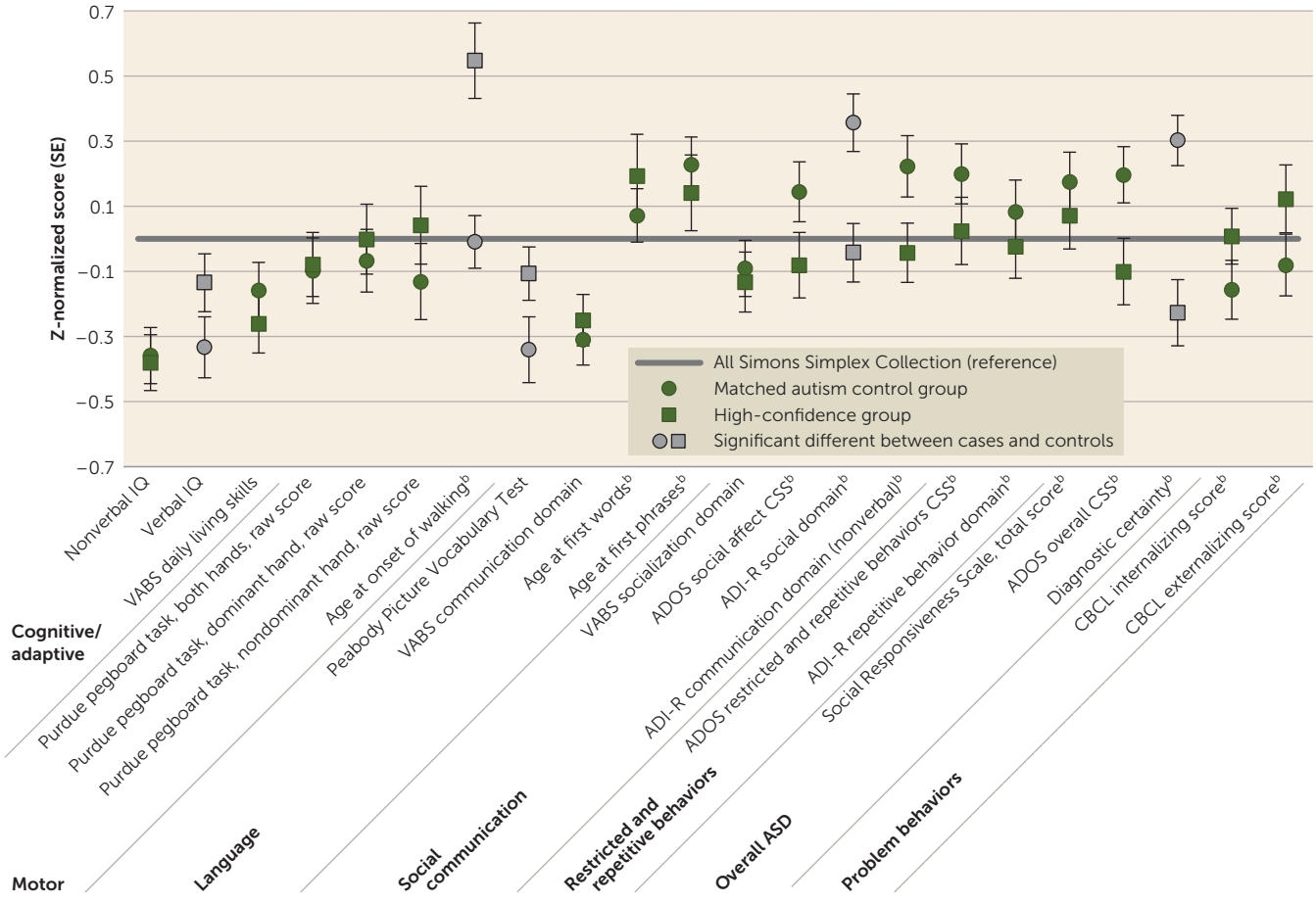
continued

TABLE 2, *continued*

Measure	Pairs (N)	Matched Autism Control Group		High-Confidence Group		Matched Autism Control Group Versus High-Confidence Group[b]			Group-by-Nonverbal IQ Interaction (Nonverbal IQ as Moderator)[b]		
						Test Statistic	p	FDR p	Test Statistic	p	FDR p
High ASD diagnostic certainty[d]	112	97	87	71	63	13.20	0.003	0.003	0.18	0.67	0.96
Family history of major psychiatric problems[e]	96	45	47	43	45	0.10	0.76	0.79	0.94	0.33	0.77
Seizures	112	6	5	13	12	2.45	0.12	0.28	4.34	0.04	0.35
Language deficit	112	40	36	24	21	7.24	0.01	0.02	0.00	0.97	0.97

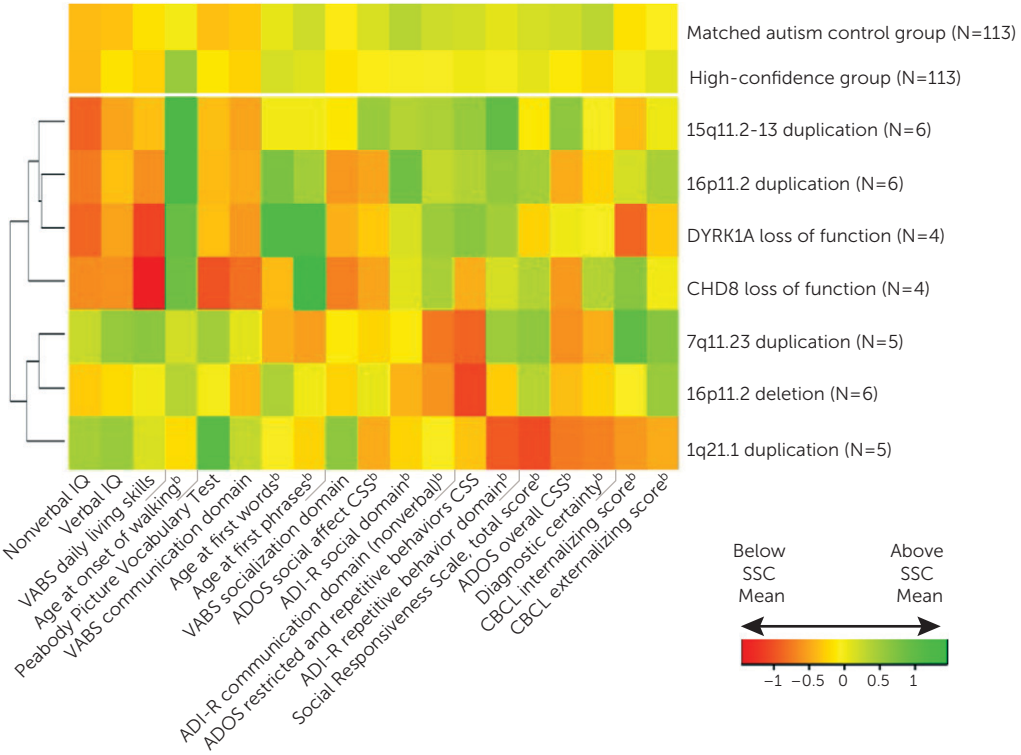
[a] ASD=autism spectrum disorder.
[b] In order to maintain the integrity of our matching procedure, if only one member of a pair was missing data on a given measure, the partner's data were also set to missing. The test statistic depends on the type of dependent variable; continuous variables (described with means) have an associated t-statistic, while categorical variables (described with proportions) have an associated χ^2 statistic. "Nonverbal IQ as moderator" refers to the interaction between group (high-confidence group versus matched autism control group) and nonverbal IQ in predicting the dependent variable and informs the question of whether group differences depend on cognitive level. FDR=false discovery rate.
[c] Module type was collapsed to modules 1 and 2 versus modules 3 and 4 for analysis.
[d] Certainty score was greater than 12, operationalized in the study as "high certainty."
[e] Controlling for ethnicity.

FIGURE 2. Phenotypic Profiles of Participants in the High-Confidence Group and the Matched Autism Control Group Relative to the Full Simons Simplex Collection Sample[a]



[a] Variables were z-normalized using the mean and standard deviation in the full Simons Simplex Collection sample as the reference. Mean z scores in each group are plotted. Gray markers indicate a significant difference between cases and controls (see Table 1). ADI-R=Autism Diagnostic Interview–Revised; ADOS=Autism Diagnostic Observation Schedule; CBCL=Child Behavior Checklist; CSS=calibrated severity score; VABS=Vineland Adaptive Behavior Scales, 2nd Edition.
[b] For these measures, a higher value is more severe or more atypical.

FIGURE 3. Profiles of Individual Conditions in the High-Confidence Group and the Matched Autism Control Group Relative to the Full Simons Simplex Collection Sample[a]



[a] De novo events found in at least four participants are shown alongside the full high-confidence sample and the matched autism control sample. Variables were z-normalized using the mean and standard deviation in the full Simons Simplex Collection (SSC) sample, and the colors in the heat map represent z scores above or below the SSC mean. Hierarchical clustering, for the purposes of presentation (indicated by the dendrogram on the left-hand side), was performed using Ward's method and Euclidian distance. ADI-R=Autism Diagnostic Interview–Revised; ADOS=Autism Diagnostic Observation Schedule; CBCL=Child Behavior Checklist; CSS=calibrated severity score; VABS=Vineland Adaptive Behavior Scales, 2nd Edition.

[b] For these measures, a higher value is more severe or more atypical.

Children from the high-confidence group also showed relative strengths in verbal and language abilities, including a smaller gap between nonverbal and verbal IQ, and were more likely to have achieved fluent expressive language abilities at the time of the SSC assessment (i.e., they were regularly using complex sentences and therefore were capable of completing module 3 or 4 of the ADOS). This suggests that once nonverbal IQ and age are taken into account, children with ASD with certain genetic abnormalities may exhibit a "muted" symptom profile with respect to language and social communication deficits relative to those with ASD with no identified genetic abnormalities. On the other hand, consistent with previous findings in individuals with intellectual disability, children from the high-confidence group were more likely to show delays in motor functioning as measured by onset of independent walking (see reference 33). In the matched ASD comparisons, for every 1-month delay in walking, there was a 17% increase in the odds of a de novo mutation being present, suggesting that age at onset of walking may be useful as a marker of potential genetic abnormality in samples with ASD (33). Furthermore, this finding of delayed gross motor milestone attainment shifts the profile of children with de novo mutations in this sample away from an exclusively ASD-specific phenotypic profile and toward a profile more similar to that of genetic syndromes associated with ASD generally.

Importantly, children with genetic abnormalities (and therefore the children selected as matched ASD controls) had lower cognitive and adaptive abilities than the rest of the SSC sample. They also tended to receive higher (worse) scores on ASD symptom measures compared with the rest of the SSC sample, mirroring decades of similar findings that children with ASD with lower IQ usually exhibit more severe impairments than those with higher IQ (27). In fact, although we sought to conduct resampling to create multiple control groups, we were only able to create one matched ASD control group because of the low number of possible matches (i.e., in some cases, it was only possible to generate one match for children with ASD-associated mutations). However, the fact that children with ASD-associated mutations were *not* more impaired on measures of social communication deficits and diagnostic certainty when compared with *relevant* controls (i.e., matched on sex, age, and nonverbal IQ) indicates that these mutations, as a group, may not actually confer specific risk for ASD-related impairment that is greater than the factors conferring risk

in the none group. This interpretation is supported by the results of the low-confidence group comparison (see the online data supplement). Alternatively, other explanatory models regarding differential thresholds for behavioral expression of ASD based on heightened risk from rare de novo mutations and/or compensatory mechanisms may be relevant to those with abnormalities in high-confidence genes and diagnosed with ASD (34, 35). Regardless, continued study of early milestone and autism symptom profiles, both in samples of heterogeneous genetic abnormalities and with specific genetic abnormalities, is required to move these findings from observational to informing risk assessment for genetic testing in clinics (36).

Limitations and Future Directions

Limitations of the SSC data set for these types of comparisons include its rigid exclusion criteria for problems that are known to be associated with pathogenic genetic abnormalities, including very low mental age and birth trauma (e.g., perinatal incidents, prematurity), exclusion of individuals who did not meet stringent ASD criteria on standardized diagnostic instruments, and the lack of contemporaneously sampled controls from different families without ASD. Thus, we note the possibility that the present findings may vary when the full range of intellectual disability and associated features within ASD is represented. That the phenotypic data collection was blinded to genetic status is a major advantage over other comparisons between "syndromic" and "idiopathic" ASD, in which clinicians' ratings on standardized instruments or measures of diagnostic certainty may be subconsciously affected by biases about, for example, whether ASD in fragile X syndrome or tuberous sclerosis is the same as "idiopathic" ASD. Therefore, our finding of lower clinician-rated diagnostic certainty for children with genetic abnormalities is robust and cannot be explained by clinician bias.

Another caveat is that although the high-confidence and matched autism control groups were matched on age, children in this study spanned a wide age range (4 to 17 years). A challenge to genetic studies requiring large samples is that it is difficult to interpret within-sample comparisons of children spanning the full range of ages and developmental stages. On the one hand, results of this study suggest that ASD symptoms are less impairing in those who are diagnosed with ASD with de novo mutations in high-confidence genes or loci than in peers with equivalent cognitive skills; on the other hand, the pattern of significant differences in early motor milestones (related to lower IQ) may suggest differences in the developmental trajectories or patterns of emergence of ASD symptoms. Indeed, the fact that the high-confidence group was characterized by later onset of independent walking than the matched autism control group indicates a very early phenotypic difference. Delayed walking is more frequently observed in individuals with intellectual disability, compared with the general population and compared with individuals with ASD, which suggests that it may serve as a marker of propensity toward later cognitive impairment. Considering that the high-confidence and matched autism control groups were matched on current nonverbal cognitive functioning, presence of this early developmental difference provides further evidence for different developmental trajectories (33). Such questions underscore the need to obtain genetic data in prospective longitudinal studies.

A third limitation of the study was the small sample sizes of participants with de novo mutations in, or including, the same ASD-associated genes and loci and our subsequent combination of all of these participants into a single group. Although a number of group-level findings still emerged as significant, Figure 3 clearly illustrates the limitations of combining individuals of such diversity. It also exemplifies the variability of phenotypic expression even within a known abnormality, already observed in many studies of these specific genetic disorders (17). While it would be interesting to make observations about the most common dnLoF mutations and dnCNVs, which included four CNV duplications, one CNV deletion, and two mutations all previously associated with ASD, there are published "genetics first" cohorts for each of these (17, 18, 20, 37–39). These studies describe wide within-cohort variability in phenotypic expression, based on type of mutation or CNV characteristics, such as deletion versus duplication, size of the error, and the specific genes involved (2). An obvious next step is to continue efforts to collect sufficient numbers of cases of specific genetic abnormalities to allow comparisons both within and across disorders, although the feasibility of this approach is limited by the relative rarity of any specific mutation. However, as our understanding of the underlying molecular neurobiology improves, grouping patients with mutations expected to affect the same pathway(s), and therefore potentially leading to a similar phenotypic outcome, may provide traction in this regard (40). Relatedly, future studies may identify common variants or familially transmitted genetic abnormalities that contribute to these biologically relevant groupings.

In conclusion, these results highlight the critical need to consider ASD-related symptoms and behaviors in the context of overall developmental level. The differences between individuals with and without de novo mutations were only revealed when sex, IQ, and age were carefully controlled in the analyses. Proper steps must be taken to account for these factors in future studies in order to advance our understanding of the range of phenotypic profiles associated with genetic findings in ASD. Studies such as these need to be replicated and extended as additional genetic abnormalities are found to be associated with ASD with high confidence. Findings from these studies will elucidate actual genotype-phenotype differences within ASD, which can be used to more carefully phenotype specific animal models for treatment targeting and to inform clinical genetic risk assessment and prognosis.

AUTHOR AND ARTICLE INFORMATION

From the Department of Psychiatry, UCSF Weill Institute for Neurosciences, University of California, San Francisco; the Pediatrics and Developmental Neuroscience Branch, NIMH, Bethesda, Md.; the Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston; the Stanley Center for Psychiatric Research and the Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Mass.; and the MRC Integrative Epidemiology Unit at the University of Bristol, School of Social and Community Medicine, Bristol, U.K.

Address correspondence to Dr. Bishop (somer.bishop@ucsf.edu).

Supported by the Intramural Program of NIMH. Dr. Robinson was funded by NIMH grant 1K01MH099286-01A1 and NARSAD Young Investigator grant 22379. Dr. Havdahl is a member of the MRC Integrative Epidemiology Unit, which is supported by the University of Bristol and the UK Medical Research Council (MC_UU_12013/1). Drs. Werling and Sanders were funded by the Simons Foundation (research award 307705 to Dr. Sanders). Dr. Werling was funded by NIMH (P50 MH106934 to Dr. Nenad Sestan) and the Autism Science Foundation (16-009 to Dr. Werling).

Dr. Bishop has received royalties from Western Psychological Services for the publication of the Autism Diagnostic Observation Schedule, 2nd Edition; all royalties received related to any research in which Dr. Bishop is involved are given to a not-for-profit agency. Dr. Willsey serves on an advisory panel for Daiichi Sankyo. The other authors report no financial relationships with commercial interests.

Received Oct. 4, 2016; revision received Dec. 14, 2016; accepted Jan. 13, 2017; published online March 3, 2017.

REFERENCES

1. de la Torre-Ubieta L, Won H, Stein JL, et al: Advancing the understanding of autism disease mechanisms through genetics. Nat Med 2016; 22:345–361
2. Watson CT, Marques-Bonet T, Sharp AJ, et al: The genetics of microdeletion and microduplication syndromes: an update. Annu Rev Genomics Hum Genet 2014; 15:215–244
3. Miles JH, Takahashi TN, Bagby S, et al: Essential versus complex autism: definition of fundamental prognostic subtypes. Am J Med Genet A 2005; 135:171–180
4. Tammimies K, Marshall CR, Walker S, et al: Molecular diagnostic yield of chromosomal microarray analysis and whole-exome sequencing in children with autism spectrum disorder. JAMA 2015; 314:895–903
5. Iossifov I, O'Roak BJ, Sanders SJ, et al: The contribution of de novo coding mutations to autism spectrum disorder. Nature 2014; 515: 216–221
6. De Rubeis S, He X, Goldberg AP, et al: Synaptic, transcriptional, and chromatin genes disrupted in autism. Nature 2014; 515: 209–215
7. 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
8. 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
9. Sanders SJ, Murtha MT, Gupta AR, et al: De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature 2012; 485:237–241
10. Thurman AJ, McDuffie A, Kover ST, et al: Autism symptomatology in boys with fragile X syndrome: a cross sectional developmental trajectories comparison with nonsyndromic autism spectrum disorder. J Autism Dev Disord 2015; 45:2816–2832
11. Frazier TW, Embacher R, Tilot AK, et al: Molecular and phenotypic abnormalities in individuals with germline heterozygous PTEN mutations and autism. Mol Psychiatry 2015; 20:1132–1138
12. Bruining H, Eijkemans MJ, Kas MJ, et al: Behavioral signatures related to genetic disorders in autism. Mol Autism 2014; 5:11
13. Moss J, Howlin P, Oliver C: The assessment and presentation of autism spectrum disorder and associated characteristics in individuals with severe intellectual disability and genetic syndromes, in The Oxford Handbook of Intellectual Disablity and Development. Edited by Burack J, Hodapp R, Iarocci G, et al. New York, Oxford University Press, 2011, pp 1–57
14. Havdahl KA, Hus Bal V, Huerta M, et al: Multidimensional influences on autism symptom measures: implications for use in etiological research. J Am Acad Child Adolesc Psychiatry 2016; 55: 1054–1063.e3
15. Chaste P, Klei L, Sanders SJ, et al: Adjusting head circumference for covariates in autism: clinical correlates of a highly heritable continuous trait. Biol Psychiatry 2013; 74:576–584
16. Bernier R, Golzio C, Xiong B, et al: Disruptive CHD8 mutations define a subtype of autism early in development. Cell 2014; 158: 263–276
17. 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
18. Bernier R, Steinman KJ, Reilly B, et al: Clinical phenotype of the recurrent 1q21.1 copy-number variant. Genet Med 2016; 18: 341–349
19. Fischbach GD, Lord C: The Simons Simplex Collection: a resource for identification of autism genetic risk factors. Neuron 2010; 68: 192–195
20. van Bon BW, Coe BP, Bernier R, et al: Disruptive de novo mutations of DYRK1A lead to a syndromic form of autism and ID. Mol Psychiatry 2016; 21:126–132
21. Sparrow S, Cicchetti DV, Balla DA: Vineland Adaptive Behavior Scales, 2nd ed. Circle Pines, Minn, AGS Publishing, 2005
22. Rutter M, LeCouteur A, Lord C: Autism Diagnostic Interview–Revised (ADI-R). Los Angeles, Western Psychological Services, 2003
23. Lord C, Rutter M, DiLavore PC, et al: Autism Diagnostic Observation Schedule (ADOS). Los Angeles, Western Psychological Services, 1999
24. Dunn DM, Dunn LM: Peabody Picture Vocabulary Test: Manual. San Antonio, Tex, Pearson, 2007
25. Hus V, Gotham K, Lord C: Standardizing ADOS domain scores: separating severity of social affect and restricted and repetitive behaviors. J Autism Dev Disord 2014; 44:2400–2412
26. Constantino JN, Gruber CP: Social Responsiveness Scale (SRS). Los Angeles, Western Psychological Services, 2005
27. Gotham K, Pickles A, Lord C: Standardizing ADOS scores for a measure of severity in autism spectrum disorders. J Autism Dev Disord 2009; 39:693–705
28. Achenbach TM, Rescorla LA: Manual for the ASEBA School-Age Forms and Profiles. Burlington, Vt, University of Vermont, Research Center for Children, Youth, and Families, 2001
29. Robinson EB, Samocha KE, Kosmicki JA, et al: Autism spectrum disorder severity reflects the average contribution of de novo and familial influences. Proc Natl Acad Sci USA 2014; 111: 15161–15165
30. Mounib E, Satchi T: Automating the selection of controls in case-control studies. Proceedings of the 25th SAS Users Group International Conference. Paper 230-25. Indianapolis, 2000
31. Benjamini Y, Hochberg Y: Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 1995; 57:289–300
32. SAS Institute: SAS, version 9.3. Cary, NC, SAS Institute, 2012
33. Bishop SL, Thurm A, Farmer C, et al: Autism spectrum disorder, intellectual disability, and delayed walking. Pediatrics 2016; 137: e20152959
34. Skuse DH: Rethinking the nature of genetic vulnerability to autistic spectrum disorders. Trends Genet 2007; 23:387–395

35. Moreno-De-Luca D, Moreno-De-Luca A, Cubells JF, et al: Cross-disorder comparison of four neuropsychiatric CNV loci. Curr Genet Med Rep 2014; 2:151–161
36. Marano RM, Mercurio L, Kanter R, et al: Risk assessment models in genetics clinic for array comparative genomic hybridization: clinical information can be used to predict the likelihood of an abnormal result in patients. J Pediatr Genet 2013; 2:25–31
37. Picinelli C, Lintas C, Piras IS, et al: Recurrent 15q11.2 BP1-BP2 microdeletions and microduplications in the etiology of neurodevelopmental disorders. Am J Med Genet B Neuropsychiatr Genet 2016; 171:1088–1098
38. Steinman KJ, Spence SJ, Ramocki MB, et al: 16p11.2 deletion and duplication: characterizing neurologic phenotypes in a large clinically ascertained cohort. Am J Med Genet A 2016; 170: 2943–2955
39. Morris CA, Mervis CB, Paciorkowski AP, et al: 7q11.23 duplication syndrome: physical characteristics and natural history. Am J Med Genet A 2015; 167A:2916–2935
40. Willsey AJ, Sanders SJ, Li M, et al: Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. Cell 2013; 155:997–1007