Title

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

Journal

American Journal of Psychiatry, 174(6)

ISSN

0002-953X

Authors

Bishop, Somer L
Farmer, Cristan
Bal, Vanessa
et al.

Publication Date

2017-06-01

DOI

10.1176/appi.ajp.2017.16101115

Peer reviewed

HHS Public Access

Author manuscript

Am J Psychiatry. Author manuscript; available in PMC 2017 August 31.

Published in final edited form as:

Am J Psychiatry. 2017 June 01; 174(6): 576–585. doi:10.1176/appi.ajp.2017.16101115.

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

Somer L. Bishop, Ph.D.[1], **Cristan Farmer, Ph.D.**[2], **Vanessa Bal, Ph.D.**[1], **Elise Robinson, Ph.D.**[3,4], **A. Jeremy Willsey, Ph.D.**[1], **Donna M. Werling, Ph.D.**[1], **Karoline Alexandra Havdahl, M.S.**[5], **Stephan J. Sanders, M.D., Ph.D.**[1], and **Audrey Thurm, Ph.D.**[2]

[1]Department of Psychiatry, UCSF Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA

[2]Pediatrics & Developmental Neuroscience Branch, National Institute of Mental Health, Bethesda, MD

[3]Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA

[4]Stanley Center for Psychiatric Research and Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA

[5]Integrative Epidemiology Unit, University of Bristol, UK

Abstract

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).

Method—This study extends work in the Simons Simplex Collection (SSC) by comparing the phenotypic profiles of ASD probands with or without identified *de novo* loss of function mutations (LoF) or Copy Number Variants (CNV) in high confidence ASD-associated genes/loci (Sanders et al., 2015). Analyses pre-emptively 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.

CORRESPONDING AUTHOR: Somer L. Bishop, Ph.D., UCSF Weill Institute for Neurosciences, 401 Parnassus Avenue, San Francisco, CA 94143, Somer.Bishop@ucsf.edu, Phone: 415-502-3349.

DISCLOSURES

Dr. Bishop has received royalties from Western Psychological Services (WPS) for the publication of the Autism Diagnostic Observation Schedule, 2nd Edition (ADOS-2). All royalties received related to any research in which Dr. Bishop are involved are given to a not for profit agency.
Dr. Farmer reports no biomedical financial interests or potential conflicts of interest.
Dr. Bal reports no biomedical financial interests or potential conflicts of interest.
Dr. Robinson reports no biomedical financial interests or potential conflicts of interest.
Dr. Willsey reports no biomedical financial interests or potential conflicts of interest.
Dr. Werling reports no biomedical financial interests or potential conflicts of interest.
Ms. Havdahl reports no biomedical financial interests or potential conflicts of interest.
Dr. Sanders reports no biomedical financial interests or potential conflicts of interest.
Dr. Thurm reports no biomedical financial interests or potential conflicts of interest.

Results—Children with *de novo* mutations (*n*=112) showed greater likelihood of motor delays during early development (i.e., later age of walking), but *less* impairment in certain measures of ASD core symptoms (parent-rated social-communication impairment and clinician-rated diagnostic certainty) 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.

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/demographic variables are properly accounted for first.

INTRODUCTION

Although the majority of children with autism spectrum disorder (ASD) do not have genetic abnormalities identifiable with currently available technology, a variety of single-gene disorders and chromosomal abnormalities have been associated with ASD and/or intellectual disability (1). Among children with ASD, those with dysmorphic features or complex medical problems (2, 3) are also 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 etiology (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 datasets 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 an associated syndrome (or a *de novo* mutation of potential pathogenic significance) are 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 only ASD or only 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 (i.e., 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 behavioral profiles (e.g., strengths, weaknesses, developmental features) have had limited success (16).

The current 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 mutations (dnLoF) or *de novo* Copy Number Variants (dnCNV) in ASD-associated genes/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 pre-emptively 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 without 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).

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 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 or dnCNV mutations. Based on the results of Illumina genotyping array and whole exome sequencing data to identify dnLoF and dnCNV, we divided the SSC probands into three groups: 1) 112 probands with at least one dnLoF or dnCNV in, or including, a high confidence ASD gene or locus (High Confidence group); 2) 292 probands with a dnLoF or dnCNV, but not in, or including, a high-confidence ASD gene or locus (Low Confidence group); and 3) 1,751 probands with no dnLoF or dnCNV in any gene or locus (None). 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 High Confidence or Matched Autism Control groups. Participant demographics are shown in Table 1. A list of the specific genetic abnormalities represented in the High Confidence group is available (Supplementary File ST1). We examined the High Confidence group as a whole, and also identified seven dnLoF or dnCNV mutations found in at least four participants; these have been previously reported separately, and include both deletions and duplications on 16q11.2, 15q11.2–13 duplications, 1q21.1 duplication, and 7q11.23 duplications (7), as well as DYRK1A LoF (20), CHD8 LoF (16). Supplementary analyses compared individuals from the Low Confidence group (group 2), 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 Supplementary Files ST2 and SF1).

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 ability level of the child. Standard scores from the Daily Living Skills domain of the Vineland Adaptive Behavior Scale, 2nd Edition (Vineland II; 21) provide a measure of independent functioning that can be used alongside cognitive ability to index 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 of independent walking, and the raw scores from the Purdue Pegboard task. **Language** was measured using age of first words (item 9) and age of first phrases (item 10) from the ADI-R, the module of the Autism Diagnostic Observation Schedule (ADOS; 23), which provides a gross estimate of expressive language level (Module 1=nonverbal/single words, Module 2=flexible phrase speech, and Modules 3 and 4=regular use of complex sentences), the Peabody Picture Vocabulary Test, 4th Edition (24) standard score, and the Vineland-II Communication Domain standard score. 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/her nonverbal mental age. **Social-communication** and **restricted and repetitive behaviors** associated with ASD were measured using total scores from the Social (A), Communication (B), and Repetitive Behavior (C) 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 the ages of 4 and 5 years 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), 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 6–15). **Behavior problems** not specific to ASD were measured using T-scores for externalizing and internalizing problems from the Child Behavior Checklist (CBCL; 28), a parent-report questionnaire. Presence of **seizures** was assessed using combined information from the SSC medical history form and the ADI-R item 85. **Family history of major psychiatric problems** was determined from the SSC Medical History form, based on presence/absence of schizophrenia, bipolar

disorder, or depressive disorder in a family member with a level of genetic relatedness at least that of first cousins (see 29).

Statistical Analysis

A randomized “nearest neighbor” approach was used to match probands with dnLoF or dnCNV mutations 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 males and females, using ranges of 10 nonverbal IQ points and 8 months of age. These ranges were selected as the narrowest range 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 SAS macro (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 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 if group differences were moderated by cognitive level. We present both uncorrected and false discovery rate (31) corrected *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 completed in SAS version 9.3 (32).

RESULTS

As has been reported in previous phenotype-genotype explorations within the SSC (5, 8), probands with dnLoF or dnCNV mutations had lower nonverbal IQ and were more likely to be female than those without (High Confidence group versus entire None group: nonverbal IQ was 75 versus 86, $p<.0001$; % female was 23% versus 12%, $p=.0002$). The results of the matching procedures are shown in Table 1. Results of the paired comparisons are also shown in Table 2, and 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 A (Social) domain Total than the Matched Autism Control group ($p_{corrected}=.01$), but the difference in ADI-R B-Nonverbal (Communication) Total scores did not survive correction ($p_{corrected}=.07$). Current ASD symptoms (ADOS Social Affect Calibrated Severity Score) did not differ significantly between the High Confidence and Matched Autism Control groups after correction ($p_{corrected}=.07$), though 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 ($p_{corrected}=.001$).

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 ($p_{corrected}=.02$) and more consistent with nonverbal IQ (nonverbal IQ-verbal IQ difference between groups, $p_{corrected}=.01$) in the High Confidence group than the Matched Autism Control group, who had larger splits between nonverbal IQ and verbal IQ. The mean split in the High Confidence group was nearly zero (0.61±16.46), compared to 7.40±16.10 in the Matched Autism Control group (Cohen’s $d=0.41$, 95% CI 0.14 to 0.68). Probands in the

High Confidence group also had significantly higher Peabody Picture Vocabulary Test scores than the Matched Autism Control group ($p_{corrected}$=.01) (a difference that was more pronounced at lower levels of nonverbal IQ, interaction $p_{corrected}$=.02), and were more likely to receive Modules 3 or 4 of the ADOS ($p_{corrected}$=.01).

Probands in the High Confidence group reportedly walked at a significantly later age than the Matched Autism Control group ($p_{corrected}$=.001). This difference depended upon 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 $p_{corrected}$=.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, versus 13.6 months in the Matched Autism Control group; at nonverbal IQ = 50, mean estimates were 17.6 and 13.6 months, respectively; at nonverbal IQ = 70, mean estimates were 16.1 and 13.5 months, respectively; and at nonverbal IQ = 90, 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 ($p_{corrected}$>.38).

Phenotypic profiles for subgroups of High Confidence probands with identified *de novo* mutations in the same locus (observed in ≥4 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 and identifiable genetic abnormalities have lower IQ and higher rates of medical problems and dysmorphology (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 datasets, though 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 or dnCNV mutations 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 to 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 females prohibited sex-based comparisons.

Results of the matched comparisons indicated that children with dnLoF or dnCNV mutations 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. 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., capable of completing Modules 3 or 4 of the ADOS). This suggests that once 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 33). In the matched ASD comparisons, for every one month delay in walking, there was a 17% increase in the odds of a *de novo* mutation being present, suggesting that age 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, 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 to 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 with comparable scores due to 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 to *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 (e.g., common variants and environmental exposures). This interpretation is supported by the results of the Low Confidence comparison (see Supplemental Information). 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 High Confidence genes diagnosed with ASD (34, 35). Regardless, continued study of these early milestone and autism symptom profiles, both in samples of heterogeneous genetic abnormalities and with specific genetic abnormalities (e.g., Fragile X), 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 dataset 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 current 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 whether, for example, ASD in Fragile X 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 one hand, results of this study suggest that ASD symptoms in those who are diagnosed with ASD with *de novo* mutations in high confidence genes or loci are less impairing compared with peers with equivalent cognitive skills; on the other hand, the pattern of significant differences in early motor milestones (related to lower IQ) might 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 to the general population and compared to individuals with ASD, suggesting it may serve as a marker of propensity toward later cognitive impairment. Considering that 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 current study was the small sample sizes of participants with *de novo* mutations in or including the same ASD associated genes/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 dnLoFs 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, though 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 impact 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 *de novo* mutations and those without 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by the Intramural Program of the NIMH, NIH (NCT00271622 and 06-M-0065). E.B.R. was funded by National Institutes of Mental Health Grant 1K01MH099286-01A1 and NARSAD Young Investigator grant 22379. K.A.H was funded by the South-Eastern Norway Regional Health Authority (2012101). S.J.S. and D.M.W. were funded by the Simons Foundation (Research Award 307705 to S.J.S.). D.M.W. was funded by the US NIMH (P50 MH106934 to Dr. Nenad Sestan) and the Autism Science Foundation (16-009 to D.M.W.)

References

1. de la Torre-Ubieta L, Won H, Stein JL, Geschwind DH. Advancing the understanding of autism disease mechanisms through genetics. Nat Med. 2016; 22:345–361. [PubMed: 27050589]
2. Watson CT, Marques-Bonet T, Sharp AJ, Mefford HC. The genetics of microdeletion and microduplication syndromes: an update. Annual review of genomics and human genetics. 2014; 15:215–244.
3. Miles JH, Takahashi TN, Bagby S, Sahota PK, Vaslow DF, Wang CH, Hillman RE, Farmer JE. Essential versus complex autism: definition of fundamental prognostic subtypes. American journal of medical genetics Part A. 2005; 135:171–180. [PubMed: 15887228]
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. [PubMed: 26325558]
5. Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, Stessman HA, Witherspoon KT, Vives L, Patterson KE, Smith JD, Paeper B, Nickerson DA, Dea J, Dong S, Gonzalez LE, Mandell JD, Mane SM, Murtha MT, Sullivan CA, Walker MF, Waqar Z, Wei L, Willsey AJ, Yamrom B, Lee YH, Grabowska E, Dalkic E, Wang Z, Marks S, Andrews P, Leotta A, Kendall J, Hakker I, Rosenbaum J, Ma B, Rodgers L, Troge J, Narzisi G, Yoon S, Schatz MC, Ye K, McCombie WR, Shendure J, Eichler EE, State MW, Wigler M. The contribution of de novo coding mutations to autism spectrum disorder. Nature. 2014; 515:216–221. [PubMed: 25363768]
6. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE, Kou Y, Liu L, Fromer M, Walker S, Singh T, Klei L, Kosmicki J, Shih-Chen F, Aleksic B, Biscaldi M, Bolton PF, Brownfeld JM, Cai J, Campbell NG, Carracedo A, Chahrour MH, Chiocchetti AG, Coon H, Crawford EL, Curran SR, Dawson G, Duketis E, Fernandez BA, Gallagher L, Geller E, Guter SJ, Hill RS, Ionita-Laza J, Jimenz Gonzalez P, Kilpinen H, Klauck SM, Kolevzon A, Lee I, Lei I, Lei J, Lehtimaki T, Lin CF, Ma'ayan A, Marshall CR, McInnes AL, Neale B, Owen MJ, Ozaki N, Parellada M, Parr JR, Purcell S, Puura K, Rajagopalan D, Rehnstrom K, Reichenberg A, Sabo A, Sachse M, Sanders SJ,

Schafer C, Schulte-Ruther M, Skuse D, Stevens C, Szatmari P, Tammimies K, Valladares O, Voran A, Li-San W, Weiss LA, Willsey AJ, Yu TW, Yuen RK, Cook EH, Freitag CM, Gill M, Hultman CM, Lehner T, Palotie A, Schellenberg GD, Sklar P, State MW, Sutcliffe JS, Walsh CA, Scherer SW, Zwick ME, Barett JC, Cutler DJ, Roeder K, Devlin B, Daly MJ, Buxbaum JD. Synaptic, transcriptional and chromatin genes disrupted in autism. Nature. 2014; 515:209–215. [PubMed: 25363760]

7. Sanders S, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D, Chu SH, Moreau MP, Gupta AR, Thomson SA, Mason CE, Bilguvar K, Celestino-Soper PBS, Choi M, Crawford EL, Davis L, Davis Wright NR, Dhodapkar RM, DiCola M, DiLullo NM, Fernandez TV, Fielding-Singh V, Fishman DO, Frahm S, Garagaloyan R, Goh GS, Kammela S, Klei L, Lowe JK, Lund SC, McGrew AD, Meyer KA, Moffat WJ, Murdoch JD, O'Roak BJ, Ober GT, Pottenger RS, Raubeson MJ, Song Y, Wang Q, Yaspan BL, Yu Timothy W, Yurkiewicz IR, Beaudet AL, Cantor RM, Curland M, Grice DE, Günel M, Lifton RP, Mane SM, Martin DM, Shaw CA, Sheldon M, Tischfield JA, Walsh CA, Morrow EM, Ledbetter DH, Fombonne E, Lord C, Martin CL, Brooks AI, Sutcliffe JS, Cook EH Jr, Geschwind D, Roeder K, Devlin B, State MW. Multiple Recurrent De Novo CNVs, Including Duplications of the 7q11.23 Williams Syndrome Region, Are Strongly Associated with Autism. Neuron. 2011; 70:863–885. [PubMed: 21658581]
8. Sanders SJ, He X, Willsey AJ, Ercan-Sencicek AG, Samocha KE, Cicek AE, Murtha MT, Bal VH, Bishop SL, Dong S, Goldberg AP, Jinlu C, Keaney JF 3rd, Klei L, Mandell JD, Moreno-De-Luca D, Poultney CS, Robinson EB, Smith L, Solli-Nowlan T, Su MY, Teran NA, Walker MF, Werling DM, Beaudet AL, Cantor RM, Fombonne E, Geschwind DH, Grice DE, Lord C, Lowe JK, Mane SM, Martin DM, Morrow EM, Talkowski ME, Sutcliffe JS, Walsh CA, Yu TW, Autism Sequencing C, Ledbetter DH, Martin CL, Cook EH, Buxbaum JD, Daly MJ, Devlin B, Roeder K, State MW. Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. Neuron. 2015; 87:1215–1233. [PubMed: 26402605]
9. Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ, Ercan-Sencicek AG, DiLullo NM, Parikshak NN, Stein JL, Walker MF, Ober GT, Teran NA, Song Y, El-Fishawy P, Murtha RC, Choi M, Overton JD, Bjornson RD, Carriero NJ, Meyer KA, Bilguvar K, Mane SM, Sestan N, Lifton RP, Gunel M, Roeder K, Geschwind DH, Devlin B, State MW. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature. 2012; 485:237–241. [PubMed: 22495306]
10. Thurman AJ, McDuffie A, Kover ST, Hagerman RJ, Abbeduto L. 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. [PubMed: 25904201]
11. Frazier TW, Embacher R, Tilot AK, Koenig K, Mester J, Eng C. Molecular and phenotypic abnormalities in individuals with germline heterozygous PTEN mutations and autism. Mol Psychiatry. 2015; 20:1132–1138. [PubMed: 25288137]
12. Bruining H, Eijkemans MJ, Kas MJ, Curran SR, Vorstman JA, Bolton PF. Behavioral signatures related to genetic disorders in autism. Molecular Autism. 2014; 5:1–12. [PubMed: 24410847]
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: Burack, J.Hodapp, R.Iarocci, G., Zigler, E., editors. The Oxford Handbook of Intellectual Disablity and Development. New York, NY: Oxford University Press; 2011. p. 1-57.
14. Havdahl KA, Hus Bal V, Huerta M, Pickles A, Oyen AS, Stoltenberg C, Lord C, Bishop SL. Multidimensional Influences on Autism Symptom Measures: Implications for Use in Etiological Research. J Am Acad Child Adolesc Psychiatry. 2016; 55:1054–1063. e1053. [PubMed: 27871640]
15. Chaste P, Klei L, Sanders SJ, Murtha MT, Hus V, Lowe JK, Willsey AJ, Moreno-De-Luca D, Yu TW, Fombonne E, Geschwind D, Grice DE, Ledbetter DH, Lord C, Mane SM, Lese Martin C, Martin DM, Morrow EM, Walsh CA, Sutcliffe JS, State MW, Devlin B, Cook EH Jr, Kim SJ. Adjusting head circumference for covariates in autism: clinical correlates of a highly heritable continuous trait. Biol Psychiatry. 2013; 74:576–584. [PubMed: 23746936]
16. Bernier R, Golzio C, Xiong B, Stessman HA, Coe BP, Penn O, Witherspoon K, Gerdts J, Baker C, Vulto-van Silfhout AT, Schuurs-Hoeijmakers JH, Fichera M, Bosco P, Buono S, Alberti A, Failla P, Peeters H, Steyaert J, Vissers LE, Francescatto L, Mefford HC, Rosenfeld JA, Bakken T, O'Roak

BJ, Pawlus M, Moon R, Shendure J, Amaral DG, Lein E, Rankin J, Romano C, de Vries BB, Katsanis N, Eichler EE. Disruptive CHD8 mutations define a subtype of autism early in development. Cell. 2014; 158:263–276. [PubMed: 24998929]

17. Hanson E, Bernier R, Porche K, Jackson FI, Goin-Kochel RP, Snyder LG, Snow AV, Wallace AS, Campe KL, Zhang Y, Chen Q, D'Angelo D, Moreno-De-Luca A, Orr PT, Boomer KB, Evans DW, Kanne S, Berry L, Miller FK, Olson J, Sherr E, Martin CL, Ledbetter DH, Spiro JE, Chung WK. The cognitive and behavioral phenotype of the 16p11.2 deletion in a clinically ascertained population. Biol Psychiatry. 2015; 77:785–793. [PubMed: 25064419]
18. Bernier R, Steinman KJ, Reilly B, Wallace AS, Sherr EH, Pojman N, Mefford HC, Gerdts J, Earl R, Hanson E, Goin-Kochel RP, Berry L, Kanne S, Snyder LG, Spence S, Ramocki MB, Evans DW, Spiro JE, Martin CL, Ledbetter DH, Chung WK. Clinical phenotype of the recurrent 1q21.1 copy-number variant. Genet Med. 2016; 18:341–349. [PubMed: 26066539]
19. Hus V, Bishop S, Gotham K, Huerta M, Lord C. Factors influencing scores on the social responsiveness scale. J Child Psychol Psychiatry. 2013; 54:216–224. [PubMed: 22823182]
20. van Bon BW, Coe BP, Bernier R, Green C, Gerdts J, Witherspoon K, Kleefstra T, Willemsen MH, Kumar R, Bosco P, Fichera M, Li D, Amaral D, Cristofoli F, Peeters H, Haan E, Romano C, Mefford HC, Scheffer I, Gecz J, de Vries BB, Eichler EE. Disruptive de novo mutations of DYRK1A lead to a syndromic form of autism and ID. Mol Psychiatry. 2016; 21:126–132. [PubMed: 25707398]
21. Sparrow, SS., Cicchetti, DV., Balla, DA. Vineland Adaptive Behavior Scales. 2. Circle Pines, MN: AGS Publishing; 2005.
22. Rutter, M., LeCouteur, A., Lord, C. Autism Diagnostic Interview-Revised (ADI-R). Los Angeles, CA: Western Psychological Services; 2003.
23. Lord, C., Rutter, M., DiLavore, PC., Risi, S. Autism Diagnostic Observation Schedule (ADOS). Los Angeles, California: Western Psychological Services; 1999.
24. Dunn, DM., Dunn, LM. Peabody picture vocabulary test: Manual. Pearson; 2007.
25. Hus, Gotham, Lord. Standardizing ADOS domain scores: separating severity of social affect and restricted and repetitive behaviors. J Autism Dev Disord. 2014; 44:2400–2412. [PubMed: 23143131]
26. Constantino, JN., Gruber, CP. Social Responsiveness Scale (SRS). Los Angeles, CA: Western Psychological Services; 2005.
27. Gotham K, Pickles A, Lord C. Standardizing ADOS scores for a measure of severity in autism spectrum disorders. Journal of Autism and Developmental Disorders. 2009; 39:693–705. [PubMed: 19082876]
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, McGrath L, Neale BM, Perlis RH, Daly MJ. Autism spectrum disorder severity reflects the average contribution of de novo and familial influences. Proc Natl Acad Sci U S A. 2014; 111:15161–15165. [PubMed: 25288738]
30. Mounib, E., Satchi, T. SUGI 25. Indianapolis, IN: 2000. Automating the selection of controls in case-control studies.
31. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological). 1995:289–300.
32. SAS Institute I. SAS. 9.3 ed. Cary, NC: SAS Institute, Inc; 2012.
33. Bishop SL, Thurm A, Farmer C, Lord C. Autism Spectrum Disorder, Intellectual Disability, and Delayed Walking. Pediatrics. 2016; 137:1–8.
34. Skuse DH. Rethinking the nature of genetic vulnerability to autistic spectrum disorders. Trends in genetics : TIG. 2007; 23:387–395. [PubMed: 17630015]
35. Moreno-De-Luca D, Moreno-De-Luca A, Cubells JF, Sanders SJ. Cross-disorder comparison of four neuropsychiatric CNV loci. Current Genetic Medicine Reports. 2014; 2:151–161.
36. Marano RM, Mercurio L, Kanter R, Doyle R, Abuelo D, Morrow EM, Shur N. 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. Journal of pediatric genetics. 2013; 2:25–31. [PubMed: 27625836]

37. Picinelli C, Lintas C, Piras IS, Gabriele S, Sacco R, Brogna C, Persico AM. 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. [PubMed: 27566550]
38. Steinman KJ, Spence SJ, Ramocki MB, Proud MB, Kessler SK, Marco EJ, Green Snyder L, D'Angelo D, Chen Q, Chung WK, Sherr EH. 16p11.2 deletion and duplication: Characterizing neurologic phenotypes in a large clinically ascertained cohort. Am J Med Genet A. 2016; 170:2943–2955. [PubMed: 27410714]
39. Morris CA, Mervis CB, Paciorkowski AP, Abdul-Rahman O, Dugan SL, Rope AF, Bader P, Hendon LG, Velleman SL, Klein-Tasman BP, Osborne LR. 7q11.23 Duplication syndrome: Physical characteristics and natural history. Am J Med Genet A. 2015; 167a:2916–2935. [PubMed: 26333794]
40. Willsey AJ, Sanders SJ, Li M, Dong S, Tebbenkamp AT, Muhle RA, Reilly SK, Lin L, Fertuzinhos S, Miller JA, Murtha MT, Bichsel C, Niu W, Cotney J, Ercan-Sencicek AG, Gockley J, Gupta AR, Han W, He X, Hoffman EJ, Klei L, Lei J, Liu W, Liu L, Lu C, Xu X, Zhu Y, Mane SM, Lein ES, Wei L, Noonan JP, Roeder K, Devlin B, Sestan N, State MW. Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. Cell. 2013; 155:997–1007. [PubMed: 24267886]

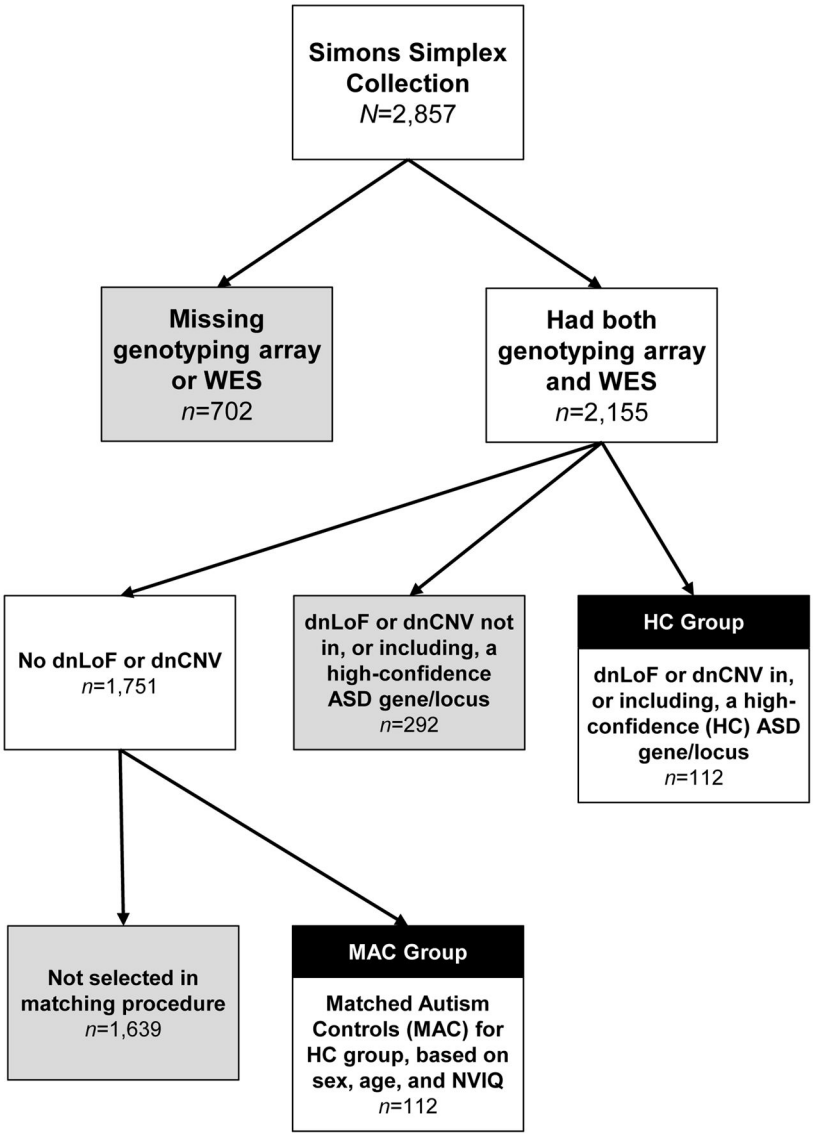

Figure 1.
Process for including participants in the HC and MAC groups

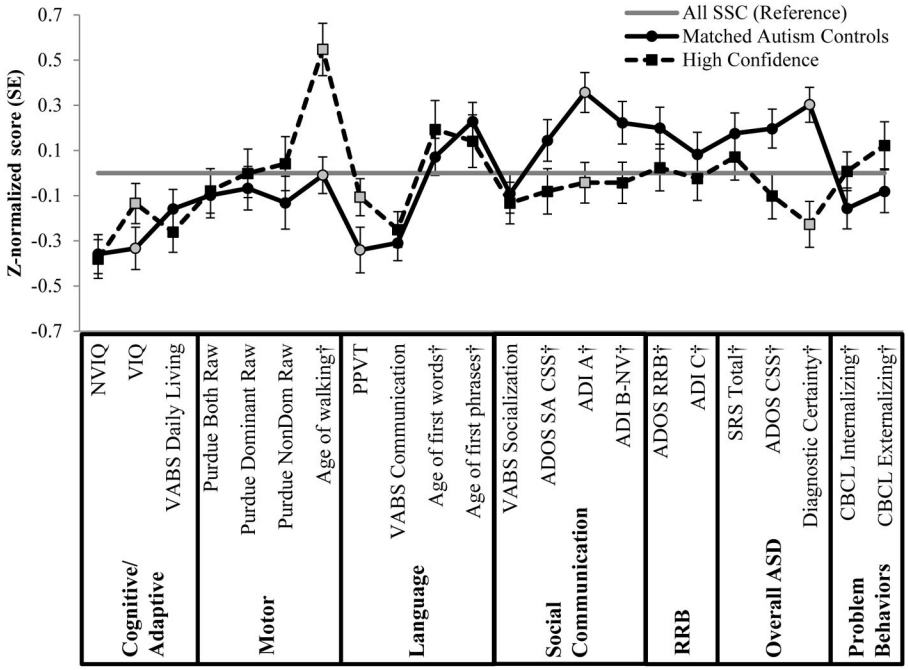

Figure 2.

Phenotypic profiles. Variables were Z-normalized using the mean and standard deviation in the full SSC sample (reference). Mean z-scores in each group are plotted. Gray markers indicate a significant difference between cases and controls (see Table 1) and a dagger (†) next to the measure name indicates that a higher value is more severe/more atypical.

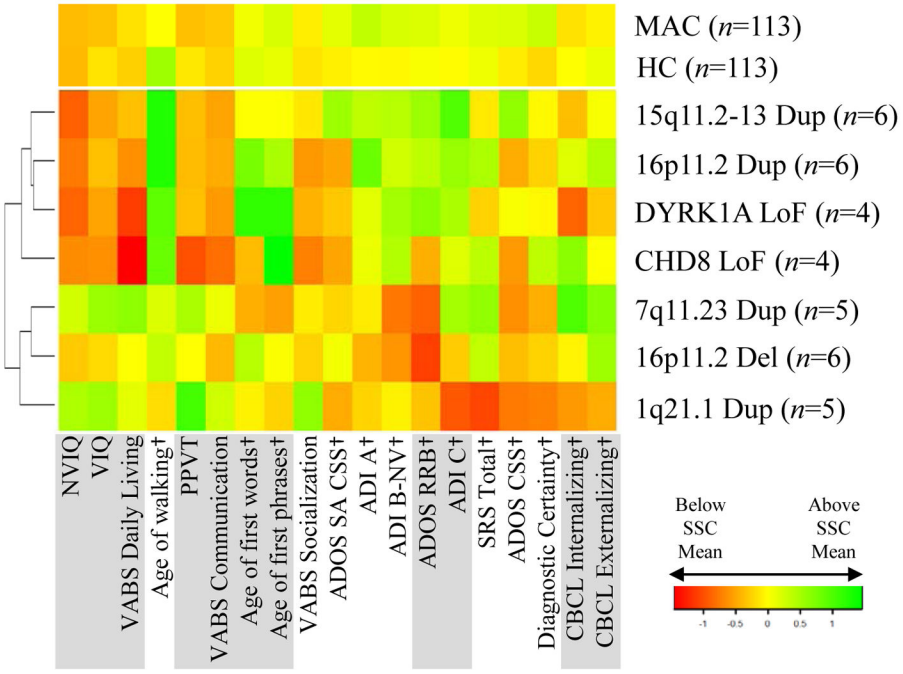

Figure 3.

Profiles of individual conditions. *De novo* events found in at least four participants are shown alongside the full High Confidence (HC) sample and the Matched Autism Controls (MAC) sample. Variables were Z-normalized using the mean and standard deviation in the full SSC sample, and the colors in the heat map represent Z scores above (green) or below (red) the SSC mean. A dagger (†) next to the measure name indicates that a higher value is more severe/more atypical. Hierarchical clustering for purpose of presentation (indicated by dendrogram on left Y axis) was performed using Ward's method and Euclidian distance.

Table 1

Participant demographics

	SSC Cohort with Genotyping Array and Whole Exome Sequencing							
	None		Low Confidence		High Confidence		High Confidence-Matched Autism Controls	
	n	**%**	***n***	**%**	***n***	**%**	***n***	**%**
N	1751		292		112		112	
Male	1546	88	245	84	86	77	86	77
White	1375	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	113.1	39.75	112.83	39.40
Nonverbal IQ	86.28	26.06	80.34	27.31	74.88	23.96	75.46	24.40

Note: The MAC group consists of participants from the None group, matched on sex, age (within 8 months), and nonverbal IQ (within 10 points) to the HC group. As a result, the MAC-HC pairs did not differ significantly on sex (p=1.0), age (p=.57), or nonverbal IQ (p=.33).

Table 2

Phenotypic comparison of High Confidence and Matched Autism Controls

	n pairs	Matched Autism Controls (MAC)		High Confidence (HC)		Matched Autism Controls versus High Confidence			Group-by-NVIQ Interaction (NVIQ as moderator)		
						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	na		na			
NVIQ (matching variable)	112	75.46	24.00	74.88	23.96	na		na			
VIQ	112	68.05	31.18	74.28	29.49	−2.81	.01	.02	1.16	.28	.75
NVIQ-VIQ Difference	112	7.40	16.10	0.61	16.46	3.12	.002	.01	1.16	.28	.75
ADI-R Age of first words (months)	112	27.28	17.43	29.71	27.18	−.80	.43	.55	3.69	.06	.39
ADI-R Age of first phrases (months)	112	49.10	23.16	46.86	31.81	.71	.48	.58	2.56	.11	.50
ADI-R Age of walking (months)	112	13.54	3.44	15.79	4.93	−4.28	<.0001	.001	12.13	.001	.02
VABS Communication Standard	112	72.63	12.17	73.49	12.19	−.86	.39	.55	0.46	.50	.92
VABS DLS Standard	112	74.25	12.77	72.81	13.22	1.11	.27	.43	0.06	.81	.97
VABS Social Standard	112	69.83	11.51	69.30	12.20	.44	.66	.71	0.04	.84	.97
VABS ABC Standard	112	70.82	10.53	69.72	10.94	1.15	.25	.42	0.01	.93	.97
PPVT Standard	109	74.86	31.24	81.77	25.23	−3.32	.001	.01	10.93	.001	.02
Purdue Pegboard Both Hands Raw	70	6.14	3.41	6.21	3.33	−.20	.85	.85	0.59	.44	.92
Purdue Pegboard Dominant Hand Raw	70	8.81	2.78	9.04	3.10	−.64	.52	.61	0.10	.75	.96
Purdue Pegboard Non-Dominant Raw	70	7.77	3.50	8.39	3.59	−1.38	.17	.34	0.11	.74	.96
CBCL Internalizing Total T	111	58.88	9.15	60.46	8.62	−1.33	.19	.35	.40	.53	.92
CBCL Externalizing Total T	111	55.70	10.57	57.85	11.60	−1.44	.15	.32	.15	.70	.96
SRS Total T	111	81.34	9.93	80.34	11.13	.71	.48	.58	.01	.92	.97
ADI-R Social Total	112	22.43	5.35	20.15	5.43	3.43	.001	.01	.18	.67	.96
ADI-R Nonverbal Communication Total	112	10.01	3.47	9.09	3.34	2.25	.03	.07	1.11	.29	.75
ADI-R RRB Total	112	6.76	2.59	6.49	2.59	.84	.40	.55	.01	.91	.97
ADOS Total CSS	112	7.79	1.53	7.29	1.82	1.26	.21	.37	.54	.46	.92
ADOS Social Affect CSS	109	7.47	1.69	7.07	1.84	2.23	.03	.07	2.48	.12	.50
ADOS RRB CSS	109	8.18	1.76	7.86	1.98	1.65	.10	.25	2.38	.12	.50
Overall Diagnostic Certainty	112	13.83	1.98	12.55	2.60	4.25	<.0001	.001	1.82	.18	.62

	n pairs	Matched Autism Controls (MAC)		High Confidence (HC)		Matched Autism Controls versus High Confidence			Group-by-NVIQ Interaction (NVIQ as moderator)		
						Test Statistic	*p*	FDR *p*	Test Statistic	*p*	FDR *p*
		Mean	*SD*	*Mean*	*SD*						
		n	*%*	*n*	*%*						
ADOS Module[a]	112					8.63	.003	.01	.20	.65	.96
1		31	28	19	17						
2		24	21	20	18						
3		53	47	71	63						
4		4	4	2	2						
High Confidence Autism Diagnosis[b]	112	97	87	71	63	13.20	.003	.003	.18	.67	.96
Family History Major Psych. Problems[c]	96	45	47	43	45	0.10	.76	.79	.94	.33	.77
Seizures (yes/no)	112	6	5	13	12	2.45	.12	.28	4.34	.04	.35
Language Deficit (yes/no)	112	40	36	24	21	7.24	.01	.02	.00	.97	.97

[a]Module was collapsed into 1/2 versus 3/4 for analysis.

[b]Certainty greater than 12.

[c]Controlling for ethnicity.

Note: na=Not applicable; NVIQ and VIQ = nonverbal and verbal IQ; ADI-R=Autism Diagnostic Interview, Revised; VABS=Vineland Adaptive Behavior Scales, Second Edition; PPVT=Peabody Vocabulary Test; CBCL=Child Behavior Checklist; SRS=Social Responsiveness Scale; ADOS=Autism Diagnostic Observation Schedule; CSS=Calibrated Severity Score; RRB=Restricted and Repetitive Behavior; FDR=false detection rate. 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 was 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. NVIQ-as-moderator refers to the interaction between group (HC versus MAC) and NVIQ in predicting the dependent variable, and informs the question of whether group differences depend on cognitive level.