



# HHS Public Access

Author manuscript

*Mol Psychiatry*. Author manuscript; available in PMC 2011 April 01.

Published in final edited form as:

*Mol Psychiatry*. 2010 October ; 15(10): 996–1005. doi:10.1038/mp.2009.41.

## “High Density SNP Association Study of the 17q21 Chromosomal Region Linked to Autism Identifies CACNA1G as a Novel Candidate Gene”

Samuel P. Strom<sup>1</sup>, Jennifer L. Stone<sup>1</sup>, John R. ten Bosch<sup>1</sup>, Barry Merriman<sup>1</sup>, Rita M. Cantor<sup>1,2,3</sup>, Daniel H. Geschwind<sup>1,2,3,4</sup>, and Stanley F. Nelson<sup>1,2,3,\*</sup>

<sup>1</sup>Department of Human Genetics, University of California, Los Angeles, CA, U.S.A

<sup>2</sup>Center for Neurobehavioral Genetics, University of California, Los Angeles, CA, U.S.A

<sup>3</sup>Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, CA, U.S.A

<sup>4</sup>Department of Neurology, University of California, Los Angeles, CA, U.S.A

### Abstract

Chromosome 17q11-q21 is a region of the genome likely to harbor susceptibility to autism (MIM[209850]) based on prior evidence of linkage to the disorder. This linkage is specific to multiplex pedigrees containing only male probands (MO) within the Autism Genetic Resource Exchange (AGRE). Previously, Stone et al.<sup>1</sup> completed a high-density SNP association study of 13.7Mb within this interval, but common variant association was not sufficient to account for the linkage signal. Here we extend this SNP-based association study to complete the coverage of the 2 LOD support interval around the chromosome 17q linkage peak by testing the majority of common alleles in 284 MO trios.

**CONCLUSIONS**—Markers within an interval containing the gene CACNA1G were found to be associated with Autism Spectrum Disorder at a locally significant level ( $p = 1.9 \times 10^{-5}$ ). While establishing CACNA1G as a novel candidate for autism, these alleles do not contribute sufficient genetic effect to explain the observed linkage, indicating there is substantial genetic heterogeneity despite the clear linkage signal. The region thus likely harbors a combination of multiple common and rare alleles contributing to the genetic risk. These data, along with previous studies of Chromosomes 5 and 7q3, suggest few if any major common risk alleles account for ASD risk under major linkage peaks in the AGRE sample. This provides important evidence for strategies to identify ASD genes, suggesting they should focus on identifying rare variants and common variants of small effect.

### Keywords

Autism; Autism Spectrum Disorder; Association; Chromosome 17q; CACNA1G

---

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:[http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

\*Corresponding Author Contact information: Stanley F. Nelson Phone: (310) 794 7981 Email: snelson@ucla.edu Street Address: 695 Charles E. Young Drive South Gonda 5554 Los Angeles, CA 90095-7088 .

## INTRODUCTION

Autism (MIM[209850]) is a Pervasive Developmental Disorder (PDD) defined by impairment along three dimensions: language development, development of social behaviors, and the presence of stereotypic or rigid behavior. The diagnosis of “autistic disorder” encompasses a broad range of phenotypically diverse conditions with wide variation along the three dimensions of impairment, making autism a particularly heterogeneous disorder. “Autistic Disorder” is commonly grouped with Asperger’s Syndrome (MIM[608638]) and Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS) under the umbrella of Autism Spectrum Disorders (ASD).

Twin and family studies have provided strong evidence of heritability and suggest a high likelihood of genetic contribution for the susceptibility to autism. The monozygotic twin concordance rate is reported as high as 90% for ASD, while sibling concordance rates are approximately 10%<sup>2, 3</sup>. This indicates a strongly heritable yet genetically complex disorder<sup>4</sup>. The sibling relative risk is approximately 15-20 fold higher than the population frequency. The inheritance pattern of ASD is not consistent with a Mendelian disease model and can likely be better explained by the involvement of multiple interacting loci and environmental factors<sup>5</sup>. However, the degree of genetic complexity not been established.

The high prevalence<sup>6, 7</sup> and strong heritability of ASD has encouraged multiple groups to complete whole-genome linkage studies<sup>8</sup> searching for genomic regions likely harboring autism susceptibility alleles. One of the few genomic regions identified in an initial linkage studies and replicated at genome-wide significance is 17q11-q21<sup>9, 10</sup>. The linkage signal was strengthened based on stratification of linkage data conditioned on the sex of affected siblings, which resulted in genome-wide significant linkage centered at 25-28 mega-bases (Mb) in multiplex families with exclusively male probands (male-only, or MO families) within the Autism Genetic Resource Exchange (AGRE)<sup>11</sup>. A replication study in a set of 109 additional MO families from AGRE showed evidence of sex specific linkage extending over the same region, and fine mapping identified a region of linkage extending an additional 18 Mb from the end of the initial linkage peak<sup>9</sup> centered at 61cM.

Previously, Stone *et al.*<sup>1</sup> tested approximately half of the linkage interval on 17q11-q21 for association to common variants via high density single nucleotide polymorphism (SNP) genotyping. While the results demonstrated suggestive evidence of association of ASD to several interesting and novel candidates, the association signals were not sufficiently strong to account for the observed linkage signal within the AGRE MO families. Thus, we sought to cover the remaining likely linkage interval by testing common DNA variants comprehensively within the remaining 17q linkage region defined by fine mapping<sup>9</sup>, testing 1975 SNPs at an average marker density of 6.3 kilo-bases (kb). This provided strong coverage of the majority of common haplotypes over the extended region of linkage, testing for ASD association within 295 genes in 284 independent trios from MO families in AGRE. We report the overall association analyses, which highlight CACNA1G as a novel candidate gene.

## MATERIALS AND METHODS

### Genetic Material and Preparation

The Autism Genetic Resource Exchange (AGRE), an organization facilitating the collection of biomaterials and phenotypic information of families with autistic individuals, provided DNA samples for this study. AGRE has a standardized set of criteria for inclusion, which have been previously published<sup>12</sup> and are available at [www.agre.org](http://www.agre.org). AGRE focuses on collecting genetic material from families with more than one individual diagnosed with Autism Spectrum Disorder (as defined by Liu *et al.*<sup>13</sup>).

For this study, both parents and one affected son were typed in 302 MO trios. After data cleaning, 296 remained (Supplemental Materials 1). Families containing individuals flagged for non-idiopathic autism because of other medical conditions such as Fragile-X Syndrome, birth trauma, and dysmorphic features, were not included in this study. Minor physical anomalies (MPA) as described by Ozgen *et al.*<sup>14</sup> were not used as exclusion criteria, as this is a newly emerging area of phenotyping and has not yet been included in the AGRE cohort or other large genetic studies such as the Autism Genome Project (AGP)<sup>15</sup>. Every family meeting these criteria and having genetic materials available through AGRE at the time of assay design were included in this study.

Ethnicity was recorded via self-report, with 79% of those reporting identified as Caucasian. These reports are consistent with population structure analyses of the genotype data using the Structure software package<sup>16</sup> (data not shown). Due to the family-based design and the out-bred nature of the population sampled (United States), ethnicity is not used as a factor in this study. This distribution will lead to an over-representation of Caucasian alleles overall, which will in turn increase the probability that discovered associated alleles will be specific to or enriched in the United States Caucasian population. The follow-up sample has a similar ethnicity distribution.

All subjects were diagnosed using the Autism Diagnosis Interview Revised (ADI-R)<sup>17</sup>. A subset of subjects were also diagnosed using the Autism Diagnostic Observation Schedule (ADOS)<sup>18</sup>. In total, 12 out of 296 individuals diagnosed as either autistic or broad spectrum on the ADI-R scored as 'Not Spectrum or Autism' on the ADOS (Table 1). These individuals were excluded from study due to their ambiguous phenotype. The 284 subjects surviving genotype cleaning and diagnostic criteria were included for all subsequent analysis.

Genomic DNA samples were obtained from the NIMH cell repository (Rutgers, Piscataway, NJ). Concentrations were determined using the Nanodrop (Wilmington, DE) instrument. The UCLA Internal Review Board (IRB) has approved all aspects of this study. A complete list of samples used in this study, along with gender, demographic, and diagnostic data are found in Supplementary Table 1.

### SNP Selection

SNPs were selected using several criteria, including linkage disequilibrium (LD) data from the Hapmap project and ability to develop a working genotyping assay. To perform the SNP

selection, we first requested all possible dbSNP genotypes that will perform well in the Golden Gate assay per manufacturer's guidelines (Illumina Inc., San Diego, CA). To maximize the amount of common variation tested while minimizing the number of markers typed, the software package Tagger19 was used to select a subset of these markers for genotyping such that the subset would cover the region of interest in the Hapmap (Build #21, July 2006) CEPH population with a minimum  $r^2$  value of 0.9. In total, 2042 SNPs were selected for genotyping in the region. Markers span 34.3Mb-47Mb on chromosome 17 (17q12-q21.33). 1975 SNPs meet the following criteria: Hardy-Weinberg Equilibrium P-Value 0.001, genotyping rate 85%, Mendelian Error Rate 1%.

### SNP Genotyping

All SNPs were genotyped using a custom "BeadArray" DNA micro-array created by Illumina, Inc. (San Diego, CA) and the Golden Gate assay design. Genotyping was completed within the Southern California Genotyping Consortium, a local installation of the Illumina genotyping system at UCLA. Array assays were performed following standard protocols<sup>20</sup>.

### Follow-Up Association

A formal replication set of sufficient size and equivalent gender stratification was not available to us. As a surrogate for formal replication, a follow-up association of 21 SNPs within CACNA1G was performed on 1046 affected trios from 556 AGRE pedigrees genotyped by the Children's Hospital of Philadelphia (CHOP, AGRE Illumina HumanHap550 data). All pedigrees in this follow-up sample are independent of the initial sample. The male to female ratio of affected individuals in the follow-up sample is 2.11:1. All available samples meeting identical criteria to the initial association were tested for association. Siblings of the original sample were analyzed separately (262 individuals available from CHOP).

### Association Analyses

Single SNPs were tested for association to ASD by performing a transmission disequilibrium test (TDT) using PLINK software.<sup>21</sup> The PLINK TDT test computes a  $\chi^2$  statistic to assess the presence and significance of transmission biases. The '--perm' option was used to calculate an empiric p-value for each SNP based on an adaptive permutation model. In this model, SNPs found to be non-associated after a low number of tests are dropped while SNPs with higher initial association are tested by up to millions of permutations.

To allow for haplotype analysis, groups of multiple markers were defined as haplotypes using the Four Gamete Test (FGT) implemented within Haploview<sup>22</sup>, identifying 340 block intervals. This test robustly assesses whether a recombination has occurred between two markers in a population. Haplotype-based TDT testing within Haploview was performed to assess the transmission bias of the blocks defined by the FGT.

## Copy Number Variant Detection

Signal intensity values in the form of Log R Ratios were used to identify regions of chromosome gain or loss, or copy number variations (CNV). Individual samples were normalized on both per sample and per marker basis and analyzed across five marker windows. Windows with mean score absolute values above three standard deviations were considered as either gain or loss, respectively. Detailed description of this method is found in Supplemental Materials 3.

## RESULTS

### Single SNP and Haplotype-based Association

To assess whether common variants in the chromosomal interval 17q11-q21 are associated with ASD, both single SNP and haplotype-based transmission disequilibrium tests were performed on 284 MO trios using 1975 markers. Full results from these tests can be found in Figure 1 and Supplemental Table 2. Tagger software within Haploview 22 was used to assess the number of independent observations made by single SNP association. Using a linkage disequilibrium threshold of  $r^2 \geq 0.8$ , a total of 1452 single SNPs were sufficient to capture 100% of the alleles present in the sample. Thus, a stringent Bonferroni cutoff of  $0.05 / 1452$  was used (corresponding to a  $-\log_{10}$  P-Value of 4.463; P-Value = 0.000034435). Two SNPs, rs757415 ( $p=0.000019$ ) and rs12603112 ( $p=0.000021$ ), met criteria for multiple testing correction (Table 2). Both markers are located within intron 9 of voltage-dependent calcium channel alpha 1G (CACNA1G, Figure 2) and are in strong linkage disequilibrium ( $r^2=0.99$ ), indicating they tag a single allele contributing to increased ASD risk. While other alleles have nominally significant association, none is associated with ASD based on interval corrected significance levels.

Haplotype-based association yielded no locally significant blocks associated with ASD given the number of independent blocks assayed (Figure 3, Table 3). The block with the highest evidence of association (P-Value =  $7.68 \times 10^{-5}$ , permutation P-Value = 0.068) was an over-transmission of a five-marker block including the two single SNPs identified as significantly associated with ASD (Figure 3). The FGT defined 14kb haplotype spans exons 7-9 of the CACNA1G gene. Sixty-one affected individuals inheriting this haplotype from both parents drive this association, suggesting it may act as, or be tightly linked to, a recessive risk allele. Complete haplotype association results are found in Supplementary Table 3. A linkage disequilibrium plot for the SNPs tested within CACNA1G was plotted (Figure 4) and demonstrates the high degree of linkage disequilibrium (LD) between SNPs in the CACNA1G interval and the haplotypic structure of this genomic interval. Causal alleles underlying the association signal may reside anywhere within the gene.

### Copy Number Variant Analysis

A comprehensive survey of copy number variation in the interval was performed to detect CNVs approximately 15kb or larger. Only one affected individual was identified as having region-wide significant copy number variant in the assayed interval using in-house developed analytical methods to analyze probe intensity data (Supplemental Materials 2, B. Merriman available on request). An approximately 28kb hemizygous loss was identified in

affected individual AU0920301, but not either parent (Supplemental Figure 1). The genomic interval spanned by this apparent de novo chromosome loss contains exon one of WFIKKN2 and all of TOB1. Neither gene was identified via association or is a clear functional candidate. To date, no common CNVs in this interval have been reported<sup>23</sup>. While this CNV may contribute to ASD in this isolated case, it does not represent a general risk factor for autism. Further study is required to determine if this loss is functional or merely a benign event or statistical false positive. In total, no novel or previously identified CNVs 15kb in were detected in affected children within the interval. Shorter length variants may be present, but were not detectable given the density of markers assayed.

### Follow-Up Association

A follow-up transmission bias test of 21 SNPs within CACNA1G in 1046 affected offspring trios yielded one marker exhibiting nominal association with ASD with an empiric P-value of 0.028 (Supplementary Table 4). The associated marker rs198547 is located within intron 7 of CACNA1G. While these data do not represent replication of the same SNPs identified in the first stage, the nominal association of a SNP within CACNA1G in this sample in conjunction with the first stage analysis stresses the need for further study of variants in this gene. Further, we note that the replication set is not ideally powered to uncover risk alleles specifically related to the MO families, as the replication set is relatively devoid of male only affected sibling pairs which was the basis of the original linkage and now association finding. The marker rs12603112 is one of the two SNPs found to be associated with ASD at a locally significant level in the initial sample. A modest over-transmission of rs12603122 was observed in affected siblings of the initial sample set (T:NT of 93:72), which is consistent with the original sample but not significant ( $p = 0.16$ ).

## DISCUSSION

Autism Spectrum Disorder (ASD) encompasses a diverse array of phenotypes. While highly heritable, ASD is a heterogeneous set of conditions as evidenced by the common distinct presentation of affected relatives<sup>2, 3, 12, 24-26</sup>. Due to this heterogeneity, few linkage regions have been confirmed by replication or validated by the discovery of risk alleles. Linkage within male-only multiplex pedigrees (MO) on chromosome 17q is one of the few genomic intervals replicated in ASD.

In this region of convincing *a priori* evidence for linkage, the common variant hypothesis was tested by performing association of 1975 markers spanning the 12.6Mb interval on chromosome 17q linked to autism spectrum disorder (ASD). With an average inter-marker distance of 6.3kb, 85% of known common haplotypes were detected, covering the complete 2-LOD support interval. In this 12.6Mb region, two single SNPs in the interval were associated with ASD at a region-wide significant level, given conservative correction for the number of tests performed. These SNPs are adjacent and in strong LD ( $r^2 0.99$ ), and are located within intron 9 of the calcium channel gene CACNA1G. While this association could be considered relatively strong (odds ratio 95% confidence interval = 1.32-2.21,  $p = 0.000019$ ) for a complex disease such as ASD in a sample of this size, it does not fully explain the strength of the linkage signal. Multiple alleles with subtle effects or, more likely,



yet to be discovered rare variants in the 17q linkage region must also contribute to ASD risk in these families.

Since we have no *a priori* knowledge of how to identify families with a more homogeneous etiology, true risk variants are not distinct from false positives occurring due to the large number of tests required to rigorously assay a region of this size. Both the reduction of true association due to partial penetrance or low effect size, and the inflation of false association due to chance, impact heavily on such studies<sup>27, 28</sup>. Thus, the primary goal of this study was to reduce the entire region of linkage to a tenable number of nominally associated common polymorphisms such that genes can be highlighted for further study — including replication and re-sequencing — to better assess which nominal associations are representative of true risk alleles.

The region covered by this study (17q12-q21.33, 33.3Mb-47.0Mb, NCBI Build 36) contains 295 known genes according to the RefSeq database<sup>29</sup>. Of these, only 8 genes are adjacent to single SNPs or haplotypes even nominally associated with ASD at  $p = 0.01$  in this study (Supplemental Tables 2 and 3). The total number of single SNPs associated (17/1975, ~1%) at this level is consistent with the number one would expect to find due to chance. The number of haplotypes associated (9/340, 2.5%) may be slightly elevated, but does not significantly differ from expectations. Thus, there does not appear to be an overall enrichment of associated SNPs or haplotypes with ASD in this interval, supporting the rare variant hypothesis.

Alleles of two SNPs (rs757415 and rs12603112) within *CACNA1G* were significantly over-transmitted with P-Values  $\approx 2.1 \times 10^{-5}$ . The calculated odds ratio of 1.30-2.21 (95% confidence interval) for these two markers is within the range of expectation for a minor effect allele, and is typical of common allele associations. Haplotype analysis identified a 5-marker block of SNPs (containing these two markers) over-transmitted to probands (T:NT = 154:92,  $p = 7.86 \times 10^{-5}$ ). While the associated SNPs within the *CACNA1G* gene interval are intronic and do not have a predicted effect on gene expression or splicing, the associated haplotype spans a large portion of the gene, including several exons (Figure 2). One marker in a follow-up association sample is nominally associated with ASD. The follow-up set does not have an ideal gender stratification pattern, and the associated SNP is not in strong LD with the markers associated in the original study. Despite these limitations, finding nominal association in a separate sample of ASD patients supports *CACNA1G* as a positional candidate gene for the disorder. Conditioning the association on only affected male sibling pairs did not increase the strength of association (data not shown). While we are cautious of over interpreting this due to the small sample size, this may indicate the effects of *CACNA1G* genetic variation are not limited to MO families.

*CACNA1G* encodes a T-type voltage gated calcium channel<sup>30-32</sup> which has been linked by a previous study to idiopathic generalized epilepsy<sup>32</sup>, indicating the gene modulates neuronal excitability and neural transmission. Varying reports have indicated a co-morbidity of epilepsy with autism at 5-38%. None of the probands in this study are known to have co-morbidity with a seizure disorder, but it is reasonable to consider that uncharacterized mutations with different phenotypic effects in this gene could contribute to ASD. We

determined the relative frequency of the putative risk alleles in a subset of the AGRE sample for which EEG information was available. 208 of the genotyped affected individuals had EEG data available. There was no significant increase in the frequency of individuals with the associated allele (rs12603112) and EEG abnormality Supplemental Table 5). A larger sample size is required to assess the impact of CACNA1G haplotypes and abnormal brain activity or with more refined and complete phenotypic data.

Recent research has implicated calcium signaling and homeostasis as a possible molecular mechanism of autism<sup>33, 34</sup>, and several studies have linked mutations in other voltage-gated calcium channels to syndromic<sup>35, 36</sup> and non-syndromic autism<sup>37</sup>. CACNA1G is a strong functional candidate especially as a sexually dimorphic risk allele. There is known sexually dimorphic expressivity of this gene due to estrogen which may account for the association signal being detected in MO families<sup>38</sup>. Further, rare mutations in related calcium channels have been previously described<sup>33</sup>. Mutations in CACNA1C - a voltage gated calcium channel gene - cause Timothy Syndrome<sup>36</sup>, which can include autistic features. A CACNA1F mutation has been identified in a pedigree with night blindness, in which some of the male probands are also affected with autism<sup>35</sup>. Heterozygous mutations in CACNA1H were detected in 6 of 461 cases of non-syndromic autism<sup>37</sup>. Thus, by analogy and the work presented here, CACNA1G is a strong functional candidate and should be assessed for rare variants and spontaneous mutations.

These data support CACNA1G as a strong candidate gene for ASD in at least a subset of cases. While the results of a follow-up association analysis in 1046 affected offspring trios (556 new independent pedigrees) supports CACNA1G as an ASD candidate gene, replication in a large gender matched sample is required to assess the veracity of the association. Re-sequencing to detect potential rare variants in this gene is warranted. Within the analysis of the remainder of the interval, genes warranting further replication in the tested interval, identified as containing nominally associated SNPs or haplotypes include: CACNB1, DUSP3, KIF18B, SKAP1, and CRHR1. Of these, three (CACNB1, DUSP3, and CRHR1) have known transcriptional patterns or biological activities suggesting they play a role in brain function, and are biological and positional candidates for ASD also meriting further study to assess potential risk alleles. Copy Number Variant detection identified a single individual with an apparent spontaneous hemizygous loss spanning approximately 28kb. Two genes, TOB1 and WFIKKN2, are potentially disrupted by this event. However, the implications of such a loss are not clear given the isolated nature of the event and lack of association in the interval. No structural variants in CACNA1G were detected.

This study has identified the calcium channel subunit gene CACNA1G as a novel candidate gene for Autism Spectrum Disorder. Further investigation, including replication to confirm common variants and DNA re-sequencing to identify rare variants, is required to establish the role CACNA1G may play in the etiology of the disorder. Combined with our previous studies<sup>1</sup>, the majority of common variants in the replicated ASD linkage region on chromosome 17q have now been tested for association. These findings demonstrate that while multiple loci in the interval may contribute to the disorder, the sum of the effects of individual common variants alone cannot explain the strength and replication of linkage.



Rather, gene-gene interactions, gene-environment interactions, rare variants, or other mechanisms not yet implicated must contribute to the etiology of this complex disease.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

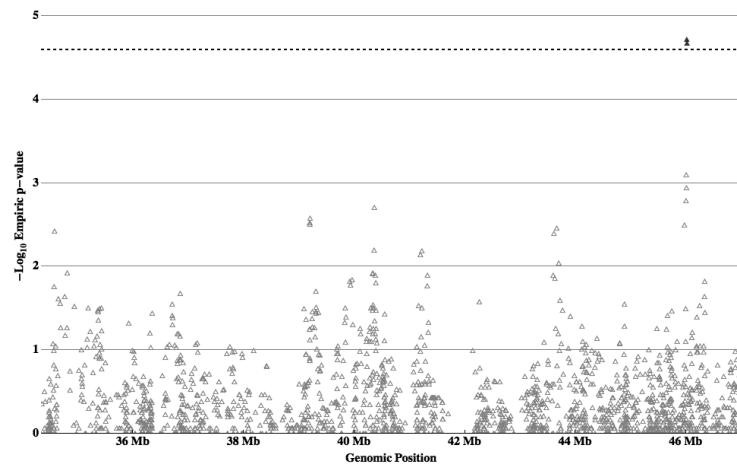
We thank Dr. Vlad Kustanovich at AGRE and Dr. Sarah Spence formerly at AGRE for their support and guidance. We gratefully acknowledge the resources provided by the Autism Genetic Resource Exchange (AGRE) Consortium and the participating AGRE families. The Autism Genetic Resource Exchange is a program of Autism Speaks and is supported, in part, by grant 1U24MH081810 from the National Institute of Mental Health to Clara M. Lajonchere (PI). AGRE Illumina HumanHap550 data were generated at the Children's Hospital of Philadelphia and the University of Pennsylvania via collaboration between the labs of Dr. Hakon Hakonarson and Dr. Maja Bucan. These data were provided to AGRE by Drs. Bucan and Hakonarson. We thank them for their contribution. The Autism Genetic Resource Exchange is a program of Autism Speaks and is supported, in part, by grant MH64547 from the National Institute of Mental Health to Daniel H. Geschwind (PI). This work was supported by MH64547, NS052108, and a grant from Cure Autism Now.

## References

1. Stone J, Merriman B, Cantor R, Geschwind D, Nelson S. High density SNP association study of a major autism linkage region on chromosome 17. *Hum Mol Genet.* 2007; 16(6):704–715. [PubMed: 17376794]
2. Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E. Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med.* 1995; 25(1):63–77. [PubMed: 7792363]
3. Ritvo E, Freeman B, Mason-Brothers A, Mo A, Ritvo A. Concordance for the syndrome of autism in 40 pairs of afflicted twins. *Am J Psychiatry.* 1985; 142(1):74–77. [PubMed: 4038442]
4. Risch N. Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet.* 1990; 46(2):222–228. [PubMed: 2301392]
5. Risch N, Spiker D, Lotspeich L, Nouri N, Hinds D, Hallmayer J, et al. A genomic screen of autism: evidence for a multilocus etiology. *Am J Hum Genet.* 1999; 65(2):493–507. [PubMed: 10417292]
6. Chakrabarti S, Fombonne E. Pervasive developmental disorders in preschool children: confirmation of high prevalence. *Am J Psychiatry.* 2005; 162(6):1133–1141. [PubMed: 15930062]
7. Rice C. Prevalence of autism spectrum disorders--autism and developmental disabilities monitoring network, six sites, United States, 2000. *MMWR Surveill Summ.* 2007; 56(1):1–11.
8. Wassink T, Brzustowicz L, Bartlett C, Szatmari P. The search for autism disease genes. *Ment Retard Dev Disabil Res Rev.* 2004; 10(4):272–283. [PubMed: 15666342]
9. Cantor R, Kono N, Duvall J, Alvarez-Retuerto A, Stone J, Alarcón M, et al. Replication of autism linkage: fine-mapping peak at 17q21. *Am J Hum Genet.* 2005; 76(6):1050–1056. [PubMed: 15877280]
10. Yonan A, Alarcón M, Cheng R, Magnusson P, Spence S, Palmer A, et al. A genomewide screen of 345 families for autism-susceptibility loci. *Am J Hum Genet.* 2003; 73(4):886–897. [PubMed: 13680528]
11. Stone J, Merriman B, Cantor R, Yonan A, Gilliam T, Geschwind D, et al. Evidence for sex-specific risk alleles in autism spectrum disorder. *Am J Hum Genet.* 2004; 75(6):1117–1123. [PubMed: 15467983]
12. Geschwind D, Sowinski J, Lord C, Iversen P, Shestack J, Jones P, et al. The autism genetic resource exchange: a resource for the study of autism and related neuropsychiatric conditions. *Am J Hum Genet.* 2001; 69(2):463–466. [PubMed: 11452364]
13. Liu J, Nyholt D, Magnussen P, Parano E, Pavone P, Geschwind D, et al. A genomewide screen for autism susceptibility loci. *Am J Hum Genet.* 2001; 69(2):327–340. [PubMed: 11452361]

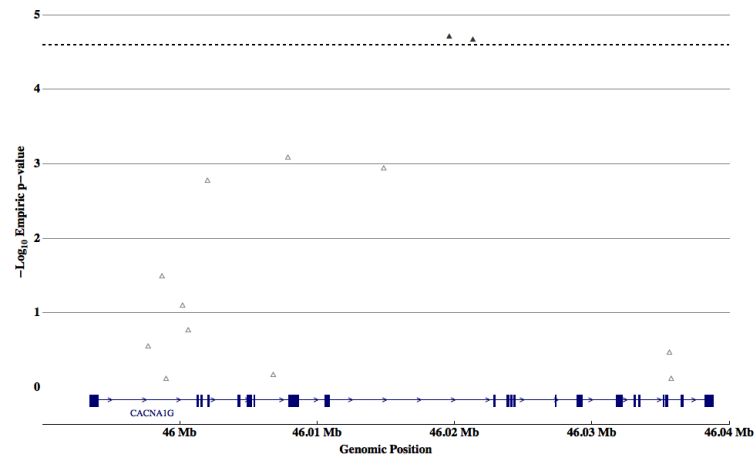
14. Ozgen H, Hop J, Hox J, Beemer F, van Engeland H. Minor physical anomalies in autism: a meta-analysis. *Mol Psychiatry*. 2008
15. Hu-Lince D, Craig D, Huentelman M, Stephan D. The Autism Genome Project: goals and strategies. *Am J Pharmacogenomics*. 2005; 5(4):233–246. [PubMed: 16078860]
16. Pritchard J, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000; 155(2):945–959. [PubMed: 10835412]
17. Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord*. 1994; 24(5):659–685. [PubMed: 7814313]
18. Lord C, Rutter M, Goode S, Heemsbergen J, Jordan H, Mawhood L, et al. Autism diagnostic observation schedule: a standardized observation of communicative and social behavior. *J Autism Dev Disord*. 1989; 19(2):185–212. [PubMed: 2745388]
19. de Bakker P, Yelensky R, Pe'er I, Gabriel S, Daly M, Altshuler D. Efficiency and power in genetic association studies. *Nat Genet*. 2005; 37(11):1217–1223. [PubMed: 16244653]
20. Oliphant A, Barker D, Stuelpnagel J, Chee M. BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. *Biotechniques*. 2002; (Suppl):56–58. 60–51. [PubMed: 12083399]
21. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet*. 2007; 81(3):559–575. [PubMed: 17701901]
22. Barrett J, Fry B, Maller J, Daly M. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005; 21(2):263–265. [PubMed: 15297300]
23. Iafrate A, Feuk L, Rivera M, Listewnik M, Donahoe P, Qi Y, et al. Detection of large-scale variation in the human genome. *Nat Genet*. 2004; 36(9):949–951. [PubMed: 15286789]
24. Ma D, Cuccaro M, Jaworski J, Haynes C, Stephan D, Parod J, et al. Dissecting the locus heterogeneity of autism: significant linkage to chromosome 12q14. *Mol Psychiatry*. 2007; 12(4):376–384. [PubMed: 17179998]
25. Szatmari P. Heterogeneity and the genetics of autism. *J Psychiatry Neurosci*. 1999; 24(2):159–165. [PubMed: 10212560]
26. Alarcón M, Yonan A, Gilliam T, Cantor R, Geschwind D. Quantitative genome scan and Ordered-Subsets Analysis of autism endophenotypes support language QTLs. *Mol Psychiatry*. 2005; 10(8):747–757. [PubMed: 15824743]
27. Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet*. 2003; 33(Suppl):228–237. [PubMed: 12610532]
28. Carlson C, Eberle M, Kruglyak L, Nickerson D. Mapping complex disease loci in whole-genome association studies. *Nature*. 2004; 429(6990):446–452. [PubMed: 15164069]
29. Pruitt K, Tatusova T, Maglott D. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res*. 2005; 33(Database issue):D501–504. [PubMed: 15608248]
30. Blanks A, Zhao Z, Shmygol A, Bru-Mercier G, Astle S, Thornton S. Characterization of the molecular and electrophysiological properties of the T-type calcium channel in human myometrium. *J Physiol*. 2007; 581(Pt 3):915–926. [PubMed: 17446221]
31. Perez-Reyes E, Lory P. Molecular biology of T-type calcium channels. *CNS Neurol Disord Drug Targets*. 2006; 5(6):605–609. [PubMed: 17168745]
32. Singh B, Monteil A, Bidaud I, Sugimoto Y, Suzuki T, Hamano S, et al. Mutational analysis of CACNA1G in idiopathic generalized epilepsy. Mutation in brief #962. *Hum Mutat*. 2007; 28(5):524–525. [PubMed: 17397049]
33. Krey J, Dolmetsch R. Molecular mechanisms of autism: a possible role for Ca<sup>2+</sup> signaling. *Curr Opin Neurobiol*. 2007; 17(1):112–119. [PubMed: 17275285]
34. Palmieri L, Papaleo V, Porcelli V, Scarcia P, Gaita L, Sacco R, et al. Altered calcium homeostasis in autism-spectrum disorders: evidence from biochemical and genetic studies of the mitochondrial aspartate/glutamate carrier AGC1. *Mol Psychiatry*. 2008

35. Hemara-Wahanui A, Berjukow S, Hope C, Dearden P, Wu S, Wilson-Wheeler J, et al. A CACNA1F mutation identified in an X-linked retinal disorder shifts the voltage dependence of Cav1.4 channel activation. *Proc Natl Acad Sci U S A*. 2005; 102(21):7553–7558. [PubMed: 15897456]
36. Splawski I, Timothy K, Sharpe L, Decher N, Kumar P, Bloise R, et al. Ca(V)1.2 calcium channel dysfunction causes a multisystem disorder including arrhythmia and autism. *Cell*. 2004; 119(1): 19–31. [PubMed: 15454078]
37. Splawski I, Yoo D, Stotz S, Cherry A, Clapham D, Keating M. CACNA1H mutations in autism spectrum disorders. *J Biol Chem*. 2006; 281(31):22085–22091. [PubMed: 16754686]
38. Qiu J, Bosch M, Jamali K, Xue C, Kelly M, Rønnekleiv O. Estrogen upregulates T-type calcium channels in the hypothalamus and pituitary. *J Neurosci*. 2006; 26(43):11072–11082. [PubMed: 17065449]



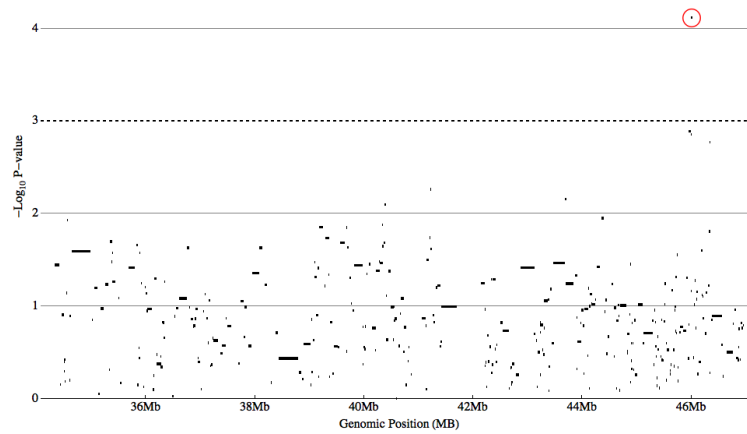
**Figure 1. Single marker Association Results**

Individual SNP transmission biases from 284 Male-Only autism trios are plotted as  $-\text{Log}_{10}$  of the empiric P-Value versus genomic position in mega-bases (Mb). Local significance threshold is represented by a black dotted line. Two markers (filled triangles) out of 1975 markers tested are associated at a locally significant level, adjacent SNPs within CACNA1G.



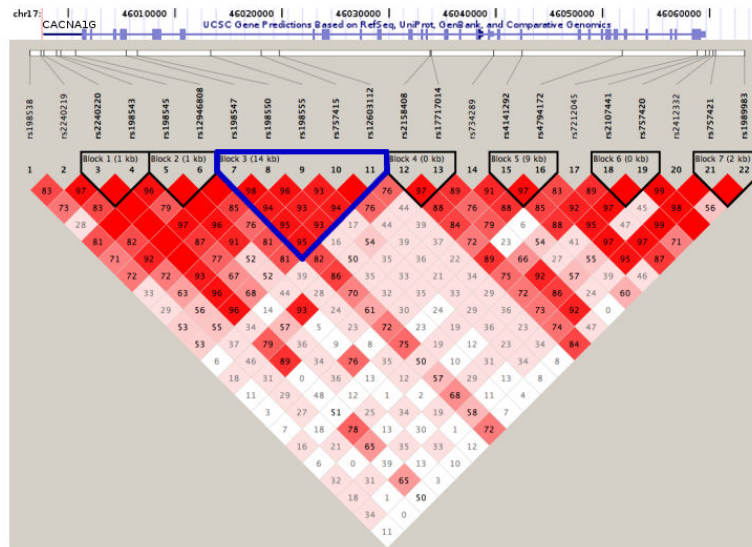
### Figure 2. Single marker Association Results for CACNA1G

Individual SNP transmission biases from 284 Male-Only autism trios of markers within the genetic interval containing the CACNA1G gene are plotted as  $-\text{Log}_{10}$  of the empiric P-Value versus genomic position in mega-bases (Mb). Local significance threshold is represented by a black dotted line. Two markers (rs757415 and rs12603112; filled triangles) are associated at a locally significant level. A gene diagram representing exons (blue rectangles), introns (blue lines), and direction of transcription (blue arrowheads) is included for reference.

**Figure 3. Haplotype Association Results**

Haplotype transmission biases from 284 Male-Only autism trios are plotted as  $-\text{Log}_{10}$  of the P-Value versus genomic position in mega-bases (Mb). For each block locus, only the block with the most significant P-Value is plotted. Nominal local significance threshold (P-Value 0.001) is represented by a black dotted line. One haplotype (circled) are associated at this nominal level. This block contains the two individual SNPs found to be associated at a locally significant level.





**Figure 4. Linkage Disequilibrium Status of SNPs within CACNA1G**  
 Haplotypes estimated from 284 Male-Only autism trios using the Four Gamete Test (FGT) are plotted as a function of  $D'$ . Approximate location of CACNA1G in relation to the markers is represented by the UCSC Genome Browser track at the top of the figure. Block 3 (highlighted in blue) is the block most associated with Autism Spectrum Disorder in this study ( $P$ -Value  $< 7.68 \times 10^{-5}$ ).

**Table 1**

**Gender, diagnosis, ethnicity, and age information**

Number of affected individuals genotyped (“N”), count and percentage of male affected individuals (“Male”), ADI-R diagnosis counts and percentages for “Autism”, “BroadSpectrum”, and “NQA”, and ADOS diagnosis counts and percentages (for available subjects) for “Autism” and “Spectrum”, count and percentage of affected individuals reporting as “Caucasian”, and the mean age of subjects upon ADI-R diagnosis are listed. Percentages reported for ADOS diagnosis reflect only those subjects with available ADOS diagnosis data.

	N	ADI-R Diagnosis		ADOS Diagnosis		Mean Age at Diagnosis	
		Autism	Broad-Spectrum	Autism	Spectrum	Males	Caucasian
Initial	Count:	281	4	182	29	290	229
	Percent:	97%	1%	86%	14%	100%	79%
Siblings	Count:	207	33	161	27	262	197
	Percent:	79%	13%	86%	14%	100%	75%
Follow-up	Count:	924	82	667	154	710	759
	Percent:	88%	8%	81%	19%	68%	73%

**Table 2**

**Single marker Association Results**

The 17 single SNPs associated with ASD with an Empiric P-Value 0.01 are shown sorted by the strength of association. Marker identifier ('SNP'), genomic position of Chromosome 17 ('Position'), Major and Minor Allele Genotypes ('Maj.', 'Min. '), Transmitted versus Non-Transmitted count ('T:NT'), Odds Ratio, empiric p-value, and gene names (if any) corresponding to each associated marker are listed. The top two markers, associated with ASD at a locally significant level, are highlighted in bold.

SNP	Position	Maj.	Min.	T:NT	Odds Ratio	Empiric P-Value	Gene
<b>rs757415</b>	<b>46019628</b>	<b>T</b>	<b>C</b>	<b>157:92</b>	<b>2.207</b>	<b>0.000019</b>	<b>CACNA1G</b>
<b>rs12603112</b>	<b>46021348</b>	<b>G</b>	<b>A</b>	<b>157:92</b>	<b>2.207</b>	<b>0.000021</b>	<b>CACNA1G</b>
rs198550	46007874	A	G	100:154	0.8352	0.0007873	CACNA1G
rs198555	46014842	A	G	107:158	0.8656	0.001118	CACNA1G
rs12946808	46002010	G	T	164:111	1.88	0.001619	CACNA1G
rs736866	40365323	A	C	113:166	0.8645	0.001955	KIF18B
rs4793026	39206034	G	A	72:113	0.8563	0.002642	DUSP3
rs11713	39201350	T	C	73:113	0.8671	0.002948	DUSP3
rs17742347	39201994	T	C	71:110	0.8699	0.003102	DUSP3
rs4793661	45975516	A	G	98:58	2.271	0.003131	EPN3
rs1377201	43667127	A	G	152:107	1.819	0.003443	SKAPI
rs17633541	34583536	T	C	29:11	5.277	0.003716	CACNB1
rs2278868	43617170	C	T	153:108	1.812	0.004006	SKAPI
rs9908100	40363477	C	T	80:116	0.9169	0.006334	KIF18B
rs7209436	41225913	T	C	114:161	0.9001	0.00645	CRHR1
rs6503448	41198241	G	A	119:164	0.9188	0.007167	
rs2924254	43703383	C	T	111:156	0.9076	0.008958	SKAPI

**Table 3****Haplotype Association Results**

The 9 Haplotype blocks associated with ASD with a P-Value 0.01 are shown sorted by the strength of association. Haplotype, frequency ('Freq.'), genomic range, ('Start' and 'End'), Transmission versus Non-transmission ('T:NT'), P-Value, genes within the genomic interval (if any), and SNPs are listed for each haplotype. The haplotype meeting nominal significance (P-Value 0.001) within CACNA1G is highlighted in bold.

Haplotype	Freq.	Start	End	T:NT	P-Value	Gene	SNPs
<b>GGGTG</b>	<b>0.384</b>	<b>46006810</b>	<b>46021348</b>	<b>154:91</b>	<b>7.68x10<sup>-5</sup></b>	<b>CACNA1G</b>	<b>rs198547;rs198550; rs198555;rs757415; rs12603112</b>
CGAA-CATG	0.16	45968836	45984457	93:54	0.0013		rs3785915;rs2306001; rs2306002;rs4793661; rs8076632;rs1132414; rs9913430;rs8065903
GG	0.497	46000610	46002010	164:111	0.0014	CACNA1G	rs198545;rs12946808
GCG	0.028	46346356	46349271	25:7	0.0017		rs16949486; rs11869222; rs8066630
GAACA	0.343	46006810	46021348	99:143	0.004	CACNA1G	rs198547;rs198550; rs198555;rs757415; rs12603112
TA	0.461	41225913	41226462	114:160	0.0055	CRHR1	rs7209436;rs9892359
CC	0.474	43693749	43703383	113:157	0.007	SKAP1	rs2938471;rs2924254
GGTA-TAGC	0.14	40388751	40406918	53:84	0.0081		rs8073976;rs1055646; rs1007190;rs3024295; rs3024293;rs2981586; rs3024275;rs17547180
GT	0.451	46000610	46002010	112:154	0.01	CACNA1G	rs198545;rs12946808

**Table 4**  
**Follow-up Association Results**

Single SNP Transmission Disequilibrium tests for the marker within CACNA1G in a follow-up sample of 1046 affected individuals from 556 independent trios with an Empiric P-Value = 0.05 is shown. Marker identifier ('SNP'), genomic position of Chromosome 17 ('Position'), Transmitted versus Non-Transmitted count ('T:NT'), odds ratio, and empiric p-value corresponding to each associated marker are listed.

SNP	Position	T:NT	Odds Ratio	Empiric P-Value
rs198547	46006810	138:176	0.7841	0.02804

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript