No evidence for linkage of liability to autism to HOXA1 in a sample from the CPEA network
No Evidence for Linkage of Liability to Autism to HOXA1 in a Sample From the CPEA Network

Bernie Devlin,1 Pamela Bennett,2 Edwin H. Cook Jr.,3 Geraldine Dawson,4 David Gonen,3 Elena L. Grigorenko,5 William McMahon,2 David Pauls,6 Moyra Smith,7 M. Anne Spence,7 CPEA Genetics Network, and Gerard D. Schellenberg8,9*

1Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania
2Division of Child and Adolescent Psychiatry, Department of Psychiatry, University of Utah, Salt Lake City, Utah
3Child and Adolescent Psychiatry, Department of Psychiatry, Laboratory of Developmental Neuroscience, Department of Obstetrics, University of Chicago, Chicago, Illinois
4Department of Psychology, University of Washington, Seattle, Washington
5Child Study Center and Department of Psychology, Yale University School of Medicine, New Haven, Connecticut
6Unit of Psychiatric and Neurodevelopmental Genetics, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts
7Department of Pediatrics, University of California, Irvine, Irvine, California
8Veterans Affairs Medical Center, Seattle, Washington
9Departments of Medicine, Neurology and Pharmacology, University of Washington, Seattle, Washington

A recent study by Ingram et al. [2000b: Teratology 62:393–405] suggests a His73 Arg polymorphism (A:G) in HOXA1 contributes substantially to a liability for autism. Using 68 individuals diagnosed with Autism Spectrum Disorders, they found a significant dearth of G homozygotes and biased transmission of G alleles from parents to affected offspring, especially from mothers. Because the connection between HOXA1 and liability to autism is compelling, we attempted to replicate their finding using a larger, independent sample from the Collaborative Programs of Excellence in Autism (CPEA) network. In our data, genotype frequencies conform to Hardy-Weinberg equilibrium; allele transmissions meet Mendelian expectations; and there is no obvious sex-biased allele transmission. Based on our sample size, calculations suggest that we would have at least 95% power to detect linkage and association even if the A:G polymorphism were to account for only 1% of the heritability of autism. Therefore, although we cannot exclude the possibility that the samples in the two studies are intrinsically different, our data from our sample argue against a major role for HOXA1 His73Arg in liability to autism.

KEY WORDS: autism; HOXA1; Asperger syndrome; pervasive developmental disorder; genetic association; autistic disorder

INTRODUCTION

Autism is a developmental disorder characterized by deficits in social interaction and language, and by stereotypic behavior and restricted interests. Symptoms are typically first observed between the ages of 1 and 3 years. Autism is the most severe of several pervasive developmental disorders, or autism spectrum disorders (ASD), that include pervasive developmental disorder not otherwise specified (PDD-NOS), Asperger syndrome, and childhood disintegrative disorder [Lord et al., 2001]. The specific etiology of autism is unknown. A substantial body of evidence indicates, however, that inheritance contributes to individuals’ liability. The strongest evidence supporting a genetic hypothesis is that concordance rates in monozygotic twins are 60% for
autism with an increase to 95% when a broader phenotype definition of social and cognitive deficits related to autism is used [Bailey et al., 1995; Santangelo and Folstein, 1999]. Concordance rates in dizygotic twins and sibs are >3% and 4.5% [Bolton et al., 1994], respectively, and for more distant cousins, the risk falls to less than 1% [Delong and Dwyer, 1988; Jorde et al., 1990, 1991]. Still these estimates are much larger than the population prevalence for autism, which is now estimated to be approximately 0.17% [Chakrabarti and Fombonne, 2001]. The large difference between risk to monozygotic twins and more distant relatives suggest that multiple interacting genes are involved. Estimates of the number of such genes range from 2–10 [Pickles et al., 1995] to 15 genes or more [Risch et al., 1999; Pritchard, 2001].

Although the genes responsible for autism are unknown, genetic linkage studies of autism families [Hoh and Ott, 2000; Buxbaum et al., 2001; Liu et al., 2001; Shao et al., 2002] have identified multiple candidate regions for autism loci, including segments of chromosomes 1p, 2q, 7q, 16p, and 19q. Perhaps the most promising locations are chromosomes 2q and 7q. The pericentromeric region of 15q has also been implicated because a number of subjects with autism have a duplication of the Prader-Willi syndrome and Angelman syndrome region, either as a duplication or as a marker chromosome [Gillberg et al., 1991; Baker et al., 1994; Flejter et al., 1996; Cook et al., 1997]. Association studies have also suggested that genes in this region are involved in autism [Cook et al., 1997; Martin et al., 2000]. Candidate genes, selected because of postulated relevance based on the function of the encoded protein, have been examined as autism genes. These include the serotonin transporter [Cook et al., 1997], reelin [Persico et al., 2001], and WNT2 [Lupski, 1998].

HOXA1 and HOXB1 have also been tested recently as candidate genes for autism spectrum disorder [Ingram et al., 2000b]. These genes encode homeobox transcription factor proteins that regulate neuronal development, possibly during neural tube formation [Murphy and Hill, 1991]. The selection of HOXA1 and HOXB1 as autism candidate genes stemmed from postulated similarities between autism and developmental abnormalities induced by in utero exposure to thalidomide and valproate. Several investigators have suggested that thalidomide [Miller et al., 1998] and valproate [Christianson et al., 1994; Williams and Hersh, 1997; Williams et al., 2001] exposure in utero appear to increase the risk of autism. In all thalidomide cases with autism, ear malformations but not limb abnormalities are present [Miller et al., 1998], suggesting that the teratogen-induced damage associated with autism occurs between Day 20 and Day 24 after conception, the period of time when HOXA1 and HOXB1 are expressed. Other evidence supporting these genes as candidates is a single autopsy of a subject with autism and prominent impairment in motor function, in which loss of the facial nucleus and superior olive were noted, deficits that are also observed in animals treated with valproate [Rodier et al., 1996] and in HOXA1 null mice [Carpenter et al., 1993; Mark et al., 1993]. Also, in rats, pre-natal valproate treatment results in reduced numbers of Purkinje cells in the cerebellum [Ingram et al., 2000a], a feature consistently observed in autopsy series of autism subjects [Ritvo et al., 1986; Bailey et al., 1998; Kemper and Bauman, 1998].

To test HOXA1 and HOXB1 as candidate genes for autism, Ingram et al. [2000b] examined polymorphisms in each gene in a group of 68 subjects with ASD. For the $\text{His}^{73}\arg$ polymorphism (A:G) in HOXA1, they found a significant deviation from Hardy-Weinberg equilibrium resulting from an excess of heterozygotes, and biased transmission of alleles from parents to affected offspring. The distribution of HOXB1 genotypes showed no significant association with autism. To confirm the role of HOX genes in autism, replication studies are required using independent subject groups. In the present study, the HOXA1 $\text{His}^{73}\arg$ polymorphism was genotyped in a large sample of individuals with autism, PDD-NOS, and Asperger syndrome and their families. The results of the study presented here do not replicate the findings of the Ingram et al. [2000b].

MATERIALS AND METHODS

Simplex and multiplex autism families were recruited from ongoing research projects at five Collaborative Program of Excellence in Autism (CPEA) network sites: University of California Irvine, University of Chicago, University of Utah, University of Washington, and Yale University. Two hundred thirty one families were genotyped. Approximately 59% were simplex families and the rest were multiplex families. Of the multiplex families, a few were multigenerational; nonetheless, linkage/association information was almost always confined to a nuclear family. Diagnosis was based on the Autism Diagnostic Interview (ADI) [Lord et al., 1994], the Autism Diagnostic Observation Scale [Lord et al., 2000], DSM-IV, and clinical evaluation to rule out known medical causes of autism. Of the genotyped families, 221 had at least one affected progeny genotyped for the HOXA1 polymorphism. For most families, DNA was available from both parents (84%). The ancestry of most of the families is European (88%); for the next largest fraction (6%), parents had different ancestries.

HOXA1 genotypes were determined essentially as described by Ingram et al. [2000b]. Because genotyping was carried out at four different sites, a panel of 20 control samples obtained from the Coriell Cell Repository was genotyped at each site and the results compared to insure that genotyping methods across sites produced identical results. In addition, four samples genotyped by Ingram et al. [2000b] were obtained to insure that the genotypes determined here were consistent with the previous work.

For our analyses, two groups of affected individuals were considered. The first group included all individuals diagnosed with autism, PDD-NOS, or Asperger syndrome (broad diagnosis). This diagnostic spectrum was similar to that used by Ingram et al. [2000b]. The second group included only individuals with autism (narrow diagnosis). In addition, we grouped together those (few)
individuals who were genotyped but were lacking diagnostic information and those individuals who were genotyped and had sub-threshold autism symptoms (diagnosis unknown).

Markers and pedigrees were evaluated for Mendelian errors using the PedCheck program [O'Connell and Weeks, 1998]. Only two genotyping errors were noted, out of more than 1,000 genotypes, and these were set to missing. To test for differential transmission of the A vs. G alleles, a generalization [Rabinowitz and Laird, 2000] of the TDT test [Spielman et al., 1993], as implemented in the program FBAT (http://www.biostat.harvard.edu/~flat/) was used. By using a general framework for conditional inference, FBAT can analyze nuclear families with any combination of affected and unaffected individuals, including unknown phenotypes. The program also infers missing parental genotypes when the missing genotypes can be unambiguously inferred from the observed genotypes in the family [Rabinowitz and Laird, 2000]. As a default analysis, we chose an additive model, which is often powerful even when the true model deviates from additivity, and we also evaluated a recessive model. The data were analyzed for both broad and narrow diagnosis. All analyses ignored the possibility of linkage in the region surrounding HOXA1 (tests would be slightly anticonservative if this assumption were false).

To evaluate the power of the sample to detect linkage and association, we first characterized the structure of the family sample for the broad diagnosis, namely the number of genotyped parents (0, 1, or 2) and the number of affected and unaffected progeny who were genotyped. Using this family structure, we calculated power using the program PowerFBAT (http://www.biostat.harvard.edu/~flat/), which also requires specification of the frequency of the liability allele, the penetrances of two genotypes (i.e., probability of being affected given an individual is carrying 0 [f0] and 2 [f2] liability alleles), and a parameter (γ) that specifies the penetrance, and therefore the genetic model, for the heterozygote. For our calculations, power was evaluated for the additive, dominant, multiplicative and recessive models, and assumed the liability allele accounted for 1% of the variability of autism. See Risch [1990] for the formula for the additive model; parameters for the other genetic models follow from some algebra. As a critical value, we took γ = 0.05 and evaluated 1,000 simulations, which generated data using the additive model (f0 = 0.00076; f2 = 0.02847; γ = 0.5; autism prevalence = 0.005) as the true model for HOXA1 liability.

RESULTS

We tested the hypothesis that HOXA1 alleles are involved in liability to autism [Ingram et al., 2000b]. Our sample consisted of 231 simplex and multiplex families with 318 offspring who were diagnosed as affected and who were genotyped. Most of the affected subjects had a diagnosis of autism (N = 250; 78.6%), with a smaller fraction having a diagnosis of Asperger disorder (N = 15; 4.7%) or PDD-NOS (N = 53; 16.7%). Of the offspring who were genotyped (N = 587), many were diagnosed as unaffected (N = 216; 36.8%) and a smaller number were either not diagnosed or had sub-threshold autistic symptoms; these were categorized as 'diagnosis unknown' (N = 53; 9.0%).

In our sample of pedigree founders, mostly parents of children with autism, the A’ allele at HOXA1 occurred with frequency 84.7% (718/848) and the G’ allele occurred with frequency 15.3% (130/848). Founder genotypes conformed to Hardy-Weinberg expectations: 302 are AA, 114 are AG, and 8 are GG, compared to 304.2, 109.9, and 9.96 expected (χ² = 0.56; df = 1; P = 0.46). The genotype frequencies of the founders were not significantly different from Ingram et al. [2000b] convenience sample of unaffected individuals from the population (Table I; χ² = 3.84; df = 2; P = 0.15). Note that three founders from our sample were diagnosed as affected, and their genotypes are AA, AA, and AG.

For the broad diagnostic category, analogous to that used by Ingram et al. [2000b] 95 out of 221 families were informative for allele transmissions. No significant deviation from the null hypothesis of 50:50 A:G transmissions was seen in our data (Z = 0.637; P = 0.524). Similar results were obtained for the narrow diagnosis of autism (Z = 1.16; P = 0.244). In this case, 80 out of 221 families were informative. For completeness, we also examined transmissions under the assumption of a recessive model of inheritance. Again, no significant differential transmissions occurred, either for the broad (Z = 1.13; P = 0.26) or narrow (Z = 1.77; P = 0.08) diagnosis.

Because Ingram et al. [2000b] evaluated transmissions from mothers and fathers to their male and female offspring, we did likewise (Table II). Two kinds of matings must be considered, a single heterozygous parent (Table II) vs. both parents heterozygous (Table II). Although all transmissions are informative for the former, transmissions of alleles to heterozygous children cannot be assigned unambiguously to their specific, heterozygous parents. Therefore, for the overall transmissions (Table II), we included only the unambiguous transmissions. Although Ingram et al. [2000b] found differential transmissions of the G allele, especially from mothers to their affected offspring and even to their unaffected female offspring, we found no excess transmission of the G allele (Table II). To the contrary, the A allele was transmitted more commonly, although not significantly. Seventy-six times, A was transmitted, whereas G was transmitted 63 times (χ² = 2.12; df = 1; P = 0.54). Mendelian expectations were realized, stochastically, for all parent/child combinations in Table II, regardless of affection status.

---

**TABLE I. Genotype Frequencies From Pedigree Founders in Our Sample Versus Those From a Convenience Sample of Non-Affected Individuals**

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Founder sample</th>
<th>Convenience sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>302 (308.4)</td>
<td>93 (86.6)</td>
</tr>
<tr>
<td>AG</td>
<td>114 (109.3)</td>
<td>26 (30.7)</td>
</tr>
<tr>
<td>GG</td>
<td>8 (6.2)</td>
<td>0 (1.8)</td>
</tr>
</tbody>
</table>

*Convenience sample from Ingram et al. [2000b]. Parenthetical numbers are expected values assuming the samples are drawn from the same population.*
It is important to ask if the sample used here is comparable to that of Ingram et al. [2000b]. In terms of presentation of symptoms, the two samples are quite similar. Of the 57 probands in the Ingram sample, 40 (70.2%) were diagnosed with autism, with the remainder having a diagnosis of Asperger syndrome or PDD-NOS; in our sample of 318 children who were diagnosed as affected and who were genotyped, 78.6% have autism. The gender ratios (male:female) of the genotyped affected are similar for both samples (3.2:1 in the Ingram et al. [2000b] sample vs. 4.2:1 for our sample). The sex ratio for unaffected children is somewhat divergent, 1:3 vs. 1.3:1, respectively, although the Ingram et al. [2000b] sample is small: 4 men and 12 women. Conditional on being diagnosed as unaffected, the expectation is approximately 1:1, so our sample matches expectation quite well. The sex ratio for the unknown category (Table II) lies between the ratios for affected and unaffected offspring, 2.4:1, which fits our expectation because many of these children have autistic features that are not sufficiently severe to generate a diagnosis.

The sampling strategy of Ingram et al. [2000b] differed somewhat from the sampling strategy here. They recruited strictly multiplex families, concentrating on second and third degree relative pairs, whereas a portion of our sample was multiplex and most of these families contained affected sibling pairs. When we analyzed only the multiplex pedigrees in our sample, we again found no evidence for linkage and association (additive: $P = 0.85$; recessive: $P = 0.50$). Likewise, the Ingram et al. [2000b] families were strictly of European origin. When we restrict the analyses to only those families in which both parents are of European ancestry, we again find no significant association (additive: $P = 0.56$; recessive: $P = 0.33$).

Power analysis was carried out to determine whether our sample was large enough to replicate the findings of Ingram et al. [2000b] To evaluate power, we assume the G allele occurs with probability 0.153 in the general population (the observed value), and use the family structure for the sample (see Materials and Methods).

We further assume HOXA1 accounts for 1% of the variability of the autism-spectrum phenotype [e.g., Risch 70 (Risch, 1990) equals $\approx 2.0$] and set the critical $P$-value to 0.05. For our sample, results from Power-PBAT show $>95\%$ power to detect an association of HOXA1 alleles for additive, dominant, multiplicative and recessive models (these results were corroborated by using approximations for family structure and formulas in Knapp [1999]). Such calculations, of course, assume the sample is representative of the population, and that genotyping and diagnostic errors are absent.

**DISCUSSION**

Ingram et al. [2000b] put forth a compelling case for HOXA1 defects contributing to liability to autism. Some of the most convincing circumstantial evidences are that HOXA1 plays a role in neurodevelopmental processes and is expressed during early embryonic development. Complementing these observations, their data suggest biased transmission of G vs. A allele, from heterozygous parents to affected offspring. This single nucleotide difference produces a histidine to arginine substitution in the finished protein.

Unfortunately, the results presented here fail to lend any support for a direct connection between the $^{\text{His73Arg}}$ polymorphism and liability for autism. Some of the most convincing circumstantial evidences are that HOXA1 plays a role in neurodevelopmental processes and is expressed during early embryonic development. Complementing these observations, their data suggest biased transmission of G vs. A allele, from heterozygous parents to affected offspring. This single nucleotide difference produces a histidine to arginine substitution in the finished protein.

Unfortunately, the results presented here fail to lend any support for a direct connection between the $^{\text{His73Arg}}$ polymorphism and liability for autism. Contrary to the results of Ingram et al. [2000b], we find no dearth of G homozygotes in our sample and transmission of the A and G alleles from heterozygous parents to offspring meet Mendelian expectations. Unlike their results, we find no gender-specific biased transmission of alleles to either affected or unaffected offspring. The greatest similarity between our results and theirs is transmission of alleles from female parents to affected offspring. This single nucleotide difference produces a histidine to arginine substitution in the finished protein.

Unfortunately, the results presented here fail to lend any support for a direct connection between the $^{\text{His73Arg}}$ polymorphism and liability for autism. Contrary to the results of Ingram et al. [2000b], we find no dearth of G homozygotes in our sample and transmission of the A and G alleles from heterozygous parents to offspring meet Mendelian expectations. Unlike their results, we find no gender-specific biased transmission of alleles to either affected or unaffected offspring. The greatest similarity between our results and theirs is transmission of alleles from female parents to affected offspring, which shows a strong bias in their sample and a weaker, non-significant bias in ours. Even in this case, however, the results diverge because the bias is for different alleles.

There are several plausible explanations for these disparate results. The simplest one is that the results of Ingram et al. [2000b] reflect a false positive. In the
search for the genetic determinants of complex disease, such an explanation is unremarkable. Alternatively, we may have failed to detect the effect of this \textit{HOXA1} polymorphism. Although this possibility cannot be ruled out, it seems less plausible because our sample is substantially larger than the original study and has ample power to detect the effect size suggested by the Ingram et al. (2000b) results. Thus, all other things being equal, our failure to replicate the previous study is telling. Not only are our results incompatible with a strong effect of this \textit{HOXA1} polymorphism on autism liability, the patterns in the data are opposite those of the original study, which saw an excess of \textit{G} transmissions compared to our stochastically greater transmission of \textit{A} alleles.

The important caveat in the previous passage is “all other things being equal.” A third explanation for the disparate findings is that the different samples reflect different study populations, a possibility that is difficult to rule out. The samples look similar in terms of their composition of autistic and autistic spectrum individuals, male/female ratios, and ethnicity. The only clear distinction is the sampling framework, in that the entire sample of Ingram et al. (2000b) was drawn from multiplex families whereas only 41% of our sample was drawn from known multiplex families, mostly affected sibling pairs. Nonetheless, analysis of our multiplex families lent no support for \textit{HOXA1} as a liability locus for autism. Because of the clinical heterogeneity of this disorder, it is not possible to exclude either systematic or random variation in sampling from the syndrome. Moreover, recent modeling of complex susceptibility loci makes it clear that the contribution of each locus in a multiplicative disorder (as suggested for autism, given the rapid increase in relative risk from siblings to MZ twins) includes relatively weak loci, which some but not all samples may detect. Thus, we cannot exclude the possibility that the \textit{His73Arg} polymorphism or some other polymorphic site within \textit{HOXA1} could be a minor contributor to autism susceptibility, but that the effect size is too small to observe with our sample. What is clear, however, is that the sample evaluated in this study does not provide any support for \textit{HOXA1 His73Arg} as a substantial contributor to autism liability.

\section*{ACKNOWLEDGMENTS}

This research was funded by a program project grant from the National Institute of Child Health and Human Development and the National Institute on Deafness and Communication Disability (P01HD34565), which is part of the NICHD/NIDCD Collaborative Program of Excellence in Autism. B.D. was supported by MH57881 from the National Institute of Mental Health. Work at the University of Utah was supported by PO1 HD35476 to W.M.N., and Public Health Service research grant M01-RR00064 from the National Center for Research and Resources and the Devonshire Foundation. We thank Christopher Stodgell for supplying 4 DNA samples for genotype comparisons, and S. Guter, Z.-Y. Yang, E. Loomis, and C. Romano for excellent technical support.

\section*{REFERENCES}

After this manuscript was submitted, Li et al. (2002) described an association study of the alleles of the \textit{HOXA1} \textit{His73Arg} polymorphism and autism using 110 multiple autism families. No positive evidence of an association of \textit{HOXA1} and autism was observed, a finding that is in agreement with the work described above.


\section*{NOTE IN ADDED PROOF}


672 Devlin et al.


