No association between the EN2 gene and autistic disorder

H Zhong, F J Serajee, R Nabi, A H M Mahbubul Huq

AUTISTIC SPECTRUM DISORDERS (MIM 209850), WHICH INCLUDE AUTISM, ASPERGER SYNDROME, AND PERVERSIVE DEVELOPMENTAL DISORDER (PDD) NOT OTHERWISE SPECIFIED, ARE CHARACTERISED BY IMPAIRMENT IN COMMUNICATIONS AND SOCIAL INTERACTIONS AND THE PRESENCE OF STEREOTYPED BEHAVIOURS. THE AETIOLOGY OF AUTISTIC DISORDER IS UNKNOWN, BUT FAMILY AND TWINS STUDIES HAVE SHOWN A HIGH MONOZYGOTIC TO DIZYGOTIC TWIN RISK RATIO AND A SIB RELATIVE RISK BETWEEN 50 AND 100, SUGGESTING THAT INHERITANCE OF AUTISM IS COMPLEX, BUT THE PREDISPOSITION TO DEVELOP IT IS LARGELY GENETICALLY DETERMINED.

A NUMBER OF MORPHOLOGICAL ABNORMALITIES INCLUDING INCREASED BRAIN SIZE AND DEVELOPMENTAL ABNORMALITIES OF THE CEREBRAL CORTEX, BRAINSTEM, AND CEREBELLM HAVE BEEN REPORTED IN AUTISM. REPORTED BRAIN STEM AND CEREBELLAR ANOMALIES INCLUDE HYPOPLASIA OF LOBULES VI AND VII, DECREASED PURKINJE CELL DENSITY, OLIVARY HYPOPLASIA, AND NEURONAL HETEROPTOPIAS. THE MORPHOLOGICAL ABNORMALITIES DESCRIBED IN THE BRAINSTEM AND CEREBELLUM OF AUTISTIC SUBJECTS SUGGEST THAT GENES INVOLVED IN CEREBELLAR DEVELOPMENT ARE CANDIDATE GENES IN AUTISM.

EN2, A HUMAN HOMOLOGUE OF DROSOPHILA ENGRAILED GENE, IS A HOMEBOX GENE WITH AN ESSENTIAL ROLE IN THE DEVELOPMENT OF THE MIDBRAIN AND CEREBELLUM. MICE HOMOZYGOUS FOR A TARGETED DELETION OF THE EN2 HOMEBOX REGION WERE VIALBE BUT SHOWED ABNORMAL FOLLATION OF THE CEREBELLUM. PETIT ET AL. STUDIED TWO RESTRICTION FRAGMENT LENGTH POLYMORPHISMS IN THE EN2 GENE IN AUTISM AND FOUND AN ASSOCIATION BETWEEN AUTISM AND A PVUII POLYMORPHISM IN THE 5TH REGION OF EN2 IN 100 AUTISTIC CHILDREN AND 100 CONTROL CHILDREN.

IN THIS STUDY, WE HAVE ATTEMPTED TO REPLICATE THE ASSOCIATION BETWEEN THE EN2 GENE AND AUTISM USING FAMILY BASED LINKAGE AND ASSOCIATION STUDIES IN 196 MULTIPLEX FAMILIES WITH AUTISM. THE EN2 GENE MAPS TO HUMAN CHROMOSOME 7q36. INFORMATION REGARDING A NUMBER OF PROMOTER, EXON, AND INTRON POLYMORPHISMS IN THE EN2 GENE IS AVAILABLE IN JSNP, A DATABASE OF COMMON GENE VARIATIONS IN THE JAPANESE POPULATION. IN ADDITION, A NUMBER OF UNCONFIRMED SNPS IN THE EN2 GENE ARE LISTED IN dBSNP. WE INVESTIGATED A G/A SNP (JSNP ID IMS-JST089559) IN EXON 1 OF THE EN2 GENE, AS THIS SNP IS LOCATED NEAR THE PVUII POLYMORPHISM STUDIED BY PETIT ET AL., ALTERS THE AMINO ACID COMPOSITION OF THE EN2 PROTEIN, AND IS A COMMON POLYMORPHISM WITH ALLELE FREQUENCIES SUITABLE FOR GENETIC ASSOCIATION STUDIES.

SUBJECTS AND METHODS

FAMILY ASCERTAINMENT AND DIAGNOSTIC CRITERIA

DNA SAMPLES FROM 204 FAMILIES WERE OBTAINED FROM THE AUTISM GENETIC RESOURCE EXCHANGE (AGRE). AGRE, WHICH WAS CREATED BY THE CURE AUTISM NOW FOUNDATION AND THE HUMAN BIOLOGICAL DATA INTERCHANGE, IS A CENTRAL REPOSITORY OF FAMILY DNA SAMPLES FOR GENETIC STUDIES OF AUTISM. DIAGNOSES OF AGRE FAMILIES WITH AUTISM, ASPERGER SYNDROME, AND PDD WERE CONFIRMED USING THE AUTISM DIAGNOSTIC INTERVIEW--REVISED PROTOCOL (ADI-R). AGRE FAMILIES INCLUDED AT LEAST TWO AFFECTED MEMBERS WITH A DIAGNOSIS OF AUTISM, ASPERGER SYNDROME, OR PDD. CHILDREN WITH FRAGILE X SYNDROME, TUBEROUS SCLEROSIS, AND CHROMOSOMAL ANOMALIES WERE EXCLUDED. PHENOTYPIC INFORMATION ON THE PARENTS WAS NOT AVAILABLE. HOWEVER, ADI-R PHENOTYPIC DATA INCLUDING ADI-R ITEMS AND ALGORITHM SCORES ARE AVAILABLE FOR AFFECTED SIBS FROM THE AGRE WEB SITE (WWW.AGRE.ORG). FOR QUANTITATIVE TRAIT STUDIES, WE USED FOUR ADI-R ITEMS THAT HAVE SHOWN INCREASED PHENOTYPIC CORRELATIONS BETWEEN SIBS AS QUANTITATIVE TRAITS FOR LINKAGE AND ASSOCIATION STUDIES.

EN2 GENOTYPING

WE GENOTYPED A G/A SNP (JSNP ID IMS-JST089559; JSNP: A DATABASE OF COMMON GENE VARIATIONS IN THE JAPANESE POPULATION) IN EXON 1 OF THE EN2 GENE BY A RESTRICTION ENZYME BASED ASSAY USING AluI (NEW ENGLAND BIORB INC, BEVERLY, MA). THE SNP ALTERS THE AMINO ACID COMPOSITION OF EXON 1 OF THE EN2 GENE, CHANGING THE CODON 121 FROM LEUCINE (CTC) TO PHENYLALANINE (TTC). WE AMPLIFIED THE TARGET REGION WITH THE FOLLOWING PRIMERS: AGCGTTCCGAGTCCGAGC (FORWARD) AND CAAGCCTGGC GAAGCAG (REVERSE) WITH A PCR PROGRAMME OF 95°C FOR 5 MINUTES, 37 PCR CYCLES OF 95°C FOR 30 SECONDS, 58°C FOR 45 SECONDS, AND 72°C FOR 30 SECONDS. FOR THE PRESENT STUDY, A TOTAL OF 890 SUBJECTS (670 CHILDREN) WERE AVAILABLE.

STATISTICAL AND GENETIC ANALYSIS

THE DATA WERE PREPARED USING MEGA2 AND ASPEX. MARKER ALLELE FREQUENCIES WERE OBTAINED BY COUNTING PARENTAL GENOTYPES. MAP DISTANCES WERE OBTAINED FROM MARSHFIELD AND GENETHON. LINKAGE ANALYSIS WAS PERFORMED BY GENEHUNTER 2.1, ASPEX, AND MERLIN SOFTWARE PACKAGES. NONPARAMETRIC QTL ANALYSES, HASSEMAN-ELSTON (HE) REGRESSIONS, AND PARAMETRIC QTL ANALYSES, HASSEMAN-ELSTON (HE) REGRESSIONS.
RESULTS
We had a total of 226 affected sib pairs and 207 discordant sib pairs; 181 families had two affected sibs and 15 families had three affected sibs. Eight families with identical twins were excluded from the analysis. However, these families were genotyped and served to check for genotyping errors. The program Pedcheck and MERLIN were used to find Mendelian errors.14 15 The detectable genotype errors in the sample were less than <0.1%. There was no deviation from Hardy-Weinberg disequilibrium.

We analysed sharing at the EN2 locus with the ASPEX software package. There was no overall increased identity by descent sharing in affected sib pairs at the EN2 locus, although there was increased sharing of the paternal allele in affected sib pairs compared to discordant sib pairs in a subset of families where all affected sibs have developed useful phrase speech. Sharing data for maternal and paternal alleles in affected and discordant sib pairs are shown in table 1.

We performed non-parametric sib pair analyses to test for linkage of autism to the EN2 gene using the software package GENEHUNTER. Two point linkage analysis using all 196 families yielded a NPL score of 0.22 at the EN2 locus. The AGRE families have been used in a genome screen for autism or subgroups with different overall levels of language and variance component analyses were computed for the quantitative sib pair data. Children from the autistic families were classified as affected or unknown and were tested for association through the use of a transmission disequilibrium test (TDT) within the GENEHUNTER package version 2.1. Association of the quantitative traits was evaluated by the QTDT (Quantitative TDT), which uses a variance component model that partitions association into between and within family components.15

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Sharing at EN2 locus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affected sib pairs</td>
<td></td>
</tr>
<tr>
<td>No of sib pairs</td>
<td>Sharing at paternal allele (%)</td>
</tr>
<tr>
<td>All families</td>
<td>225</td>
</tr>
<tr>
<td>Useful phrase speech</td>
<td>96</td>
</tr>
<tr>
<td>No useful phrase speech</td>
<td>124</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Transmission disequilibrium test (TDT) for EN2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmitted</td>
<td>Not transmitted</td>
</tr>
<tr>
<td>All families</td>
<td>206</td>
</tr>
<tr>
<td>Strict autism families</td>
<td>205</td>
</tr>
<tr>
<td>Autism-PDD families</td>
<td>67</td>
</tr>
<tr>
<td>Useful phrase speech</td>
<td>107</td>
</tr>
<tr>
<td>No useful phrase speech</td>
<td>99</td>
</tr>
</tbody>
</table>

that, on a daily basis, involves phrases of three words or more, at least sometimes includes a verb, and is comprehensible to other people). The NPL score of 1.58 at the EN2 locus. Analysis of families, in which at least one sib had no speech or no functional use of three word phrases, resulted in a NPL score of 0.33 at the EN2 locus. We also stratified the families according to whether all affected sibs met the strict criteria for autism. There was no increased evidence of linkage or association when analyses were performed separately with strict autism sib pairs or broader phenotype sib pairs.

For QTL analysis, we performed non-parametric multipoint linkage analyses in 196 families focusing on ADI-R items for expressive language level and a composite algorithm score for stereotypical behaviour. We studied four ADI-R items that have shown increased phenotypic correlations between sibs as quantitative traits for linkage studies.16 17 ADI-R items A12 (age at first word), A13 (age at phrase speech), and A19 (useful phrase speech level) were used as language traits. ADI-R composite algorithm score D Total was used as a stereotyped behaviour trait. Age at first word refers to words repeated and consistently used for the purpose of communication with reference to a particular concept. For age at first phrase, the phrase must consist of two words, one of which must be a verb. Subjects without any word or phrase speech by interview time were coded by their current age if <8 years or 96 months if >8 years. Non-parametric QTL and Haseman Elston regression studies with GENEHUNTER and variance component analyses with MERLIN did not show any evidence of linkage with p values less than 0.05.

TDT analysis of the exon 1 variant in 196 autistic disorders families did not show any evidence of association (p=0.58). TDT using the subgroup of families with a strict diagnosis of autism or subgroups with different overall levels of language did not increase evidence of association (table 2). We also investigated the ADI-R items A12, A13, and A19 as language traits and the ADI-R algorithm score D Total for stereotyped behaviour using the QTDT package to test association between the EN2 gene and quantitative traits. We performed 5000 permutations for each of the above traits and did not find evidence of association of EN2 with any of the quantitative traits with p values <0.1.

DISCUSSION
EN2 is a potential autistic disorder susceptibility gene. The study by Petit et al16 used a case-control approach and found an association between a PvuII polymorphism in EN2 and autism.
The PrxII polymorphism in their study was typed using the MP-4 probe, which is located within the 5’ region of the EN2 gene. They did not find any association with an SNP polymorphism that was typed using the MP5 probe, which is located within the homeobox region and extends to the end of the 3’ region. We have studied a G/A SNP (position 89582, Accession AC008060) in exon 1 of the human EN2 gene. The genomic sequence of EN2 is now available in the GenBank database (Homo sapiens PAC clone RP5-912113 from chromosome 7, Accession AC008060). Review of the genomic sequence shows PrxII sites at nucleotide positions 84338, 85789, 88661, and 87107 (Accession 008060) in the EN2 promoter, such that with the MP4 probe used by Petit et al., a PrxII polymorphism at position 86681 would result in a 1.45 bp invariant band and a two allele RFLP with a 1318 bp or a 1072 bp band. This suggests that the PrxII RFLP studied by Petit et al. (position 86681, Accession 008060) is located 2.9 kb 5’ to the G/A SNP in exon 1. The allele frequency of the exon 1 SNP (G: 0.514, A: 0.486) is similar to the allele frequency of the PrxII polymorphism in the original study. The ethnic distribution among the original and replication data appears to be similar. In the original study, subjects who met all the criteria of early infantile autism as defined in DSM III-R were included, while we included subjects with diagnosis of autism, Asperger syndrome, and PDD. However, when we analysed the subset of our sample that met strict criteria for autism in all affected sibs, we still did not find any evidence of association.

There are several possible explanations for the differences in the findings in the two studies. First, there are differences in the diagnostic definition between the two studies. In our study, diagnosis of families with autism, Asperger syndrome, and PDD were confirmed using the ADI-R, while in the original study clinical evaluation was completed using the Behavior Rating Scale. Second, there are differences in the extent of linkage disequilibrium and the susceptibility mutation in the EN2 gene. They did not find any association with an EN2 polymorphism in their study was typed using the MP5 probe, which is located 2.9 kb 5’ to the G/A SNP in exon 1. The allele frequency of the exon 1 SNP (G: 0.514, A: 0.486) is similar to the allele frequency of the PrxII polymorphism in the original study. The ethnic distribution among the original and replication data appears to be similar. In the original study, subjects who met all the criteria of early infantile autism as defined in DSM III-R were included, while we included subjects with diagnosis of autism, Asperger syndrome, and PDD. However, when we analysed the subset of our sample that met strict criteria for autism in all affected sibs, we still did not find any evidence of association.

In summary, we have failed to replicate the report by Petit et al. suggesting that the EN2 gene on 7q36 is a genetic risk factor for autistic disorder. EN2 is a small gene within two exons contained within an 8 kb genomic region. This suggests that if a susceptibility mutation exists at EN2, our study should have detected an association between the EN2 exon 1 polymorphism and the susceptibility mutation in the EN2 gene. It should be noted, however, that the extent of linkage disequilibrium can be irregular in a number of chromosomal regions. The negative finding may be the result of heterogeneity among our autistic disorder families. We stratified the families according to the severity or overall level of language to identify more homogeneous subgroups, but it may be useful to subtype the autistic disorder families using other criteria. In addition, a larger sample may be needed to detect a susceptibility locus with small effect.

The EN2 gene maps to human chromosome 7q36. Although most studies have implicated the chromosome 7q21-34 region in autism,3,4 a recent study reported a peak with a lod score of 2.13 at marker D7S483 on 7q36.5 There is also evidence of a quantitative trait locus (QTL) for language on 7q35-36.17 Although our study did not provide support for a role of the EN2 gene in autistic disorder, further studies are required to investigate the presence of an additional susceptibility locus for autistic disorders further distally on the chromosome 7q36 region where EN2 is located.

ACKNOWLEDGEMENTS

The study described here was supported in part by a grant from the Children’s Hospital of Michigan Research Foundation. We acknowledge support from Cure Autism Now and the Autism Genetic Resource Exchange ( AGRE). We gratefully acknowledge the resources provided by the AGRE consortium* and the participating AGRE families.

*The AGRE Consortium. AGRE. Scientific Advisory Board: Chairman: Daniel H Geschwind, UCLA, Los Angeles, CA; Members: W Ted Brown, NYS Institute for Basic Research in Developmental Disabilities, Long Island, NY; Maya Bucan, University of Pennsylvania, Philadelphia, PA; Joseph Buxbaum, Mount Sinai School of Medicine, NY; Edwin H Cook, University of Chicago, Chicago, IL; T Conrad Gilliam, Columbia Genome Center, NY; David A Greenberg, Mount Sinai Medical Center, NY; David H Ledbetter, University of Chicago, Chicago, IL; Bruce Miller, UCSF, San Francisco, CA; Stanley F Nelson, UCLA School of Medicine, Los Angeles, CA; Jonathan Pevsner, Kennedy Krieger Institute, Baltimore, MD; Jerome I Rotter, Cedars-Sinai Medical Center, Los Angeles, CA; Gerard D Schellenberg, University of Washington, Seattle, WA; Carol A Sproose, Children’s National Medical Center, Baltimore, MD; Rudolph E Tanzi, Massachusetts General Hospital, Boston, MA; Kirk C Wilfond, University of California, San Francisco, CA; and Jeremy M Silverman, Mount Sinai Medical School, New York, NY.

Authors’ affiliations

H Zhong, F J Serenjee, R Nobi, A H M Mahabubul Huq, Department of Pediatrics, Wayne State University, Detroit, MI 48201, USA

A H M Mahabubul Huq, Department of Neurology, Wayne State University, Detroit, MI 48201, USA

Correspondence to: Dr A H M Mahabubul Huq, Division of Neurology, Children’s Hospital of Michigan, Detroit, MI 48201, USA; ahuq@med.wayne.edu

REFERENCES