ELECTRONIC LETTER

No association between the *EN2* gene and autistic disorder

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

J Med Genet 2003;**40**:e4(http://www.jmedgenet.com/cgi/content/full/40/1/e4)

Autistic spectrum disorders (MIM 209850), which include autism, Asperger syndrome, and pervasive 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 twin 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.[1 2]

A number of morphological abnormalities including increased brain size and developmental abnormalities of the cerebral cortex, brainstem, and cerebellum have been reported in autism.[3 4] Reported brain stem and cerebellar anomalies include hypoplasia of lobules VI and VII, decreased Purkinje cell density, olivary dysplasia, and neuronal heterotopias.[3–5] 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 homeobox gene with an essential role in the development of the midbrain and cerebellum.[6] Mice homozygous for a targeted deletion of the *EN2* homeobox region were viable but showed abnormal foliation of the cerebellum.[7] Petit *et al*[8] studied two restriction fragment length polymorphisms in the *EN2* gene in autism and found an association between autism and a *Pvu*II polymorphism in the 5′ 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.[9] 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 *Pvu*II polymorphism studied by Petit *et al*,[8] alters the amino acid composition of the EN2 protein, and is a common polymorphism with allele frequencies suitable for genetic association studies.

Key points

- Autism is a neurodevelopmental disorder with a complex genetic aetiology. A previous case-control study has detected an association between a marker in the homeobox gene *EN2* and autism using 100 autistic and 100 control children (*J Med Genet* 1995;**32**:269-74). We attempted to replicate this finding using family based linkage and association studies.
- We investigated a non-synonymous single nucleotide polymorphism in exon 1 of the *EN2* gene for evidence of association and linkage to autism using 196 multiplex autistic disorder families. Affected sib pair studies showed a maximal multipoint NPL (GENEHUNTER) score of 1.31 at 7q36 at the *EN2* locus.
- However, transmission disequilibrium test did not show any association between the exon 1 variant and autistic disorder. There was also no linkage or association between language and stereotypic behaviour quantitative traits and the exon 1 variant.
- In conclusion, our studies do not support a role of the *EN2* gene in autism.

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.[10] Diagnoses of AGRE families with autism, Asperger syndrome, and PDD were confirmed using the Autism Diagnostic Interview-Revised protocol (ADI-R). All 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 *Alu*I (New England Biolab Inc, Beverly, MA).[9] 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: AGCTGTCCGAGTCCGAGC (forward) and CAAGCCTGGC GAAGCAG (reverse) with a PCR programme of 95°C for five 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.[11 12] 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.[12–14] Non-parametric QTL analyses, Haseman-Elston (HE) regressions,

Table 1 Sharing at *EN2* locus

	Affected sib pairs			Discordant sib pairs		
	No of sib pairs	Sharing at paternal allele (%)	Sharing at maternal allele (%)	No of sib pairs	Sharing at paternal allele (%)	Sharing at maternal allele (%)
All families	225	51.6	47.9	210	47.3	50.9
Useful phrase speech	96	61.9	52.6	76	40	50
No useful phrase speech	124	46.5	42.9	131	51.4	51.6

Table 2 Transmission disequilibrium test (TDT) for *EN2*

	Transmitted	Not transmitted	χ^2	p value
All families	206	195	0.30	0.58
Strict autism families	205	186	0.92	0.34
Autism-PDD families	67	85	2.13	0.14
Useful phrase speech	107	97	0.49	0.48
No useful phrase speech	99	98	0.01	0.94

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]

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 16] 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[17] and genotype data from 200 families at an average 10 cM resolution are available online at the AGRE web site (www.agre.org).[10] ADI-R phenotypic data including ADI-R items and algorithm scores are also available from the AGRE web site. For multipoint study, ADI-R phenotypic data from the AGRE families, as well as the genotypic data of 18 additional markers from chromosome 7, were merged with the *EN2* genotype data. Analysis of 192 multiplex families resulted in a multipoint NPL scores of 1.31 at the *EN2* locus (p=0.06). When the families were stratified according to the overall language level, analysis of families with affected sibs who are coded 0 on ADI-R item A19 (that is, who have functional use of spontaneous, echoed, or stereotyped language that, on a daily basis, involves phrases of three words or more, at least sometimes includes a verb, and is comprehensible to other people) resulted in a 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 stereotyped behaviour. We studied four ADI-R items that have shown increased phenotypic correlations between sibs as quantitative traits for linkage studies.[18 19] 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 al*[8] used a case-control approach and found an association between a *Pvu*II polymorphism in *EN2* and autism.

The *Pvu*II 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 *Sst*I 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 RP5-912I13 from chromosome 7, Accession AC008060). Review of the genomic sequence shows *Pvu*II sites at nucleotide positions 84338, 85789, 86861, and 87107 (Accession 008060) in the *EN2* promoter, such that with the MP4 probe used by Petit *et al*,[8] a *Pvu*II polymorphism at position 86681 would result in a 1451 bp invariant band and a two allele RFLP with a 1318 bp or a 1072 bp band. This suggests that the *Pvu*II RFLP studied by Petit *et al*[8] (position 86861, 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 *Pvu*II 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 Summarized Evaluation: BSE scale.[20] More importantly, we carried out family based linkage and association studies including TDT, while Petit *et al*[8] used a case-control approach. Case-control studies may be subject to population stratification and spurious association.[21]

In summary, we have failed to replicate the report by Petit *et al*[8] 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.[22] 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 the aetiology of autism,[17 23–27] one study reported a novel peak with a lod score of 2.13 at marker D7S483 on 7q36.[17] There is also evidence of a quantitative trait locus (QTL) for language on 7q35-36.[18] 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 Sprouse, Children's National Medical Center, Baltimore, MD; Rudolph E Tanzi, Massachusetts General Hospital, Boston, MA; Kirk C Wilhelmsen, University of California, San Francisco, CA; and Jeremy M Silverman, Mount Sinai Medical School, New York, NY.

Authors' affiliations
H Zhong, F J Serajee, R Nabi, A H M Mahbubul Huq, Department of Pediatrics, Wayne State University, Detroit, MI 48201, USA
A H M Mahbubul Huq, Department of Neurology, Wayne State University, Detroit, MI 48201, USA

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

REFERENCES

1 **Folstein S**, Rutter M. Infantile autism: a genetic study of 21 twin pairs. *J Child Psychol Psychiatry* 1977;**18**:297-321.
2 **Bailey A**, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E, Rutter M. Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med* 1995;**25**:63-77.
3 **Bailey A**, Luthert P, Dean A, Harding B, Janota I, Montgomery M, Rutter M, Lantos P. A clinicopathological study of autism. *Brain* 1998;**121**(Pt 5):889-905.
4 **Kemper TL**, Bauman M. Neuropathology of infantile autism. *J Neuropathol Exp Neurol* 1998;**57**:645-52.
5 **Courchesne E**. Brainstem, cerebellar and limbic neuroanatomical abnormalities in autism. *Curr Opin Neurobiol* 1997;**7**:269-78.
6 **Liu A**, Joyner AL. EN and GBX2 play essential roles downstream of FGF8 in patterning the mouse mid/hindbrain region. *Development* 2001;**128**:181-91.
7 **Joyner AL**, Herrup K, Auerbach BA, Davis CA, Rossant J. Subtle cerebellar phenotype in mice homozygous for a targeted deletion of the En-2 homeobox. *Science* 1991;**251**:1239-43.
8 **Petit E**, Herault J, Martineau J, Perrot A, Barthelemy C, Hameury L, Sauvage D, Lelord G, Muh JP. Association study with two markers of a human homeogene in infantile autism. *J Med Genet* 1995;**32**:269-74.
9 **Hirakawa M**, Tanaka T, Hashimoto Y, Kuroda M, Takagi T, Nakamura Y. JSNP: a database of common gene variations in the Japanese population. *Nucleic Acids Res* 2002;**30**:158-62.
10 **Geschwind DH**, Sowinski J, Lord C, Iversen P, Shestack J, Jones P, Ducat L, Spence SJ. The autism genetic resource exchange: a resource for the study of autism and related neuropsychiatric conditions. *Am J Hum Genet* 2001;**69**:463-6.
11 **Mukhopadhyay N**, Almasy L, Schroeder M, Mulvihill WP, Weeks DE. Mega2, a data-handling program for facilitating genetic linkage and association analyses. *Am J Hum Genet* 1999;**65**:A436.
12 **Hauser ER**, Boehnke M, Guo SW, Risch N. Affected-sib-pair interval mapping and exclusion for complex genetic traits: sampling considerations. *Genet Epidemiol* 1996;**13**:117-37.
13 **Kruglyak L**, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996;**58**:1347-63.
14 **Abecasis GR**, Cherny SS, Cookson WO, Cardon LR. Merlin - rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002;**30**:97-101.
15 **Abecasis GR**, Cardon LR, Cookson WO. A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 2000;**66**:279-92.
16 **O'Connell JR**, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998;**63**:259-66.
17 **Liu J**, Nyholt DR, Magnussen P, Parano E, Pavone P, Geschwind D, Lord C, Iversen P, Hoh J, Ott J, Gilliam TC. A genomewide screen for autism susceptibility loci. *Am J Hum Genet* 2001;**69**:327-40.
18 **Alarcon M**, Cantor RM, Liu J, Gilliam TC, Geschwind DH. Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families. *Am J Hum Genet* 2002;**70**:60-71.
19 **Silverman JM**, Smith CJ, Schmeidler J, Hollander E, Lawlor BA, Fitzgerald M, Buxbaum JD, Delaney K, Galvin P. Symptom domains in autism and related conditions: evidence for familiality. *Am J Med Genet* 2002;**114**:64-73.

20 **Barthelemy C**, Adrien JL, Tanguay P, Garreau B, Fermanian J, Roux S, Sauvage D, Lelord G. The Behavioral Summarized Evaluation: validity and reliability of a scale for the assessment of autistic behaviors. *J Autism Dev Disord* 1990;**20**:189-204.
21 **Lander ES**, Schork NJ. Genetic dissection of complex traits. *Science* 1994;**265**:2037-48.
22 **Abecasis GR**, Noguchi E, Heinzmann A, Traherne JA, Bhattacharyya S, Leaves NI, Anderson GG, Zhang Y, Lench NJ, Carey A, Cardon LR, Moffatt MF, Cookson WO. Extent and distribution of linkage disequilibrium in three genomic regions. *Am J Hum Genet* 2001;**68**:191-7.
23 **A full genome screen for autism with evidence for linkage to a region on chromosome 7q**. International Molecular Genetic Study of Autism Consortium. *Hum Mol Genet* 1998;**7**:571-8.
24 **Philippe A**, Martinez M, Guilloud-Bataille M, Gillberg C, Rastam M, Sponheim E, Coleman M, Zappella M, Aschauer H, Van Maldergem L, Penet C, Feingold J, Brice A, Leboyer M, van Maldergerm L. Genome-wide scan for autism susceptibility genes. Paris Autism Research International Sibpair Study. *Hum Mol Genet* 1999;**8**:805-12.
25 **Ashley-Koch A**, Wolpert CM, Menold MM, Zaeem L, Basu S, Donnelly SL, Ravan SA, Powell CM, Qumsiyeh MB, Aylsworth AS, Vance JM, Gilbert JR, Wright HH, Abramson RK, DeLong GR, Cuccaro ML, Pericak-Vance MA. Genetic studies of autistic disorder and chromosome 7. *Genomics* 1999;**61**:227-36.
26 **Risch N**, Spiker D, Lotspeich L, Nouri N, Hinds D, Hallmayer J, Kalaydjieva L, McCague P, Dimiceli S, Pitts T, Nguyen L, Yang J, Harper C, Thorpe D, Vermeer S, Young H, Hebert J, Lin A, Ferguson J, Chiotti C, Wiese-Slater S, Rogers T, Salmon B, Nicholas P, Myers RM . A genomic screen of autism: evidence for a multilocus etiology. *Am J Hum Genet* 1999;**65**:493-507.
27 **Barrett S**, Beck JC, Bernier R, Bisson E, Braun TA, Casavant TL, Childress D, Folstein SE, Garcia M, Gardiner MB, Gilman S, Haines JL, Hopkins K, Landa R, Meyer NH, Mullane JA, Nishimura DY, Palmer P, Piven J, Purdy J, Santangelo SL, Searby C, Sheffield V, Singleton J, Slager S. An autosomal genomic screen for autism. Collaborative linkage study of autism. *Am J Med Genet* 1999;**88**:609-15.