The metabotropic glutamate receptor 8 gene at 7q31: partial duplication and possible association with autism

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

Family and twin studies have shown a high monozygotic to dizygotic twin risk ratio in autistic disorder and a slight relative risk between 50 to 100, suggesting that the predisposition to develop autism is largely genetically determined. The family, twin, and linkage data suggest that inheritance of autism is complex, and multiple genes interacting in variable combinations in additive, multiplicative, epistatic, or as yet unknown fashions are involved.

There is strong support for a role of the chromosome 7q21-34 region in the etiology of both autism and language disorders. Several whole genome screens have found some evidence of excess allele sharing in an overlapping 60 cM region of 7q21-34 in autism. FOXP2, a gene on 7q31, is mutated in a severe form of speech and language impairment segregating within a single large pedigree, but a recent study found no association with autism. A recent report has found nominally significant association of the WNT2 gene at 7q31 with autism but with a subsequent study has reported no association.

A recent study has suggested that the reelin gene at 7q22 may be one of the genes contributing to possible linkage between chromosome 7 and autistic disorder, but most likely not the major one.

The 7q31 region contains the GRM8 gene, a good positional and functional candidate for susceptibility to autism. Based upon animal model, neurochemical, and neuropharmacological studies, autism has been proposed to be a hypoglutamatergic disorder. There is a report of alterations in the glutamate neurotransmitter system in postmortem brains of autistic subjects. A protein form of synaptic dependent synaptic plasticity, long term depression, which is triggered by activation of metabotropic glutamate receptors, is selectively enhanced in the hippocampus of mutant mice lacking FMRP. Given that many fragile X syndrome patients exhibit autistic features, this finding suggests that the metabotropic glutamate receptors may play a role in susceptibility to autism. Finally, a recent study has reported linkage and association of the inotropic GluR6 gene on chromosome 7 with autism.

Glutamate is a major excitatory neurotransmitter in human brain and activates both the inotropic (NMDA, AMPA, and kainate) and metabotropic glutamate receptors (mGluRs). GRM8 is a member of the mGluRs. The mGluRs are a family of G protein coupled receptors that have been divided into three groups. Group I includes GRM1 and GRM5 and these receptors activate phospholipase C. Group II (GRM2 and GRM3) and group III (GRM4, 6, 7 and 8) are negatively coupled to cyclic AMP production, but differ in their agonist selectivities.

A previous study has shown that the GRM8 gene encompasses approximately 1000 kb of DNA at the boundary of the 7q31.3-q32.1 bands and has at least seven exons. The sequence of the complete GRM8 gene is now available in the GenBank database. Comparison of the GRM8 cDNA (Accession NM_000845.1) with human chromosome 7 working draft sequence of the 7p14 region show that the gene spans 804 657 bp and includes 10 exons. The GRM8 gene has three spliced isoforms including a possible secreted isoform. The 3′ part of the GRM8 gene codes for the region of the protein involved in interaction with intracellular proteins and receptor signalling. We investigated SNPs in intron 8, exon 10, and the 3′ region of the GRM8 gene for evidence of association with autism using 196 multiplex autistic disorder families. TDT and quantitative TDT studies showed a nominally significant difference in the transmission of the SNPs and haplotypes to the autistic subjects.

These data suggest the presence of a susceptibility mutation in linkage disequilibrium with variants in the GRM8 gene.

**Key points**

- Several independent genome scans have shown linkage to the chromosome 7q21-32 region. The 7q31 region contains the metabotropic glutamate receptor 8 (GRM8) gene, a good positional and functional candidate for susceptibility to autism.
- We screened five single nucleotide polymorphisms (SNPs) in exons 4, 6, 7, and 8, but these SNPs were uninformative and not useful for association studies. Investigation of SNPs in exons 3, 4, and 5 was also complicated by the presence of an intrachromosomal segmental duplication of the GRM8 gene that resides on 7p14 and contains nearly identical sequence to exons 3 to 5.
- We investigated informative SNPs in intron 8, exon 10, and the 3′ region of the GRM8 gene for evidence of association with autism using 196 multiplex autistic disorder families. TDT and quantitative TDT studies showed a nominally significant difference in the transmission of the SNPs and haplotypes to the autistic subjects.
- These data suggest the presence of a susceptibility mutation in linkage disequilibrium with variants in the GRM8 gene.

**METHODS**

**Subjects**

DNA samples from 196 families were obtained from the Autism Genetic Resource Exchange ( AGRE). AGRE, developed and maintained by the Cure Autism Now Foundation, is a central repository of family DNA samples for genetic studies of autism. AGRE families included at least two affected members with a diagnosis of autism, Asperger syndrome, or pervasive developmental disorder (PDD) not otherwise specified. Diagnoses of AGRE families were confirmed using the Autism Diagnostic Interview-Revised protocol (ADI-R). The AGRE families have been used in a genome screen for autism and genotype data from 200 families at an average 10 cM resolution are available online at the AGRE web site (www.agre.org). ADI-R phenotypic data including ADI-R items and algorithm scores are also available from the AGRE. DNA samples from 196 families were obtained from the Autism Genetic Resource Exchange ( AGRE). AGRE, developed and maintained by the Cure Autism Now Foundation, is a central repository of family DNA samples for genetic studies of autism. AGRE families included at least two affected members with a diagnosis of autism, Asperger syndrome, or pervasive developmental disorder (PDD) not otherwise specified. Diagnoses of AGRE families were confirmed using the Autism Diagnostic Interview-Revised protocol (ADI-R). The AGRE families have been used in a genome screen for autism and genotype data from 200 families at an average 10 cM resolution are available online at the AGRE web site (www.agre.org). ADI-R phenotypic data including ADI-R items and algorithm scores are also available from the AGRE web site. For the present study, a total of 890 subjects were available.

**GRM8 genotyping for linkage and association analyses**

Genotyping of a G/A SNP (dbSNP ID 237731) in intron 8 of the GRM8 gene was performed by a restriction enzyme based

![Image](http://jmg.bmj.com/40/4/e42)
Statistical and genetic analysis

The programs Pedcheck and MERLIN were used to find Mendelian errors.22–24 Mistyping analyses were also performed using SimWalk2.25 Samples with apparent recombination between the markers were genotyped. Marker allele frequencies were obtained by counting parental genotypes. Map distances were obtained from Marshfield and Genethon. Linkage analysis (NPL scores) was performed by GENEHUNTER 2.1 software packages.21 Linkage disequilibrium between the markers was analysed by the SimWalk2 and GOLD software packages using parental gametic haplotypes.24–26 The standardised pairwise disequilibrium value $D'$ and coefficient of disequilibrium D were calculated. The coefficient of disequilibrium, D, is the difference between the observed haplotype frequency and the frequency expected under statistical independence. The $D'$ measure is a proportion of the maximum value of D, whose range extends from –1 to +1, with –1 and +1 representing complete LD and 0 representing free association.

Family based association analyses were performed using the transmission disequilibrium test (TDT), where preferential allelic transmission from heterozygous parents to affected offspring is tested by applying the (b-c)/(b+c) statistics and the $\chi^2$ test.27 TDT was performed by randomly selecting one affected subject from each multiplex family. In this situation, TDT is a valid test for association. Haplotypes of the SNPs located in intron 8, exon 10, and GRM8 3' region were constructed on the basis of transmission patterns in families in which both the parents were genotyped. When TDT is applied to multilocus haplotypes, a bias may be introduced in some families for which both parents have the same heterozygous genotype at some locus.28 Elimination of all cases in which the two parents are identically heterozygous, whether or not the offspring can reconstruct phase, results in a conservative test but a robust estimate of gene effect as estimated by transmission ratio. To eliminate this bias, all families with doubly heterozygous parents were discarded so that the standard TDT may be used. For each haplotype, a test with 1 df for excess transmission of that haplotype was calculated. Finally, a global test was also performed with H-1 df, by summing the chi-squared values for each haplotype and multiplying the sum by (H-1)/H, where H is the number of haplotypes for which transmission data are available. This statistic is approximately chi-squared with H-1 degrees of freedom.29–34

We used the QTDT (Quantitative TDT) software package to test association between the GRM8 haplotypes and quantitative traits.26–29 ADI-R items A12 (age at first word) and A13 (age at phrase speech) 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 <6 years or 72 months if >6 years. These three ADI-R items have shown increased phenotypic correlations between sibs,30–32 and have been used in previous trait linkage studies.33,34 QTDT uses a variance component model that partitions association into between and within family components.35–37 When an orthogonal model of association is used, QTDT treats linkage and association separately. Consequently, in contrast to other methods that provide tests of disequilibrium only in minimal family configurations,35,38 the orthogonal method does not detect linkage in the absence of disequilibrium in nuclear families of any configuration. In QTDT, evidence of association is evaluated by a likelihood ratio test. To accommodate parent of origin effects, the additive component of variance can be partitioned into a component that reflects the influence of the quantitative locus (QTL) carried on the maternally derived chromosome and a component that reflects the influence of the QTL carried on the paternally derived chromosome. The likelihood ratio test can be used to test for a QTL on either the maternally or paternally derived chromosome. Alternatively a full modelling for association with alleles derived from either or both parents can be conducted by comparing the model allowing for parent of origin effects with a model in which both maternal and paternal additive variance components are constrained to equal 0. In addition to the parent of origin analysis, other options that are available in the QTDT analysis include use of covariates such as sex, parental phenotype, or user defined covariates and various association models.

RESULTS

Screening of exonic variants and partial duplication of the GRM8 gene

A review of the SNP databases showed that 10 SNPs are available in exons 1, 4, 7, 8, 9, and 10 of the GRM8 gene. In the initial screening of four variants in exon 6 (G/A variant at cDNA position 1232, dbSNP: rs2234947), exon 7 (C/T variant at cDNA position 1521, dbSNP: rs769198), and exon 8 (a T/G variant at cDNA positions 1700, dbSNP: rs2234948 and a C/G variant at cDNA position 2215, dbSNP: rs769201), all subjects had only one allele. The same allele was present in the autistic subjects and unaffected sibs suggesting that these variants do not represent mutation in the affected subjects. The annotations in the SNP databases (NCBI dbSNP, the Human Genome Variation Database, and the Japanese SNP database) also showed that the frequencies of the rare allele of these SNPs were approximately 0.01 or less. These findings suggest that these SNPs represent rare variants and are not useful for association studies in our sample of 196 families.

Genotyping of a variant in exon 4 (apparent C/G variant, cDNA position 960, dbSNP: rs2234946) showed homoyzogous alleles in 175 subjects from 50 families with autism spectrum disorder. This was corroborated by sequencing of exon 4 in 16 subjects with autistic disorder. The apparently homoyzogous genotype in all subjects tested suggested the presence of a pseudogene for GRM8 with nearly identical sequence for exon 4. Blast analysis using exon 4 sequence showed 99% identity with sequences from RP13-157F18 clone (Accession number AC104057.4) which maps to 7p14. Blast analysis also showed the presence of duplicated exons 3–5 on 7p14, suggesting duplicative transposition of 49 kb of genomic DNA from 7q31 to 7p14. The duplicated sequences contain exons and introns as well as repeat elements (HindIII). The GRM8 DNA segment spanning part of intron 2, exons 3–5, and part of intron 5 from nucleotide (nt) 19084809 to nt 20034838 of the GRM8 gene was used as a probe.
There is evidence that the C terminal region of the gene is involved in the regulation of the GRM8 function. We investigated three apparently non-functional but informative SNPs in the GRM8 gene for evidence of association with autism. The G/A SNP (dbSNP: rs2237731) in intron 8 of the GRM8 gene is located 16.3 kb from exon 9 and approximately 70 kb from exon 10. The C/T polymorphism (dbSNP: rs712723) in exon 10 and the C/A G/A SNPs were in strong linkage disequilibrium (intron 8 and exon 10: n=750, df=1, \( \chi^2 = 381, p=0.0000, D=0.169, \) and \( D'=0.788 \); intron 8 and 3′ SNP: n=749, df=1, \( \chi^2 = 674, D=0.168, \) and \( D'=0.972 \)). The transmission of alleles and haplotypes of three SNPs from heterozygous parents to affected offspring was examined to determine their influence on susceptibility to autism. To eliminate this bias, all families with doubly heterozygous parents were discarded so that the standard TDT may be used. There was a nominally significant association between autistic subjects and the intron 8 G/A SNP (p=0.056), the exon 10 C/T and 3′ SNPs were in strong linkage disequilibrium. The allele frequencies of the exon 10 SNP were 0.42 for the C allele and 0.58 for the T allele. There was no deviation from Hardy-Weinberg equilibrium. The allele frequencies of the intron 8, exon 10, and 3′ SNPs were 0.43 for the A allele and 0.57 for the G allele. The transmission of alleles and haplotypes of three SNPs from heterozygous parents to affected offspring was examined to determine their influence on susceptibility to autism. The transmission of alleles and haplotypes of three SNPs from heterozygous parents to affected offspring was examined to determine their influence on susceptibility to autism. To test for linkage of autism to the intron 8, exon 10, and 3′ SNPs using the software package GENEHUNTER. We performed two point non-parametric sib pair analysis to test for linkage of autism to the intron 8, exon 10, and 3′ SNPs using the software package GENEHUNTER. For example, the mean age at first diagnosis among autistic subjects was 3.2 years. Among subjects with autistic disorders, the GRM8 haplotypes were associated with differences in the language quantitative scores in the ADL-R.

![Figure 1](image-url)
word (ADI-R item A12) for subjects who are homozygous for the GTG haplotype; which was more often transmitted to subjects with autistic disorder, was 33.5 months (n=81, SD 20), whereas the mean age at first word for subjects who are homozygous for the ACA haplotype, which was less often transmitted to subjects with autistic disorder, was 28 months (n=21, SD 18).

Quantitative TDT studies
The QTDT package was used to test association between the haplotypes of GRM8 gene and the quantitative traits. Age at first word, age at phrase speech, and ADI-R scores for stereotypic behaviour were used as quantitative traits. Some types of non-normality in phenotypic distribution, such as marked leptokurtosis, inflate type 1 error rates in variance component studies, especially when there is high phenotypic correlations among sibs. The quantitative traits, age at first word (n=275, mean=31.1 months, SD=17.7, range 10-85, kurtosis -0.246, and skewness 0.7), age at phrase speech (n=233, mean=45 months, SD=18.4, range 11-108, kurtosis -0.207, and skewness 0.434) and composite algorithm score for stereotypic behaviour (n=377, mean=5.53, SD=2.6, range 0-12, kurtosis -0.083, and skewness 0.361) did not show such leptokurtosis.

QTDT uses a variance component model and evaluates association by a likelihood ratio test. The variance components include common environment shared by all family members, non-shared environment unique to each family member, polygenic effects that are a function of relatedness between family members, and additive major gene effect that represents the additive effect of linkage to a major gene. There was evidence for polygenic effects for the three quantitative traits (age at first word, age at phrase speech, and stereotypic behaviour) to quantitative traits (the most significant result: GTG haplotype and age at phrase speech, p=0.2; nine tests, overall Bonferroni significance level 0.87). Using sex as a covariate and a saturated model for parent of origin effects where paternal and maternal alleles were modelled separately (options "–cs –of " in the qtdt analysis), the most significant association was between the GTG haplotype and age at first word (p=0.0021, nine tests, overall Bonferroni significance level 0.018). There was significant difference in the transmission of maternal and paternal alleles (p=0.0007, nine tests, overall Bonferroni significance level 0.006). The QTDT program includes a permutation framework, which allows the derivation of empirical p values for the sample being evaluated. These permutations condition on the trait distribution, linkage, and familiality, and correct for small sample sizes or deviation of quantitative traits from multivariate normality. When we performed 1000 permutations for each of the above traits, the most significant association was between the GTG haplotype and age at first word (p=0.009).

DISCUSSION
GRM8 is a member of metabotropic glutamate receptors, which negatively modulate glutamate transmission and thus serve to prevent pathological changes in neuronal hyperexcitability and homeostasis. Thus, dysfunction of the GRM8 gene may result in neuronal damage. Recent studies have shown presynaptic and glial cell localisation of GRM8 in several regions of the brain including pontine grey, lateral reticular nucleus of the thalamus, olfactory bulb, and pyriform cortex and to a lesser extent hippocampus, caudate, cerebellum, and cerebral cortex. These regions have been consistently reported to be anatomically abnormal in autistic subjects.

TDT studies were suggestive of linkage disequilibrium between autistic disorder and variants in the GRM8 gene. Our studies showed nominally significant association of the autistic disorder with SNPs and haplotypes of the GRM8 3’ gene region, which is consistent with a susceptibility mutation in the GRM8 or a nearby gene. So far we have performed linkage disequilibrium and association analyses with informative but apparently non-functional SNPs. The modest level of evidence may be a chance finding, but it is also consistent with a gene

<table>
<thead>
<tr>
<th>Table 1</th>
<th>TDT studies of GRM8 SNPs and haplotypes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allele or haplotype</td>
<td>T</td>
</tr>
<tr>
<td>Intron 8 SNP</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>G</td>
</tr>
<tr>
<td>Exon 10 SNP</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>T</td>
</tr>
<tr>
<td>3’ SNP</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>G</td>
</tr>
<tr>
<td>Haplotype of intron 8 and exon 10 SNPs</td>
<td>AC</td>
</tr>
<tr>
<td></td>
<td>AT</td>
</tr>
<tr>
<td></td>
<td>GC</td>
</tr>
<tr>
<td></td>
<td>GT</td>
</tr>
<tr>
<td>Haplotype of exon 10 and 3’ SNPs</td>
<td>CA</td>
</tr>
<tr>
<td></td>
<td>TA + CG</td>
</tr>
<tr>
<td></td>
<td>TG</td>
</tr>
<tr>
<td>Haplotype of intron 8, exon 10, and 3’ SNPs</td>
<td>ACA</td>
</tr>
<tr>
<td></td>
<td>ATA + GTA</td>
</tr>
<tr>
<td></td>
<td>ATG</td>
</tr>
<tr>
<td></td>
<td>GCA</td>
</tr>
<tr>
<td></td>
<td>GTG</td>
</tr>
</tbody>
</table>

T indicates transmitted and NT indicates not transmitted.
of moderate effect or a gene acting epistatically with other genes. Linkage disequilibrium studies using additional variants throughout the gene as well as studies in a larger sample are necessary to fully assess the role of the **GRM8** gene in susceptibility to autism.

Quantitative phenotype scores contain more information than is provided by dichotomous traits. The 7q linkage evidence obtained by analysing the quantitative language items was several fold stronger than that obtained by the qualitative 10 cM scan based on the narrow or broad autism diagnosis. In this report, the QTDT evidence for association of language quantitative traits (age at first word or age at phrase speech) in autism with haplotypes of the **GRM8** gene was stronger than that provided by TDT studies of the discrete traits.

We also report here an intrachromosomal segmental duplication of the **GRM8** gene that resides on 7p14 and contains a nearly identical sequence to exons 3-5 including introns 3 and 4. This region of 7p14 contains duplicated stretches of mosaic and juxtaposed sequences originating from different parts of chromosome 7 and other chromosomes (data not shown). The mechanism of the duplication is unclear, but junctions of the duplicated fragments contain Line-like repeats L1, L2, and simple repeat elements.

The duplicated region corresponding to **GRM8** exons 3-5 are 99% homologous with the **GRM8** exons with six nucleotide differences between the exons and the duplicated segment. The nucleotide differences do not introduce any stop codon but change the amino acid composition at cDNA (Accession NM_000845.1) positions 859 (GCC to TGC: arginine to cystine), 1085 (CGA to CAA: arginine to glutamine), and 1100 (CGA to CAA: arginine to glutamine). Genotyping of an apparent G/C polymorphism corresponding to the cDNA (Accession NM_000845.1) position 960 (fig 1B, “g” in **GRM8** and “c” at the duplication) showed heterozygous genotypes in 175 subjects suggesting that the duplication is present in all subjects. Genotyping of an apparent G/A (“a” in **GRM8** and “g” at the duplication) variant in intron 4 showed homozygous GG sequences in some subjects, suggesting that this variant is polymorphic in some subjects at the **GRM8** locus.

This segmental duplication may impact linkage and association studies involving the **GRM8** gene. The SNP maps in the databases over the exons 3-5 regions of the **GRM8** gene are hampered by the duplication, leading to the annotation of SNPs that simply represent variation between duplicated segments and the **GRM8** gene. Some SNPs cannot be tracked owing to lack of unique sequence information. A recent report indicates that such duplicated sequences may predispose to large scale mitotic polymorphisms associated with complex genetic traits. There is also a report of large deletions at four different sites in chromosomes 7 and 8 in kindreds with autism.

Although it appears that the segmental duplication of the **GRM8** gene is present in all subjects, fluorescence in situ hybridisation studies using appropriate probes will be required to resolve whether the duplication is mosaic or polymorphic. Additional studies will also help to clarify whether this duplication predisposes to mutations at the **GRM8** gene locus by nucleotide exchange from the duplicated sequences at 7p14.

**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, MD, PhD, UCLA, Los Angeles, CA; Members: W Ted Brown, MD, PhD, FACMG, NYS Institute for Basic Research in Developmental Disabilities, Long Island; Maya Bucan, PhD, University of Pennsylvania, Philadelphia, PA; Joseph Buxbaum, PhD, Mount Sinai School of Medicine, NY, NY; Edwin H Cook, MD, University of Chicago, Chicago, IL; T Conrad Gilliam, PhD, Columbia Genome Center, NY, NY; David A Greenberg, PhD, Mount Sinai Medical Center, NY, NY; David H Ledbetter, PhD, University of Chicago, Chicago, IL; Bruce Miller MD, UCSC, San Francisco, CA; Stanley F Nelson, MD, UCLA School of Medicine, Los Angeles, CA; Jonathan Peiris, PhD, Kennedy Krieger Institute, Baltimore, MD; Jerome J Rotter, MD, Cedars-Sinai Medical Center, Los Angeles, CA; Gerard D Schellenberg, PhD, University of Washington, Seattle, WA; Carol A Sprouse. EdD, Children’s National Medical Center, Baltimore, MD; Rudolph E Tanzi, PhD, Massachusetts General Hospital, Boston, MA; Kirk C Wilhelmsen, MD, PhD, University of California, San Francisco, CA. AGRE Coordinators: Jeremy M Silberman, PhD, Mount Sinai Medical School, New York, NY.

**Authors’ affiliations**

F J Serunjogi, H Zhong, 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, 3901 Beaubien Blvd., Detroit, MI 48201, USA; ahuq@med.wayne.edu

**REFERENCES**


40 Genetics, Estivill X. A polymorphic genomic duplication on human chromosome 15 is a susceptibility factor for panic and phobic disorders. Cell 2001;106:367-79.