Family-Based Association Study Showing that Immunoglobulin A Nephropathy Is Associated with the Polymorphisms 2093C and 2180T in the 3' Untranslated Region of the *Megsin* Gene

YOU-JI LI,* YONG DU,* CAI-XIA LI,† HUI GUO,‡ JOSEPH C.K. LEUNG,§ MAN F. LAM,§ NIANSHENG YANG,* FENGXIAN HUANG,* YUN CHEN,‡ JI-QIAN FANG,† PATRICK H. MAXWELL,¶ KAR N. LAI,§ and YIMING WANG‡

**Department of Nephrology, 1st Affiliated Hospital, †Department of Medical Statistics, and ‡Department of Medical Genetics, Zhongshan Medical College, Sun Yat-Sen University, Guangzhou, People's Republic of China; §Department of Medicine, Queen Marry Hospital, The University of Hong Kong, Hong Kong; and ¶Department of Medicine, Hammersmith Hospital, Imperial College, London, United Kingdom.*

Abstract. Immunoglobulin A nephropathy (IgAN) is considered to be a multifactorial disease with genetic and environmental factors contributing to its pathogenesis. The genes involved in susceptibility and progression of the disease have not yet been clearly elucidated. *Megsin* (SERPINB7) is an important candidate gene, predominantly expressed in glomerular mesangium and upregulated in IgAN. To investigate the potential role of this and other genes in IgAN, patients with biopsy-proven IgAN were recruited, as were family members, for a family-based association study. The genotypes of the polymorphisms C2093T and C2180T within the 3' untranslated region of the gene were determined by polymerase chain reaction–restriction fragment length polymorphism and direct sequencing. The results were analyzed by transmission disequilibrium test (TDT) and haplotype relative risk (HRR). TDT analyses revealed that *Megsin* 2093C and 2180T alleles were significantly more transmitted from heterozygous parents to patients than expected (C2093T: 127 trios, $P = 0.034$, C2180T: 100 trios, $P = 0.002$). Extended TDT showed increased cotransmission of the 2093C and 2180T alleles (232 families, $P < 0.001$). HRR revealed that the 2093C and 2180T alleles were more often transmitted to patients ($P = 0.014$, <0.001, respectively). Genetic variation in *Megsin* confers susceptibility to IgAN.

IgA nephropathy (IgAN) is the most common primary glomerular disease (1–3). The prevalence is probably higher among Southeast Asians (3–5), although the exact incidence is unknown and available data are likely to be heavily influenced by differing criteria for nephrological assessment and renal biopsy (6). The etiology is considered to be multifactorial. Evidence for a genetic contribution has come from the observation that some families are multiply affected in an apparently autosomal dominant fashion, although this is uncommon. Furthermore, a linked locus has been demonstrated in a number of families (7). Epidemiologic studies suggest a significant difference in the incidence/prevalence among different ethnic groups and support significant familial aggregation (6). Data from mouse studies, including the outbred strain ddY (8) and knockout and transgenic approaches (9–11), have shown that specific genetic alterations in a model organism give rise to a phenotype that shares features with the human disease.

Megsin, also known as SERPINB7, serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 7 (http://bioinfo.weizmann.ac.il, GenBank ID AF027866), is a predominantly mesangial expressed gene and belongs to the serpin (serine proteinase inhibitor) superfamily (12). Immunohistochemistry and *in situ* hybridization studies have shown that it is upregulated in IgAN when compared with normal renal tissue and several other forms of glomerulonephritis. The upregulation coincides with mesangial proliferation and extracellular matrix expansion (12–14). *Megsin* is therefore a good candidate for involvement in human IgAN. However, a previous case-control study of a polymorphism, C2093T, did not demonstrate a significant association (15).

Here we report the assembly of a DNA collection in China suitable for family-based analysis and investigate whether there is evidence for association of disease with transmission of two polymorphisms of *Megsin*, C2093T and C2180T, located in the 3' untranslated region (UTR) of the gene.

Received October 24, 2003. Accepted April 14, 2004.
Correspondence to Prof. Yiming Wang, Dept of Medical Genetics, Zhongshan Medical College, 74 Zhongshan Road II, Guangzhou 510089, PR China. Phone: 86-20-87333136; Fax: 86-20-87335785; E-mail: ywzhong@hotmail.com

1046-6673/1507-1739
Journal of the American Society of Nephrology

DOI: 10.1097/01.ASN.0000130858.00688.46

Materials and Methods

Patients, Families, Normal Control Subjects

The patients were recruited from centers in Guangzhou and Hong Kong, China. After informed consent was obtained, peripheral blood was drawn into EDTA tubes. Clinical data were also recorded (age, gender, duration of observation, proteinuria, serum creatinine, BP). The parents of the patients were also recruited, as well as unaffected siblings if one or both of the parents were unavailable. Informed consent was obtained from all family members. Blood samples of the patients, and blood or saliva of the parents or siblings, were collected in EDTA tubes or on 4-mm filter papers in the case of saliva collection. Blood samples from gender- and age-matched unrelated normal subjects were also obtained on a voluntary basis from Guangdong province.

DNA Extraction

Genomic DNA was extracted with QIAamp DNA blood Maxi kits according to the manufacturer's instructions. DNA was extracted from saliva with chelex-100 as described by Ohhashi *et al.* (16).

Genotyping of Megsin C2093T and C2180T

C2093T was genotyped by PCR amplification of a 256-bp fragment in the 3' UTR of the gene with a 9700 thermocycler (Eppendorf 5331, Germany). The primer sequences were: sense: 5'-TTG TTG ACC TAT GAA GAT TTT AGA G-3', antisense: 5'-AAA CTT ATA AAC TAC ACA GCA TAT GA-3'. The reaction mixture contained l x PCR buffer, l.5 mmol/L $MgCl_2$, 200 mmol/L deoxynucleotide triphosphates (dNTPs), l unit *Taq*DNA polymerase (MBI), 10 pmol of each primer and 100 to 200 ng genomic DNA. The PCR amplification reaction consisted of a cycle at 95°C for 10 min, followed by 35 cycles of denaturation at 94°C for 15 s, annealing at 51°C for 15 s, and extension at 72°C for 30 s. A final extension was performed at 72°C for 5 min. The PCR products were digested with restriction endonuclease *Hae*III (New England Biolabs) and electrophoresed on 2% agarose gel containing ethidium bromide (1.5 μg/ml). The 2093C allele produces 114- and 142-bp fragments; the T allele has no digestion site, as described previously (15). The genotyping results were also confirmed by direct sequencing in randomly selected samples.

C2180T was genotyped by PCR amplification with the same primers described above or the following primers: sense: 5'-TCT TTT AAC TGT TGG CAG TTG TT-3' and antisense: 5'-CAA AGA AAG CCC TAG TTG TCC-3'. The PCR products were sequenced on an ABI 3100 Genetic Analyzer with a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems).

Statistical Analyses

Family-Based Studies. The TRANSMIT program was used for the transmission disequilibrium test (TDT) analysis (17). Two TDT methods were used: the classical TDT of transmission from heterozygous parents in complete trios (patients plus both parents) (18), and the extended TDT, which used information from all families (complete trios regardless of the heterozygosity status of the parents, plus single-parent families) (17). Haplotype relative risk (HRR) was analyzed as described elsewhere (19,20).

Case-Control Analysis. χ^2 analysis was used when comparing allele frequencies between the groups. Hardy-Weinberg equilibrium was tested by a χ^2 test with 1 degree of freedom.

Species Comparisons. We aligned the sequence of human with rat and mouse by using multizMm3Rn2, available at http://genome.ucsc.edu. The sequences of cow, sheep, and horse were obtained from GenBank and aligned with human sequence by the pairs alignment tool BioEdit, available at http://www.ncbi.nlm.nih.gov/.

Results

A total of 423 patients with IgAN were recruited over an initial 5-yr period. IgAN was diagnosed by World Health Organization criteria (21), with a renal biopsy performed to assess mesangial IgA deposition via immunofluorescence. Recruitment required that patients were younger than 60 yr old, and did not have significant hepatic disease or Henoch Schonlein purpura. The clinical data of the patients at the time of renal biopsy are listed in Table 1.

A case-control analysis of 423 patients suggested that the 2093C allele frequency was significantly higher among the patients than controls ($P = 0.006$) (Table 2). Although this would be consistent with an effect of the 2093C allele on susceptibility and/or manifestations of IgAN, an important limitation of such comparisons is population stratification between the affected subjects and the control population. Our study was designed to address this issue by examining whether there was distortion from the predicted chance transmission of parental alleles. TDT analysis of this genetic variant could be undertaken for 127 complete trios; in the other 28 complete trios, both of the parents were homozygotes (Tables 3 and 4). The TDT showed that the C allele was significantly more transmitted to patients from their parents than would be expected with a P value of 0.034 (Table 4). An extended TDT analysis, including the 28 trios where the parents were homozygous and 83 families in which only one parent was available, gave a similar result ($P = 0.032$) (Tables 3 and 4). The HRR showed that the C

Table 1. Clinical data of the patients at time of renal biopsy[a]

Gender (F/M)	228/195
Age (yr)	30.1 ± 9.9
Duration of disease (mo)	20.6 ± 34.3
Systolic BP (mmHg)	122.4 ± 43.7
Diastolic BP (mmHg)	78.8 ± 41.8
Proteinuria (g/d)	0.99 ± 1.79
Scr (μmol/L)	109 ± 124.0
Serum CHOL (mmol/L)	5.47 ± 2.29
Serum TG (mmol/L)	1.43 ± 0.96

[a] Values expressed as mean ± SD or number. Scr, serum creatinine; CHOL, cholesterol; TG, triglyceride.

Table 2. Allele frequencies (%) of C2093T in patients and control subjects

	Allele (%)		χ^2	P
	C	T		
IgAN ($n = 423$)	70.7	29.3	7.546	0.006
Control ($n = 201$)	62.9	37.1		

allele is more often transmitted to patients ($P = 0.014$) (Table 5).

To further examine the role of genetic variation in the *Megsin* 3' UTR, we examined a second polymorphism. TDT analysis of the C2180T polymorphism could be undertaken for 100 complete trios in which at least one of the parents was heterozygous, and an extended TDT method including all complete trios and 90 single-parent families gave a total of 241 families (Tables 3 and 6). Both TDT analyses showed that the 2180T alleles were significantly more transmitted to patients from their parents than would be expected ($P = 0.002$ and $P < 0.001$, respectively) (Table 6). The HRR showed that the 2180T allele is more often transmitted to patients ($P < 0.001$) (Table 5).

Because of the differences in the parental heterozygosity status in the two polymorphisms, some families were only entered into TDT analysis for one marker. We therefore also undertook TDT analyses restricted to the 232 families that were typed for both polymorphisms (Tables 7 and 8). This showed a significant overcotransmission of the 2093C and 2180T alleles ($P = 0.0003$) (Table 8).

Table 3. Families assembled and entered into statistical analyses

	C2093T	C2180T
Complete trios genotyped	155	151
Complete trios entered classical TDT	127	100
Single parent families entered extended TDT	83	90
Total families	238	241

Discussion

By using both TDT and HRR analysis for complete trios, our results show that the *Megsin* 2093C and 2180T alleles are transmitted more frequently from heterozygous parents to affected patients than would be expected statistically. The same result was obtained when we extended the analysis to include incomplete trios. We also observed an increased rate of cotransmission of the 2093C and 2180T alleles in the families typed for both. HRR also showed a significantly increased transmission rate of the 2093C and 2180T alleles to the patients. Our case-control analysis also showed that the 2093C allele was present at a significantly higher frequency in the patient group, and conversely, the T allele was present in more of the normal subjects.

These results indicate that in this population the *Megsin* 2093C and 2180T alleles are associated with a predisposition to clinically overt IgAN. A previous case-control study in a different ethnic group that used 110 patients and 104 controls did not observe any difference in allele frequency in the two

Table 4. TDT analyses of *Megsin* C2093T alleles transmitted to patients

	Family Number	Transmitted Allele	Observed	Expected	χ^2	*P*
Classical TDT	127	C	169	155.5	4.472	0.034
		T	85	98.5		
Extended TDT	238	C	323	306.54	4.582	0.032
		T	153	169.46		

Table 5. HRR analyses of C2093T and C2180T transmitted and nontransmitted alleles

Trios Entered Classical TDT	C2093T			C2180T		
	C	T	Total	C	T	Total
Transmitted	169[a]	85	254	55	145[b]	200
Nontransmitted	142	112	254	89	111	200

[a] Haplotype-based HRR $\chi^2 = 6.045$, $P = 0.014$, HRR = 1.568.
[b] Haplotype-based HRR $\chi^2 = 12.543$, $P < 0.001$, HRR = 2.114.

Table 6. TDT analyses of *Megsin* C2180T alleles transmitted to patients

	Family Number	Transmitted Allele	Observed	Expected	χ^2	*P*
Classical TDT	100	C	55	72	9.3226	0.002
		T	145	128		
Extended TDT	241	C	101	126.45	15.298	<0.001
		T	381	355.55		

Table 7. Families that entered into the extended TDT for co-transmission of alleles of C2093T and C2180T[a]

C2093T Trios + C2180T Trios	C2093T SPF + C2180T SPF	C2093T Trios + C2180T SPF	C2093T SPF + C2180T Trios	Total
125	56	25	26	232

[a] SPF, single parent family.

Table 8. TDT analysis[a] of C2093T–C2180T haplotypes

	Family Number	Transmitted Haplotype	Observed	Expected	χ^2	*P*
Extended TDT	232	2093C–2180C	26.457	30.392	1.650	0.199
		2093T–2180C	83.543	97.925	5.911	0.015
		2093C–2180T	295.543	268.904	12.924	0.0003
		2093T–2180T	58.457	66.779	2.785	0.095

[a] Global χ^2 test: $\chi^2 = 13.431$, $P = 0.004$.

groups (15). Several explanations could account for this discrepancy. One possibility is that the sample size may have been too small to detect a relatively modest effect on susceptibility, especially given the difficulties in matching patient and control groups in such studies.

Cloned by a group of Japanese scientists in 1998 (12), *Megsin* rapidly emerged as a candidate gene that might be associated with, or contributing to, various mesangial lessons. It was further supported by transgenic mouse studies showing that overexpression leads to progressive mesangial matrix expansion and an increase in the number of mesangial cells, accompanied by augmented immune complex, Ig, and complement deposition (11). *In vitro* assays have identified that plasmin is one of its substrates and have confirmed its proteinase inhibitory activity (12). Glomerular mesangial cells play a key role in maintaining the normal structure and function of the glomerulus by mediating extracellular matrix remodeling and immune complex disposal. Therefore, one may speculate that by inhibiting these plasmin/proteolytic activities, *Megsin* would contribute to the pathologic processes, which eventually lead to clinically evident IgAN.

Located on chromosome 18q21.3, close to the other serpin gene cluster, the *Megsin* gene contains 8 exons and 7 introns, spanning a 20-kb genomic region, which encodes a predicted peptide of 380 amino acids (22). The transcriptional start site is located 391 bp upstream of the start codon. Its identified regulatory regions are a promoter region located within a 4021-bp interval in the 5' region and a protein A binding motif, CTGATTCAC, located at -120 to -112 (22). The two polymorphisms studied here are located in the 3' UTR, in which regulatory sequences have not yet been identified. Comparing the predicted orthologs in different species shows heterogeneity for the equivalent nucleotides to human *Megsin* 2093 and 2180 positions. The 2093 position is C in mouse, rat, cow and horse, but T in sheep. The 2180 position is T in mouse, rat, and sheep, but A in horse.

Currently, how the 2093C and 2180T alleles are associated with susceptibility is unclear, and we do not know whether the C/T substitutions themselves confer the effect or whether it is due to other variants nearby with which they are in linkage disequilibrium, but the results suggest that genetic variation or variations of *Megsin*, which could be in the 3' UTR, confers susceptibility to IgAN. Further studies will be necessary to precisely locate the causal variation or variations and the mechanism through which the susceptibility is conferred. Interestingly, we observed a 2093C allele frequency of 62.9% among our Chinese control subjects. This is significantly higher than the reported frequency of 48.5% in a recent study in the Hungarian population (15) ($P = 0.001$). It is plausible that this could contribute to a higher prevalence of this disease in the Chinese population.

The study presented here represents what is to our knowledge the first family-based association study investigating the relationship of *Megsin* C2093T and C2180T with IgAN, and our findings strongly suggest a role for variation at the *Megsin* 3' UTR with disease susceptibility. Furthermore, we have assembled a genetic resource that should be of considerable value in assessing the contribution of other genetic variants to this important condition.

Acknowledgment

This project was supported by the National Natural Science Foundation of China (30170434), Guangdong Provincial Natural Science Foundation (013140), and the China Medical Board, Rockefeller Foundation (98-677).

References

1. D'Amico G: The commonest glomerulonephritis in the world: IgA nephropathy. *Q J Med* 64: 709–727, 1987
2. Maisonneuve P, Agodoa L, Gellert R, Stewart JH, Buccianti G, Lowenfels AB, Wolfe RA, Jones E, Disney AP, Briggs D, McCredie M, Boyle P: Distribution of primary renal diseases leading to end-stage renal failure in the United States, Europe,

and Australia/New Zealand: Results from an international comparative study. *Am J Kidney Dis* 35: 157–165, 2000
3. Levy M, Berger J: Worldwide perspective of IgA nephropathy. *Am J Kidney Dis* 12: 340–347, 1988
4. Julian BA, Waldo FB, Rifai A, Mestecky J: IgA nephropathy, the most common glomerulonephritis worldwide: A neglected disease in the United States? *Am J Med* 84: 129–132, 1988
5. Donadio JV, Grande JP: IgA nephropathy. *N Engl J Med* 347: 738–748, 2002
6. Hsu SI, Ramirez SB, Winn MP, Bonventre JV, Owen WF: Evidence for genetic factors in the development and progression of IgA nephropathy. *Kidney Int* 57: 1818–1835, 2000
7. Gharavi AG, Yan Y, Scolari F, Schena FP, Frasca GM, Ghiggeri GM, Cooper K, Amoroso A, Viola BF, Battini G, Caridi G, Canova C, Farhi A, Subramanian V, Nelson-Williams C, Woodford S, Julian BA, Wyatt RJ, Lifton RP: IgA nephropathy, the most common cause of glomerulonephritis, is linked to 6q22–23. *Nat Genet* 26: 354–357, 2000
8. Imai H, Nakamoto Y, Asakura K, Miki K, Yasuda T, Miura AB: Spontaneous glomerular IgA deposition in ddY mice: An animal model of IgA nephritis. *Kidney Int* 27: 756–761, 1985
9. Zhang Z, Kundu GC, Zheng F, Yuan CJ, Lee E, Westphal H, Ward J, DeMayo F, Mukherjee AB: Insight into the physiological function(s) of uteroglobin by gene-knockout and antisense-transgenic approaches. *Ann N Y Acad Sci* 923: 210–233, 2000
10. Zhang Z, Kundu GC, Yuan CJ, Ward JM, Lee EJ, DeMayo F, Westphal H, Mukherjee AB: Severe fibronectin-deposit renal glomerular disease in mice lacking uteroglobin. *Science* 276: 1408–1412, 1997
11. Miyata T, Inagi R, Nangaku M, Imasawa T, Sato M, Izuhara Y, Suzuki D,Yoshino A, Onogi H, Kimura M, Sugiyama S, Kurokawa K: Overexpression of the serpin megsin induces progressive mesangial cell proliferation and expansion. *J Clin Invest* 109: 585–593, 2002
12. Miyata T, Nangaku M, Suzuki D, Inagi R, Uragami K, Sakai H, Okubo K, Kurokawa K: A mesangium-predominant gene, megsin, is a new serpin upregulated in IgA nephropathy. *J Clin Invest* 102: 828–836, 1998
13. Suzuki D, Miyata T, Nangaku M, Takano H, Saotome N, Toyoda M, Mori Y, Zhang SY, Inagi R, Endoh M, Kurokawa K, Sakai H: Expression of megsin mRNA, a novel mesangium-predominant gene, in the renal tissues of various glomerular diseases. *J Am Soc Nephrol* 10: 2606–2613, 1999
14. Nangaku M, Miyata T, Suzuki D, Umezono T, Hashimoto T, Wada T, Yagi M, Nagano N, Inagi R, Kurokawa K: Cloning of rodent megsin revealed its up-regulation in mesangioproliferative nephritis. *Kidney Int* 60: 641–652, 2001
15. Szelestei T, Bahring S, Kovacs T, Vas T, Salamon C, Busjahn A, Luft FC, Nagy J: Association of a uteroglobin polymorphism with rate of progression in patients with IgA nephropathy. *Am J Kidney Dis* 36: 468–473, 2000
16. Ohhashi A, Aoki T, Matsugo S, Simasaki C: PCR-based typing of human buccal cell's DNA extracted from whole saliva and saliva stains. *Nippon Hoigaku Zasshi* 47: 108–118, 1993
17. Clayton D: A generalization of the transmission/disequilibrium test for uncertain-haplotype transmission. *Am J Hum Genet* 65: 1170–1177, 1999
18. Spielman RS, McGinnis RE, Ewens WJ: Transmission test for linkage disequilibrium: The insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 52: 506–516, 1993
19. Falk CT, Rubinstein P: Haplotype relative risks: An easy reliable way to construct a proper control sample for risk calculations. *Ann Hum Genet* 51: 227–233, 1987
20. Terwilliger JD, Ott J: A haplotype-based "haplotype relative risk" approach to detecting allelic associations. *Hum Hered* 42: 337–346, 1992
21. Habib R, Niaudet P, Levy M: Schonlein-Henoch purpura nephritis and IgA nephropathy. In: *Renal Pathology*, edited by Tisher C, Brenner BM, Philadelphia, Lippincott, 1994, pp 472–523
22. Inagi R, Miyata T, Nangaku M, Ueyama H, Takeyama K, Kato S, Kurokawa K: Transcriptional regulation of a mesangium-predominant gene, megsin. *J Am Soc Nephrol* 13: 2715–2722, 2002