IgA nephropathy, the most common cause of glomerulonephritis, is linked to 6q22–23

Ali G. Gharavi[1,4], Yan Yan[1], Francesco Scolari[5], F. Paolo Schena[6], Giovanni M. Frasca[7], Gian Marco Ghiggeri[8], Kerry Cooper[3], Antonio Amoroso[9], Battista Fabio Viola[5], Graziana Battini[10], Gianluca Caridi[8], Cristina Canova[11], Anita Farhi[1], Vairavan Subramanian[1], Carol Nelson-Williams[1], Sue Woodford[13], Bruce A. Julian[14], Robert J. Wyatt[13] & Richard P. Lifton[1,2,3]

End-stage renal disease (ESRD) is a major public health problem, affecting 1 in 1,000 individuals and with an annual death rate of 20% despite dialysis treatment[1,2]. IgA nephropathy (IgAN) is the most common form of glomerulonephritis, a principal cause of ESRD worldwide[1,2]; it affects up to 1.3% of the population[3–6] and its pathogenesis is unknown. Kidneys of people with IgAN show deposits of IgA-containing immune complexes with proliferation of the glomerular mesangium (Fig. 1). Typical clinical features include onset before age 40 with haematuria and proteinuria (blood and protein in the urine), and episodes of gross haematuria following mucosal infections are common; 30% of patients develop progressive renal failure[6–9]. Although not generally considered a hereditary disease, striking ethnic variation in prevalence[1–6,10] and familial clustering[11–16], along with subclinical renal abnormalities among relatives of IgAN cases[9,14–16], have suggested a heretofore undefined genetic component. By genome-wide analysis of linkage in 30 multiplex IgAN kindreds, we demonstrate linkage of IgAN to 6q22–23 under a dominant model of transmission with incomplete penetrance, with a lod score of 5.6 and 60% of kindreds linked. These findings for the first time indicate the existence of a locus with large effect on development of IgAN and identify the chromosomal location of this disease gene.

We studied multiplex IgAN kindreds from Italy (24 kindreds) and the United States (6 kindreds). All were ascertained through biopsy-documented IgAN cases and had at least one more affected member. Relatives provided medical records and were screened for haematuria, proteinuria and renal function. Relatives were classified as affected if they had IgAN on renal biopsy, or, alternatively, haematuria (≥5 red blood cells/high power field) or proteinuria

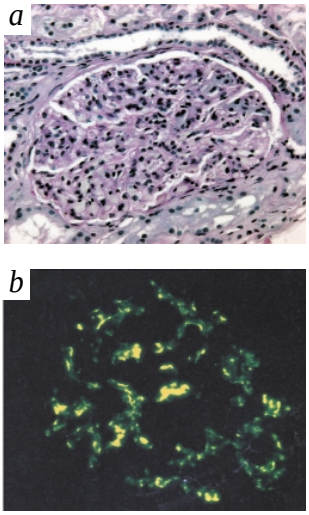

Fig. 1 Pathology of IgAN. A section of a renal biopsy of a patient with IgAN is shown. ***a***, Periodic acid Schiff staining of a glomerulus revealing mesangial expansion and proliferation. ***b***, Immunofluorescence staining of a glomerulus with anti-IgA antiserum demonstrating IgA deposition in the mesangium.

Table 1 • Lod scores for linkage of IgAN to candidate loci

Locus	location	(cM)	Nearest loci	(cM)	Multipoint lod (α=1)	Multipoint lod (α=0.5)
Ig κ-chain cluster	2p12	(115)	*D2S2972*	(114)	–6.9	–0.3
			D2S340	(120)		
Ig λ-chain cluster	22q11	(14)	*D22S420*	(15)	–2.6	–0.45
			GCT10C10	(18)		
Ig heavy chain cluster	14q32.3	(124)	GATA168F06	(113)	–3.5	–0.05
			GATA136B01	(126)		
Ig joining chain	4q25	(77)	*D4S3248*	(70)	–4.5	0.3
			D4S2367	(78)		
MHC	6p21.3	(45)	GATA163B10	(43)	–8.0	–0.5
			D6S1281	(44)		
galactosyltransferase						
T1	2q32	(182)	*D2S1776*	(173)	–4.5	0.3
			D2S1391	(186)		
T2	1q31	(209)	*D1S518*	(202)	–2.9	–1.8
			D1S1660	(212)		
T3	3q25	(178)	*D3S1763*	(176)	–4.8	–2.5
			D3S2440	(181)		
T4	6p21	(45)	GATA163B10	(43)	–8.0	–0.5
			D6S1281	(44)		
IgA Fc receptor (CD89)	19q13.4	(97)	*D19S589*	(88)	–2.8	0.5
			D19S210	(100)		
uteroglobin	11q13	(65)	*D11S1985*	(58)	–11.5	–1.1
			D11S1765	(62)		
			D11S1883	(65)		

Lod scores for linkage of IgAN to indicated candidate loci are shown under dominant models of transmission specifying locus homogeneity or heterogeneity with α=0.5.

[1]*Howard Hughes Medical Institute and Departments of* [2]*Genetics and* [3]*Medicine, Yale University School of Medicine, New Haven, Connecticut, USA.* [4]*Department of Medicine, Mount Sinai School of Medicine, New York, New York, USA.* [5]*Division and Chair of Nephrology, Spedali Civili, University of Brescia, Brescia, Italy.* [6]*Institute of Nephrology, University of Bari, Polyclinic, Bari, Italy.* [7]*Nephrology Division, Policlinico S. Orsola, Bologna, Italy.* [8]*Laboratory and Department of Nephrology, Giannina Gaslini Institute, Genoa, Italy.* [9]*Servizio di Genetica e Cattedra di Genetica, IRCCS Burlo Garofolo, Trieste, Italy.* [10]*Dialysis Service, Denzenzano Hospital, Desenzano, Italy.* [11]*Nephrology Department, St Orsola University Hospital, Bologna, Italy.* [13]*Crippled Children's Research Foundation, University of Tennessee Health Sciences Center, Memphis, Tennessee, USA.* [14]*Department of Medicine, University of Alabama at Birmingham, Alabama, USA. Correspondence should be addressed to R.P.L. (e-mail: richard.lifton@yale.edu).*

(≥3+ proteinuria on urine dipstick) on at least three occasions, or ESRD without other identifiable causes. Relatives ≥40 years with normal urinalysis and no history of renal disease were classified as unaffected; unaffected relatives <40 years were classified as phenotype unknown. The 30 kindreds included 94 affected members, 48 unaffected members and 21 with phenotype unknown owing to young age (Fig. 2). Among affected members, 60 had biopsy-documented IgAN (12 with ESRD), 29 had persistent haematuria/proteinuria (16 with episodic gross haematuria) and 5 had idiopathic ESRD. The mean age of diagnosis was 33 years and the male:female ratio was 1.5:1. All phenotypes were assigned prospectively before initiation of genotyping. The familial clustering in these pedigrees is consistent with multifactorial determination or autosomal dominant transmission with reduced penetrance.

Suggested abnormalities in IgAN have included a maladaptive immune response or altered structure or clearance of IgA (refs 8,9). Consequent candidate genes include the immunoglobulin genes[9,17], the major histocompatibility locus[9,18], the galactosyl transferases[8,19], the IgA Fc receptor[20] (CD89) and uteroglobin[21]. Accordingly, we genotyped polymorphic markers closely linked to these genes. Analysis of the results did not provide evidence of linkage of IgAN to any of these loci under parametric or non-parametric models (Table 1). Because gene-targeting studies in mice have raised uteroglobin as a candidate[21], we also screened the coding sequence of this gene for mutations but identified none. These findings do not implicate any of these candidate loci in familial IgAN.

We next carried out a genome-wide analysis of linkage, genotyping polymorphic markers distributed at 10-cM intervals. To test for a locus

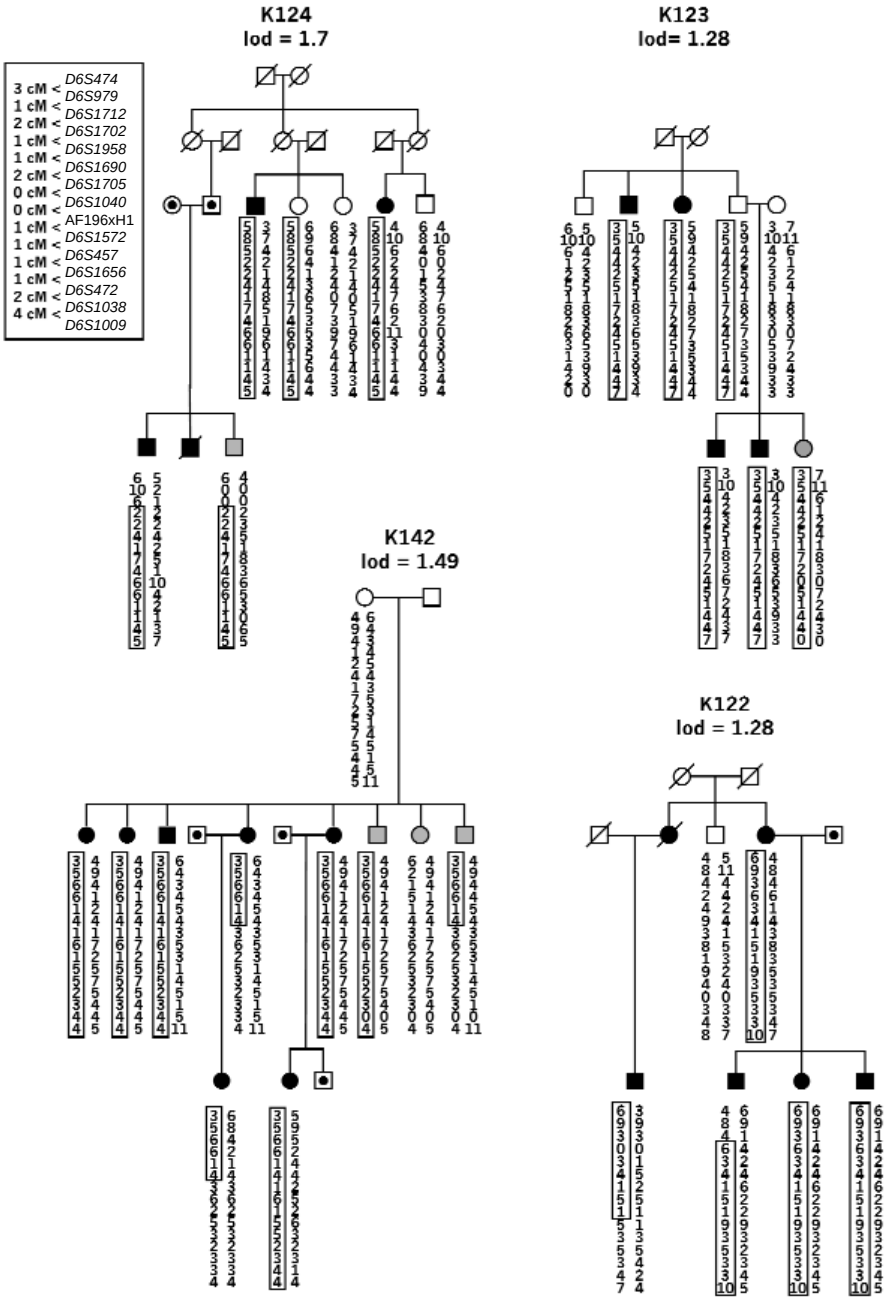

Fig. 2 IgAN kindreds. The structures of 4 of the 30 multiplex IgAN kindreds are shown. Affected individuals are indicated by the filled symbols, unaffected individuals by unfilled symbols and individuals of phenotype unknown (unaffected with age <40) by shaded symbols. Dotted symbols denote individuals who were not studied and were classified as phenotype unknown. Deceased individuals (diagonal lines) are of unknown phenotype unless indicated by filled symbols. Below the symbol for each individual, genotypes for 15 markers at 6q22–23 are shown in their chromosomal order. The identity of each marker and the distance between adjacent loci is indicated in the box at the upper left. Inferred haplotypes shared by relatives with IgAN are indicated by boxes. Below each kindred number, the multipoint lod score for linkage to *IGAN1* at *D6S1040* under the dominant model is shown. Pedigrees and genotypes for remaining kindreds are available (Figs A–D, see http://genetics.nature.com/supplementary_info/).

Table 2 • Pairwise lod scores for linkage of markers at 6q22–23 to IgAN

Locus	Location (cM)	lod (θ)
D6S474	119	1.6 (0.14)
D6S423	120	0.3 (0.23)
D6S979	122	0.2 (0.30)
D6S1712	123	0.1 (0.41)
D6S408	124	0.1 (0.05)
D6S1702	125	3.3 (0.01)
D6S1715	126	0.1 (0.40)
D6S1958	126	1.5 (0.08)
D6S407	127	2.9 (0.05)
D6S1690	127	0.4 (0.22)
D6S1705	129	2.5 (0.04)
D6S1040	129	3.7 (0.00)
AF196xH1	129	3.0 (0.00)
AF162GT1	130	0.9 (0.14)
D6S1572	130	1.4 (0.13)
D6S262	130	0.7 (0.14)
D6S457	131	1.5 (0.14)
D6S1656	132	1.9 (0.14)
D6S472	133	1.1 (0.06)
D6S413	133	0.9 (0.13)
D6S1038	135	1.0 (0.01)
D6S270	137	0.5 (0.15)
D6S1272	137	0.2 (0.40)
D6S1009	139	1.1 (0.10)

Maximum lod scores for linkage of indicated loci to IgAN at indicated recombination fractions (θ) in all 30 kindreds under the dominant model of transmission are shown.

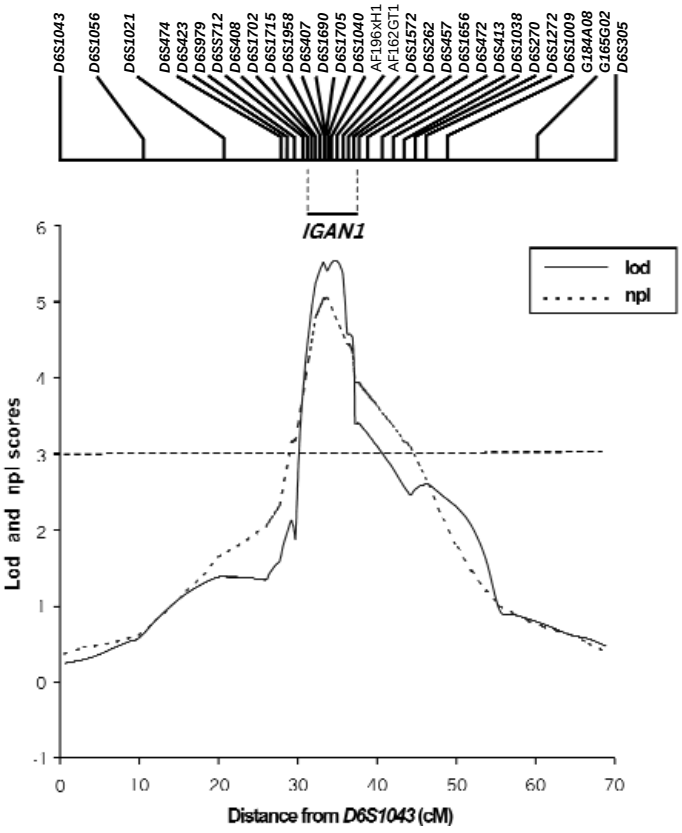

Fig. 3 Linkage of IgAN to 6q22–23. The location of 24 polymorphic marker loci across 6q22–23 that were used in multipoint linkage are shown at the top of the figure. Below, the multipoint lod and NPL scores for linkage of IgAN across this interval in the 30 IgAN kindreds are shown. The multipoint lod score was calculated under the dominant model of transmission with incomplete penetrance allowing for locus heterogeneity. The filled bar indicates the lod-1 support interval for *IGAN1*.

Table 3 • Multipoint lod scores for linkage to *IGAN1* in individual kindreds

Kindred	Multipoint lod at *IGAN1*
124	1.70
142	1.49
122	1.28
123	1.28
107	0.80
131	0.73
104	0.69
100	0.59
160	0.55
102	0.50
129	0.29
103	0.29
108	0.19
171	0.12
176	0.11
111	0.05
133	0.05
174	0.02
121	–0.02
127	–0.15
125	–0.15
126	–0.24
175	–0.25
109	–0.55
136	–0.90
172	–1.00
173	–1.00
128	–1.20
101	–1.30
106	–1.50

The lod score for linkage of IgAN to the *IGAN1* locus under the dominant model of transmission is shown for each kindred studied.

with large effect on the trait, we carried out multipoint analysis of linkage under models of locus homogeneity or heterogeneity, specifying IgAN as an autosomal dominant trait with disease allele frequency of 0.001, phenocopy rate of 0.01 and estimated penetrance of 75%. Analysis identified no interval with lod score greater than 1.0 under locus homogeneity; under models of locus heterogeneity, five intervals yielded initial multipoint lod scores greater than 1.0. Following typing of further markers in each interval, the lod score remained greater than 1.0 in two intervals. One locus, which we now call *IGAN1*, is located at 6q22–23. We genotyped 24 polymorphic markers in the 20-cM interval bounded by loci *D6S474* and *D6S1009*. All yielded positive lod scores in pairwise analysis, with a maximum lod score of 3.7 at *D6S1040* (Table 2). Multipoint linkage including all loci typed on chromosome 6 yielded a maximum lod score of 5.6 at *D6S1040* with an estimated 60% of kindreds linked (Fig. 3 and Table 3). This lod score exceeds the threshold of 3.28 for significant linkage in the setting of locus heterogeneity and has approximately the same significance as a lod score of 5.3 under models of locus homogeneity[22]. The lod -1 support interval places *IGAN1* in the 6.5-cM interval between *D6S1702* and *D6S262*.

This evidence for linkage is robust to alternative analyses. The lod score remains greater than or equal to 5.0 for penetrance estimates ranging from 20% to 85% and for phenocopy rates ranging from 0.03 to 0.0001. The lod score remains 5.2 (α=0.59) in analysis of affected members only. Similarly, non-parametric analysis of allele-sharing among affected individuals achieves a non-parametric lod (NPL) score of 5.1 ($P=4.8\times10^{-6}$; Fig. 3), which is significant at the genome-wide level[23]. Secondary analysis using a more stringent phenotype, classifying as affected only biopsy-documented cases plus those with either episodic gross haematuria or idiopathic ESRD, continues to provide strong evidence of linkage at *IGAN1* (multipoint lod score 5.5, α=0.65). Both Italian (lod score 3.2, α=0.50) and US kindreds (lod score 2.7, α=0.98) contribute and show peak lod scores in the same interval. Locus homogeneity is 1,000-fold less likely than locus heterogeneity, and recessive models of transmission at *IGAN1* are rejected.

The other interval with lod score greater than 1.0 lies at 3p23–24; this interval did not reach statistical significance (lod score 2.8, α=0.33; NPL score 2.9, P=0.003), and this result was not altered by multilocus analysis. Re-analysis of genome-wide data in kindreds partitioned into those with positive or negative lod scores at *IGAN1* did not yield further loci suggestive of linkage.

These results provide highly significant evidence for linkage of IgAN to 6q22–23 under a dominant model with incomplete penetrance and locus heterogeneity. The observed reduced penetrance is compatible with an inherited susceptibility whose expression is dependent on environmental or genetic modifiers. There are no apparent clinical or demographic features that distinguish families with positive or negative lod scores at *IGAN1*. *IGAN1* lies 80 cM distal to the major histocompatibility locus. Examination of sequence databases indicates that the linked interval does not contain any compelling candidates for IgAN. So far, we have found no evidence of linkage disequilibrium between marker alleles at 6q22–23 and IgAN.

Progress in identification of the *IGAN1* gene can proceed on several fronts. Genes identified in this interval can be tested for their role in IgAN and polymorphisms can be used to search for linkage disequilibrium. It is of interest that although IgAN is the most common renal disease among the Zuni native Americans[11], it is rare in individuals of African ancestry[10]. These observations suggest use of linkage disequilibrium and admixture disequilibrium mapping in IgAN cases in these respective populations[24]. Identification of the molecular defect underlying *IGAN1* may provide insight into the pathogenesis, diagnosis and treatment of this cause of renal failure.

Methods

Patients. We ascertained IgAN kindreds through index cases with biopsy-documented disease. Relatives were also examined. Eight Italian kindreds were from the Brescia district of northern Italy and 16 were of diverse geographic origin in Italy. Four kindreds in the US were from Kentucky, one was from Connecticut and one was from Washington State. Eighteen kindreds have been reported[9,12–16]; the relationship of these kindreds to those reported previously is indicated (Table A, see http://genetics.nature.com/supplementary_info/). The study was approved by the Human Investigation Committee at Yale University.

Genetic analysis. We prepared genomic DNA from venous blood of kindred members as described[25]. We initially genotyped 384 highly polymorphic markers distributed at intervals of approximately 10 cM using fluorescent primers to direct PCR from genomic DNA followed by fractionation and data collection using an ABI 377 DNA sequencer[25]. Investigators blinded to affection status scored all genotypes using the Genotyper 1.1.1 software. Gene and marker locations were specified according to databases at the National Center of Biotechnology Information, the Whitehead Institute for Genomic research and the Genome Database web sites.

We carried out analysis of linkage under parametric and non-parametric models, and pairwise and multipoint parametric analysis using LINKAGE (ref. 26) v. 5.0 and GENEHUNTER (ref. 27) v. 2.0, respectively. We specified a dominant model of transmission with reduced penetrance (estimated at 75%), allowing for locus heterogeneity with disease allele frequency of 0.001 and phenocopy rate of 0.01 (in secondary analysis using a more stringent phenotype, the phenocopy rate was accordingly reduced to 0.001). NPL scores were calculated using GENEHUNTER. Intervals showing NPL or lod scores >1 in the initial genome-wide analysis as well as several candidate loci were analysed in further detail by typing additional markers. Allele frequencies at marker loci were specified according to their frequency in the families studied; owing to the high marker density, changing the estimates of allele frequencies had little effect on the resulting multipoint lod scores at *IGAN1*. Likelihood ratios of linkage under models of locus heterogeneity were calculated as LR = αLR (θ_1) + (1–α), where θ_1 is the location of the trait locus and α is the proportion of families linked to θ_1. Multilocus lod scores were calculated as LR = αLR (θ_1) + βLR(θ_2) + (1–α–β), where β corresponds to the proportion of families linked to θ_2.

To search for linkage disequilibrium, we compared the frequencies of independent alleles segregating with *IGAN1* with those found on independent unlinked chromosomes. Italian and US kindreds were analysed separately using the DISEQIN program[28].

To search for mutations in the uteroglobin gene, we used primers located in introns to direct amplification of the coding sequences of this gene by PCR (exon 1, F, 5′–GACGGAACCAGAGACGGGCCAGAGCAT–3′, R, 5′–CTGAGACTCAGCATGCCCAG–3′; exon 2, F, 5′–CTTCTCTCCTCTGTGTTGCA–3′, R, 5′–CTTGGGGAAGCCACTTCTAACCAG–3′; exon 3, F, 5′–TCCTCCTAGAGTTGACTGCAC–3′, R, 5′–GTGGACTCAAAGCATGGCAGCGGC–3′; ref. 29). The products were analysed by single-stranded conformational polymorphism analysis, as described[30].

Acknowledgements

We thank the members of the kindreds for participation; J. Budzinack for database assistance; M. Kashgarian for review of renal pathology; K. Choate for technical assistance; and N. Risch for helpful discussions. A.G.G. was supported by National Institutes of Health training grant and a gift by H. Printz. R.P.L. is an investigator of the Howard Hughes Medical Institute.

Received 14 June; accepted 14 September 2000.

1. U.S. Renal Data System, USRDS 1999 Annual Data Report, National Institute of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda.
2. Maisonneuve, P. *et al*. 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. Varis, J. *et al*. Immunoglobulin and complement deposition in glomeruli of 756 subjects who had committed suicide or met with a violent death. *J. Clin. Pathol.* **46**, 607–610 (1993).
4. D'Amico, G. The commonest glomerulonephritis in the world: IgA nephropathy. *Q. J. Med*. **64**, 709–727 (1987).
5. Julian, B.A., Waldo, F.B., 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).
6. Schena, F.P. A retrospective analysis of the natural history of primary IgA nephropathy worldwide. *Am. J. Med.* **89,** 209–215 (1990).
7. Ibels, L.S. & Gyory, A.Z. IgA nephropathy: analysis of the natural history, important factors in the progression of renal disease, and a review of the literature. *Medicine* **73**, 79–102 (1994).
8. Julian, B.A., Tomana, M., Novak, J. & Mestecky, J. Progress in the pathogenesis of IgA nephropathy. *Adv. Nephrol. Necker Hosp.* **29**, 53–72 (1999).
9. Schena, F.P. Immunogenetic aspects of primary IgA nephropathy. *Kidney Int*. **48**, 1998–2013 (1995).
10. Jennette, J.C, Wall, S.D. & Wilkman, A.S. Low incidence of IgA nephropathy in blacks. *Kidney Int*. **28**, 944–950 (1985).
11. Hoy, W.E., Hughson, M.D., Smith, S.M. & Megill, D.M. Mesangial proliferative glomerulonephritis in southwestern American Indians. *Am. J. Kidney Dis.* **21**, 486–496 (1993).
12. Julian, B.A. *et al.* Familial IgA nephropathy. Evidence of an inherited mechanism of disease. *N. Engl. J. Med.* **312**, 202–208 (1985).
13. Wyatt, R.J. *et al*. Regionalization in hereditary IgA nephropathy. *Am. J. Hum. Genet*. **41**, 36–50 (1987).
14. Scolari, F. *et al*. Familial clustering of IgA nephropathy: further evidence in an Italian population. *Am. J. Kidney Dis.* **33**, 857–865 (1999).
15. Scolari, F. *et al*. Familial occurrence of primary glomerulonephritis: evidence for a role of genetic factors. *Nephrol. Dial. Transplant*. **7**, 587–596 (1992).
16. Schena, F.P., Scivittaro, V. & Ranieri, E. IgA nephropathy: pros and cons for a familial disease. *Contrib. Nephrol*. **104**, 36–45 (1993).
17. Yano, N. *et al*. Polymorphism in the Iα1 germ-line transcript regulatory region and IgA productivity in patients with IgA nephropathy. *J. Immunol.* **160**, 4936–4942 (1998).
18. Fennessy, M. *et al*. HLA-DQ gene polymorphism in primary IgA nephropathy in three European populations. *Kidney Int*. **49**, 477–480 (1996).
19. Allen, A.C., Topham, P.S., Harper, S.J. & Feehally, J. Leukocyte beta 1,3 galactosyltransferase activity in IgA nephropathy. *Nephrol. Dial. Transplant*. **12**, 701–706 (1997).
20. van Zandbergen, G. *et al*. Reduced binding of immunoglobulin A (IgA) from patients with primary IgA nephropathy to the myeloid IgA Fc-receptor, CD89. *Nephrol. Dial. Transplant*. **13**, 3058–3064 (1998).
21. Zheng, F. *et al*. Uteroglobin is essential in preventing immunoglobulin A nephropathy in mice. *Nature Med*. **5**, 1018–1025 (1999).
22. Faraway, J.J. Distribution of the admixture test for the detection of linkage under heterogeneity. *Genet. Epidemiol.* **10**, 75–83 (1993).
23. Lander, E.S. & Kruglyak, L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genet*. **11**, 241–247 (1995).
24. Stephens, J.C., Briscoe, D. & O'Brien, S.J. Mapping by admixture linkage disequilibrium in human populations: limits and guidelines. *Am. J. Hum. Genet.* **55**, 809–824 (1994).
25. Craig, H.D. *et al*. Multilocus linkage identifies two new loci for a mendelian form of stroke, cerebral cavernous malformation, at 7p15–13 and 3q25.2-27. *Hum. Mol. Genet*. **7**, 1851–1858 (1998).
26. Lathrop, G. M. & Lalouel, J.-M. Easy calculations of lod scores and genetic risks on small computers. *Am. J. Hum. Genet*. **36**, 460–465 (1984).
27. Kruglyak, L., Daly, M.J., Reeve-Daly, M.P. & Lander, E.S. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am. J. Hum. Genet*. **58**,1347–1363 (1996).
28. Terwilliger, J.D. A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. *Am. J. Hum. Genet.* **56**, 777–787 (1995).
29. Zhang, Z. *et al*. Human uteroglobin gene: structure, subchromosomal localization, and polymorphism. *DNA Cell. Biol.* **16**, 73–83 (1997).
30. Simon, D.B. *et al*. Gitelman's variant of Bartter's syndrome, inherited hypokalaemic alkalosis, is caused by mutations in the thiazide-sensitive Na-Cl cotransporter. *Nature Genet.* **12**, 24–30 (1996).