Kidney International, Vol. 61 (2002), pp. 1947–1956

COL4A3/COL4A4 mutations: From familial hematuria to autosomal-dominant or recessive Alport syndrome

Ilaria Longo,[1] Paola Porcedda,[1] Francesca Mari, Daniela Giachino, Ilaria Meloni, Carla Deplano, Alfredo Brusco, Maurizio Bosio, Laura Massella, Giancarlo Lavoratti, Dario Roccatello, Giovanni Frascá, Gianna Mazzucco, Andrea Onetti Muda, Maura Conti, Federica Fasciolo, Christelle Arrondel, Laurence Heidet, Alessandra Renieri, and Mario De Marchi

Genetica Medica, Università di Siena, Siena; Dip. Scienze Cliniche e Biologiche and Genetica, Biologia e Biochimica, Università di Torino, Torino; Nefrologia Pediatrica P.O. Magenta, Milano; Ospedale Bambin Gesù, IRCCS, Roma; Ospedale Pediatrico Mayer, Firenze; CMID Ospedale L. Einaudi, Torino; Ospedale Maggiore, Bologna; Anatomia Patologica, Università di Torino, Torino; Anatomia Patologica, Università "La Sapienza," Roma; Nefrologia, Cagliari; and Nefrologia, Ospedale Umberto I, Brescia, Italia; and INSERM U423, Université Renée Descartes, Hôpital Necker-Enfants Malades, Paris, France

***COL4A3/COL4A4* mutations: From familial hematuria to autosomal-dominant or recessive Alport syndrome.**

Background. Mutations of the type IV collagen *COL4A5* gene cause X-linked Alport syndrome (ATS). Mutations of *COL4A3* and *COL4A4* have been reported both in autosomal-recessive and autosomal-dominant ATS, as well as in benign familial hematuria (BFH). In the latter conditions, however, clinical features are less defined, few mutations have been reported, and other genes and non-genetic factors may be involved.

Methods. We analyzed 36 ATS patients for *COL4A3* and *COL4A4* mutations by polymerase chain reaction–single strand conformational polymorphism (PCR-SSCP) and direct sequencing. Sporadic patients who had tested negative for *COL4A5* mutations were included with typical cases of autosomal recessive ATS to secure a better definition of the phenotype spectrum.

Results. We identified seven previously undescribed *COL4A3* mutations: in two genetic compounds and three heterozygotes, and one in *COL4A4*. In agreement with the literature, some of the mutations of compound heterozygotes were associated with microhematuria in healthy heterozygous relatives. The mutations of heterozygous patients are likely dominant, since no change was identified in the second allele even by sequencing, and they are predicted to result in shortened or abnormal chains with a possible dominant-negative effect. In addition, both genes showed rare variants of unclear pathogenicity, and common polymorphisms that are shared in part with other populations.

Conclusions. This study extends the mutation spectrum of *COL4A3* and *COL4A4* genes, and suggests a possible relationship between production of abnormal COL IV chains and dominant expression of a continuous spectrum of phenotypes, from ATS to BFH.

[1]I.L. and P.P. equally contributed to this work.

Key words: benign familial hematuria, collagen IV, inherited nephropathy, X-linked ATS, collagen IV genes.

Received for publication June 20, 2001
and in revised form November 7, 2001
Accepted for publication February 5, 2002

Alport syndrome (ATS) is a progressive clinically and genetically heterogeneous nephropathy often associated with deafness and/or ocular lesions, and characterized by changes of the type IV collagen alpha3/alpha4/alpha5 network of the glomerular basement membrane (GBM). Any of the genes encoding these three chains may be involved. In the common X-linked type (MIM # 301050), Barker et al [1] and several subsequent studies have detected *COL4A5* mutations in 50 to 60% of cases [2, 3; reviewed in 4], and in 85 to 90% with more sensitive techniques [5, 6]. These results indicate that *COL4A5* is the only gene responsible for X-linked ATS. Mutations of either *COL4A3* or *COL4A4* on chromosome 2q have been found in homozygous or compound heterozygous state in autosomal-recessive ATS (MIM # 203780) [7–10], and in heterozygous state in benign familial hematuria (BFH, MIM # 141200) [9–11]. Lastly, a *COL4A3* mutation has recently been identified in a single large family with autosomal-dominant ATS (MIM # 104200) that had been previously linked to the COL4A3/COL4A4 locus [12, 13]. We now report detection of eight novel mutations in six out of 36 Italian families in a comprehensive mutation screening of *COL4A3* and *COL4A4*.

METHODS

Patients

Thirty-six probands were selected who either had a defined or possible ATS diagnosis and negative mutation

Table 1. Clinical features of the 36 probands

	AR and sporadic cases			
	Males	Females	Total	AD probands
N probands	14	16	30	6
Mean age *years*	22.3	18.2	20.1	33.2
ESRD	6	4	10/30	3/6
Mean ESRD age *years*	24.2	24.2	24.2	39
Parents' consanguinity	2	0	2/30	0
Hypoacusia	7	8	15/30	2/6
Proteinuria	9	13	22/30	3/6
Ocular lesions	3	3	6/30	1/6

ESRD is end-stage renal disease.

analysis of *COL4A5* [3] (Table 1). Twelve appeared to have autosomal-recessive ATS, six exhibited apparent autosomal-dominant inheritance with father-to-son transmission, and 18 were sporadic cases. Consanguinity was present in two families. Average age of progression to end-stage renal disease (ESRD) was 24.2 years in those with autosomal-recessive ATS, and 39 years in those with autosomal-dominant ATS. Hearing loss was present in 17 of 36 probands.

Molecular analysis

All the COL4A3 and COL4A4 exons were amplified using primers and polymerase chain reaction (PCR) conditions described by Boye et al for COL4A4 [9] and by Heidet et al for COL4A3 [10]. For exon 39 of COL4A4, the following new primers were used: forward 5′-AGT ATTAATTCTGTTTTCCCCATT-3′; reverse 5′-GGGC AAAGCATGCTACAGCTT-3′. The following internal primers designed on exonic sequences (underlined) were used: *COL4A3* gene, exon 17 reverse 5′-ATTACCT TAATGCCATATTC-3′, exon 25 reverse 5′-TTCACT CCCGGAGTT-3′; *COL4A4* gene, exons 2, 42, 44 are in Boye et al [9]. Silver staining, single-strand conformational polymorphism (SSCP) was performed using either a Genephor apparatus (Pharmacia Amersham, Little Braunschweig, Germany) or vertical polyacrylamide gel electrophoresis (PAGE). All shifts were confirmed on a second PCR product. Direct sequencing of the purified PCR products was performed in both directions (PE Big dye terminator cycle sequencing kit) on an ABI 377A Automated Sequencer and analyzed with the Genescan software. All exons where mutations or rare variants were identified in one patient were determined in ten additional ATS probands and in 50 healthy controls. Reference nucleotide sequences in the GDB database were: COL4A3, 128351; COL4A4, 132673. Nucleotides were numbered starting from the A of the initiator codon, in accordance with the codified nomenclature [14].

RESULTS

Single-strand conformational polymorphism analysis of the 36 probands detected six pathogenic mutations, one in the *COL4A4* and five in the *COL4A3* genes, all in heterozygous state and previously undescribed (Tables 2 and 5). Several rare variants of unclear pathogenic significance and a number of common polymorphisms also were found (Tables 3, 4, 6 and 7). In the six probands with one identified mutation direct sequencing was performed to look for further mutations. This characterized the second allele in two probands (PAR and NCL), who both displayed a mutation in exon 37 of COL4A3. Thus, four mutations were found in two genetic compound probands, both belonging to the group of autosomal-recessives, whereas in four probands (one autosomal-dominant, two sporadic and one adopted), there was only one mutated allele in heterozygous state. Figure 1 shows the phenotype and mutation status of the six families.

COL4A3 changes

Proband PAR is a compound heterozygote for a nonsense mutation (R1037X) and a ATG→TTG change in the start codon, which probably precludes the synthesis of alpha3(IV) chain (Figs. 1 and 2; Table 2). She is a 14-year-old girl with microhematuria, proteinuria, chronic renal failure (CRF; blood creatinine 1.5 mg%) and deafness. At one year of age she had an episode of macrohematuria. Electron microscopy of the renal biopsy performed at age three showed splitting of the GBM, suggesting a diagnosis of ATS. Slow progressive renal failure became apparent in adolescence. The father, now aged 40, carries the R1037X mutation and has persistent microhematuria, but normal renal function (blood creatinine 1.1 mg%). The mother, aged 33, carries the start codon mutation and does not display either microhematuria or other signs of ATS.

Proband NCL is a genetic compound for an in-frame 27 bp deletion in exon 39 (3386del27), causing the loss of nine amino acids in the collagenous domain, and for a frame-shift insertion (3209insA) in exon 37. She is a 17-year-old girl with CRF (blood creatinine 2 mg%), bilateral mixed hearing loss and delayed growth. Microhematuria and proteinuria are reported since the age of five months, and one episode of macrohematuria at seven months. A renal biopsy at age 13 revealed GBM thickening and splitting, suggesting a diagnosis of ATS. One sister aged nine, with microhematuria (blood creatinine 0.4 mg%) and one macrohematuria episode at age two, has the same compound genotype, while a dizygotic co-twin is homozygous for the normal sequence. The father, now aged 43, carries the 3386del27 mutation, shows microhematuria with some macrohematuria episodes, and has normal renal function (blood creatinine 0.9 mg%) and normal audiogram. The mother is a nonsymptomatic carrier of the 3209insA mutation. She does not show microhematuria or proteinuria, although she had a left nephrectomy due to hydronephrosis consequent to stenosis of the ureteropelvic junction. An appar-

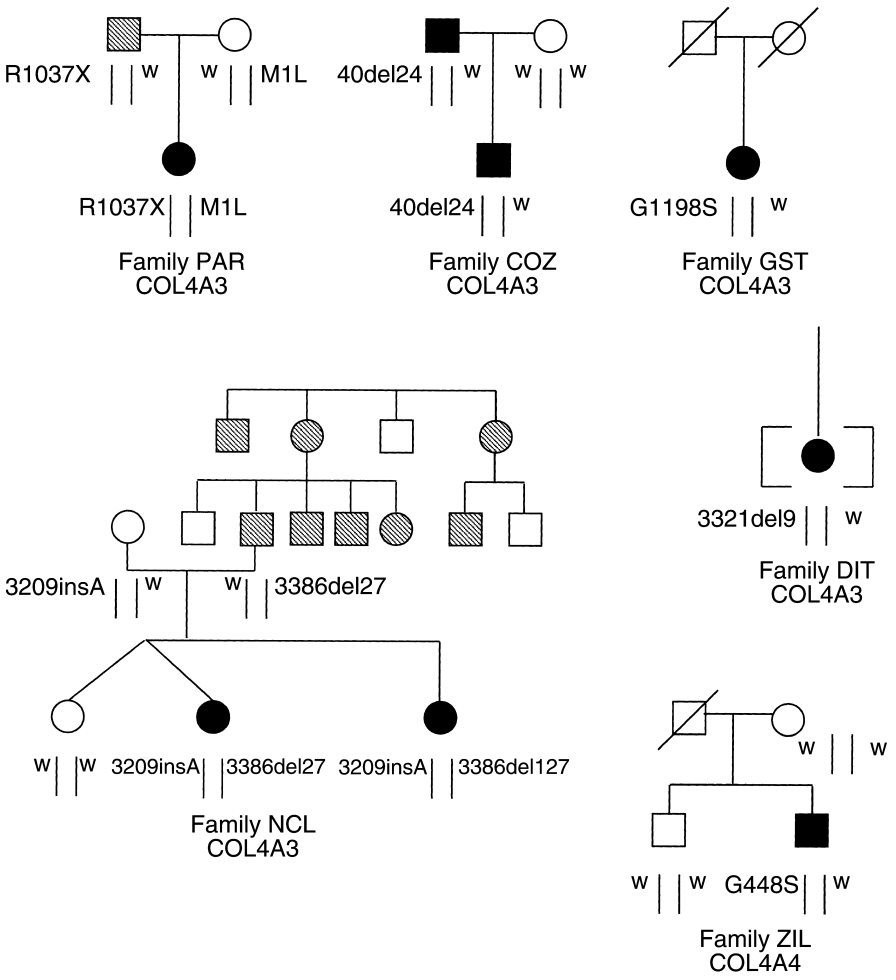

Fig. 1. Families with identified *COL4A3* and *COL4A4* mutations. Mutations are indicated below or above each individual; all affect the *COL4A3* gene, except for G448S in family ZIL, which affects *COL4A4*. Abbreviation is: W, wild type allele. Symbols are: striped symbols, microhematuria only; black symbols, microhematuria plus proteinuria or hypacusis or chronic renal failure.

ently autosomal dominant form of microhematuria segregates in the paternal branch of the pedigree; however, this unfortunately was not available for analysis (Fig. 1).

In proband COZ an in-frame deletion in exon 1 (40del24) caused the loss of eight amino acids in the signal peptide. He is a 14-year-old boy with microhematuria and mild proteinuria (0.2 to 0.8 g/day) since age four. His renal function is presently normal (creatinine clearance 86 mL/min/1.73 m^2) and he has neither ocular signs nor neurosensorial deafness. The same mutation was found in the 46-year-old father, who has had micro- and macrohematuria since childhood. He has suffered from mild bilateral, high-tone and probably job-induced hypacusis since age 39, and has normal renal function. No mutation was identified (by either SSCP or sequencing) in the allele inherited from the mother, who has a normal urinalysis.

Proband GST has a heterozygous missense G1198S change in exon 42. This mutation was absent in a control group of 50 healthy individuals. She is a 39-year-old woman with microhematuria (30 RBC/microscopic field), proteinuria (0.4 g/day) and unimpaired renal function. She began to suffer from a relapsing steroid-resistant nephrotic syndrome at age two, which required immunosuppressive treatment up to the age of eight. Although this clinical course was considered unusual for a COL(IV)-associated disorder, the early onset of proteinuria and persistent microhematuria with non-nephrotic proteinuria at age 28 prompted a renal biopsy to be performed, where electron microscopy revealed irregular thinning of the GBM. The father died at 50 and the mother at 56 with no sign of renal disease. Her 88-year-old aunt had a diagnosis of nephritis but was not biopsied, and no data are available on her relatives.

In proband DIT we identified a nine-nucleotide deletion (3321del9) causing the loss of a Ser-Pro-Gly tripeptide in the collagenous domain. She is a 32-year-old woman who presented with microhematuria and proteinuria (0.7 g/day). A renal biopsy performed at age 31 revealed short segments of thinning of the GBM and segmental glomerulosclerosis. At age 32 she was given angiotensin-converting enzyme (ACE) inhibitors as antiproteinuric treatment. Renal function is presently normal. Since she is adopted, no information is available on her relatives.

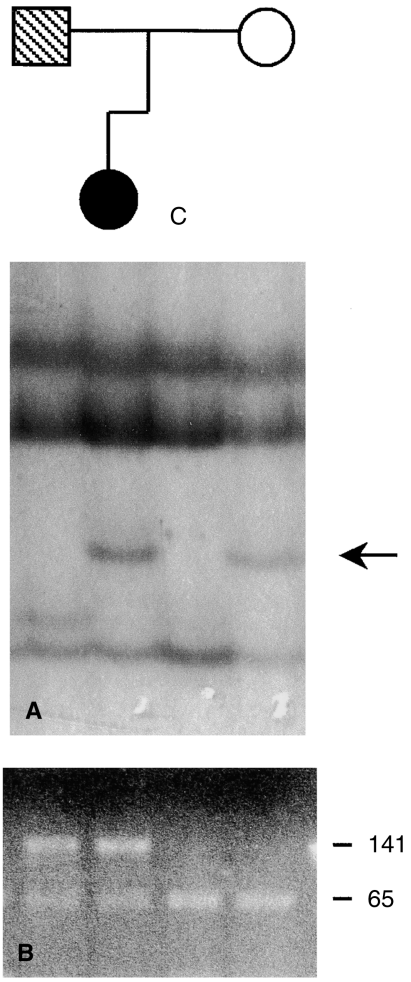

Fig. 2. Mutation inheritance of the compound heterozygous genotype in the PAR family. (*A*) Single-strand conformational polymorphism (SSCP) pattern of exon 1. Arrow indicates the shifted conformer in the proband and the mother. (*B*) Taq1 digestion of exon 37. Fragment size is indicated in bp on the left. The 206 bp polymerase chain reaction (PCR) product is digested by Taq I enzyme in two fragments of 141 and 65 bp. The R1037X mutation causes the loss of the Taq I site.

The three in-frame deletions occurred at sites of direct DNA repeats in the normal COL4A3 sequence. This may have favored a "polymerase slippage" error (not shown). Our identified mutations are likely to have various functional consequences (Fig. 3). The glycine substitution (G1198S) and the two in-frame deletions in collagenous domain (3321del9 and 3386del27) are expected to produce abnormal chains that can be incorporated into abnormal tropocollagen monomers. By contrast, the frameshift (3209insA) and stop codon (R1037X) are null mutations. The same is true for the start codon mutation (M1L), since the next methionine is localized in exon 11, 208 amino acids downstream, and no upstream methionine codon can be found before a stop at −16. Finally, a strongly reduced secretion of the protein is expected for the in-frame deletion (40del24) affecting the signal peptide. With the PSORT program (www.expasy.ch/tools/), an extracellular location of the native protein is correctly predicted with score 0.820, while for the mutant protein this score is reduced to 0.370. For this mutation, therefore, a quantitative effect seems more likely than the qualitative effect predicted for the other in-frame mutations.

Three more COL4A3 missense changes with no clear pathogenic significance were identified (Table 3). Two were private non-glycine substitutions in the triple-helical domain, not found even in the extended series of 46 patients nor in the 50 controls. A third variant, Q1495R, was found in the NC domain. In this case, a glutamine residue conserved in alpha3(IV) chain of several species is replaced by an ancestral arginine shared by alpha1(IV) and alpha5(IV) chains (Fig. 4). This variant was identified in a sporadic patient and in three families where it was present in all affected relatives and some healthy ones as well (data not shown). Two synonymous and one intronic variants were present each in a single patient. Finally, we found eight common polymorphisms (Table 4), including six previously reported in a different population [7, 10]. Five affect non-glycine residues in the triple helical domain, one a glycine as well in the triple helical domain (G43R [10]), one is located in the NC, and one is synonymous.

COL4A4 changes

None of these changes are predicted to be null mutations (Tables 5, 6 and 7). A pathogenic role is likely for the G448S substitution in the triple helical domain found in proband ZIL, but in none of the 50 controls (Figs. 1 and 3). ZIL is a 35-year-old man with microhematuria and proteinuria, who has suffered from non-progressive CRF since age 28. His blood creatinine is presently 1.57 mg% and clearance 100 mL/min. A renal biopsy at age 28 showed non-specific thickening and splitting of the GBM. The mutation causes the loss of a StuI restriction site, but since his father died at 66, it could only be analyzed in his healthy mother and brother. Both tested negative. Haplotype analysis showed that the normal brother received a different paternal allele (not shown).

Three other private missense changes were non-glycine substitutions and their pathogenic role remains to be established (Table 6). R877Q and P1257R affect the triple helical domain. The third substitution, I967V, falls within the eighteenth interruption of the collagenous domain. Thirteen additional COL4A4 variants were considered as polymorphisms. Four of these—two previously characterized by Boye et al [9] and two first reported here—result in amino acid substitutions in the triple helical domain, while six—comprising the GGG→GGA polymorphism at codon 1198 in exon 39 [11]—were synonymous changes. The other three were novel intronic polymorphisms.

Table 2. Mutations in the *COL4A3* gene

Exon	Mutation	Nucleotide change	Effect on coding sequence	Predicted effect on protein	Proband
1	M1L	c.1 A>T	Met→Leu at 1	Unknown[a]	PAR
1	40del24	Deletion of 24 bp at c.40	Deletion of 8 amino acids (from 14) in signal peptide	Reduced production	COZ
37	R1037X	c.3109 C>T	Arg1037→Stop	Termination at 1037	PAR
37	3209insA	Insertion of A at c.3209	Frame-shift	Termination at 1078	NCL
38	3321del9	Deletion of 9 bp at c.3321	Deletion of 3 amino acids (from 1107) in the CD	Chain shortening	DIT
39	3386del27	Deletion of 27 bp at c.3386	Deletion of 9 amino acids (from 1129) in the CD	Chain shortening	NCL
42	G1198S	c.3592 G>A	Gly1198→Ser in the CD	Interruption of continuous Gly-X-Y	GST

CD is collagenous domain.
[a] Likely p0 as there is no alternative downstream AUG codon

Table 3. Rare variants of unknown significance in the *COL4A3* gene

Exon	Variant	Nucleotide change	Effect on coding sequence	Frequency		Reference
				Patients	Controls	
23	G484G	c.1452 G>A	Synonymous Gly 484	1/46	0/50	This study
38	P1109S	c.3325 C>T	Pro1109→Ser in the CD	1/46	0/50	This study
34	P942P	c.2987 C>T	Synonymous Pro 942	1/46	0/50	This study
46	D1347E	c.4041 C>A	Asp1347→Glu in the CD	1/46	0/50	[10]
49	Q1495R	c.4645 A>G	Gln1495→Arg in the NCD	4/46	0/50	This study
48(IVS)	g.4463–11 A>T	g.4463–11 A>T	—	1/46	0/50	This study

Abbreviations are: CD, collagenous domain; NCD, non-collagenous domain.

Table 4. Polymorphisms in the *COL4A3* gene

Exon	Variant	Nucleotide change	Effect on coding sequence	Heterozygosity	Restriction enzyme	Reference
2	G43R	c.127 C>G	Gly→Arg at 43 in the first Gly of CD	0.4	MspI	[10]
7	P141L	c.422 C>T	Pro→Leu at 141 in the CD	0.23	BsrI	This study
9	E162G	c.485 A>G	Glu→Gly at 162 in the CD	0.25	EcoNI	[10]
17	D326Y	c.976 G>T	Asp→Tyr at 976 in the CD	0.1	MboII	[10]
21	L399L	c.1195 T>C	Synonymous Leu 399 in the CD	0.1	—	This study
21	R408H	c.1223 G>A	Arg→His at 408 in the CD	0.1	Fnu4HI	[10]
22	H451R	c.1352 A>G	His→Arg at 451 in the CD	0.1	BstUI	[10]
48	L1474P	c.4421 T>C	Leu→Pro at 1474 in the NCD	0.1	—	[7, 10]

DISCUSSION

Alport syndrome may be caused by mutations in any of three *COL4A3*, *COL4A4*, or *COL4A5* genes encoding tissue-specific collagen IV chains. Since 1990 extensive studies of *COL4A5* in the common X-linked semidominant ATS type have characterized more than 300 mutations and identified the genotype-phenotype correlation [15–18]. More limited studies have implicated *COL4A3* and *COL4A4* in autosomal-recessive [7–10, 19–23] and autosomal-dominant ATS [12, 13], as well as in BFH [9–11]. The wide phenotypic spectrum of collagen IV disorders, the several underlying loci, and mutational heterogeneity make molecular diagnosis difficult, particularly in small families and sporadic cases, where formal genetic analysis fails to suggest what gene to study and what genotype to expect (homozygous/genetic compound vs. heterozygous).

In view of these ascertainment ambiguities, we adopted rather permissive inclusion criteria in the study of autosomal ATS, as in our previous study of X-linked ATS [3]. Together with cases displaying typical features of recessive ATS, that is, rapidly sex-independent progressive nephritis or parental consanguinity [8–10], we included sporadic male and female patients who had tested negative for *COL4A5* mutations. Our two-step analysis of *COL4A3* and *COL4A4* identified many changes, only eight of which could be confidently considered pathogenic. None was in homozygous state, which would have been useful to evaluate their role. In *COL4A3*, four mutations were found in compound heterozygotes and three in simple heterozygotes, as the single *COL4A4* mutation.

In *COL4A3*, the start codon- and nonsense changes (proband PAR) as well as the frame-shift (maternal mutation in NCL) likely result in complete loss of function.

Table 5. Mutations in the *COL4A4* gene

Exon	Mutation	Nucleotide change	Effect on coding sequence	Predicted effect on protein	Proband
20	G448S	c.1550 G>A	Gly→Ser at 448	Interruption of continuous Gly-X-Y	ZIL

Table 6. Rare variants of unknown significance in the *COL4A4* gene

				Frequency	
Exon	Variant	Nucleotide change	Effect on coding sequence	Patients	Controls
14(IVS)	871 + 18 ins T	Insertion of T at g.871 + 18	—	1/46	0/50
30	R877Q	c.2838 G>A	Arg→Gln in the CD	1/46	0/50
32	I967V	c.3107 A>G	Ile→Val in interruption XVIII	1/46	0/50
32	G975G	c.3133 G>A	Synonymous Gly 975	1/46	0/50
40	P1257R	c.3978 C>G	Pro→Arg in the CD	1/46	0/50

All data are from the current study.

Table 7. Polymorphisms in the *COL4A4* gene

Exon	Variant	Nucleotide change	Effect on coding sequence	Heterozygosity	Restriction enzyme	Reference
17 (IVS)	1029 + 36 A>C	c.1029 + 36 A>C	—	0.1	—	This study[c]
18	L365L	c.1303 T>C	Synonymous Leu 365	0.3	MseI	This study
21	P482S	c.1652 C>T	Pro→Ser at 482 in the CD	0.4	HaeIII	[9][a]
23	G545A	c.1842 G>C	Gly→Ala at 545 in the CD	0.1	SphI	[9][b]
28 (IVS)	2592-5 C>T	g.2592-5 C>T	—	0.1	DdeI	[9][a]
33	L1004P	c.3219 T>C	Leu→Pro at 1004 in the CD	0.4	MspI	[9][b]
39	G1198G	c.3802 G>A	Synonymous Gly 1198	0.4	HaeIII	[11]
	K1228K	c.3892 G>A	Synonymous Lys 1228	0.1	—	[9][a]
42	P1360P	c.4288 A>G	Synonymous Pro 1360	0.1	HpaII	[9][a]
	M1327V	c.4190 A>G	Met→Val at 1327 in the CD	0.1	BsrI	[9][a]
47	V1516V	c.4753 A>G	Synonymous Val 1516	0.5	—	[9][a]
46 (IVS)	4731-8T>C	g.4731-8 T>C	—	0.5	—	[9][a]
48	F1644F	c.5140 C>T	Synonymous Phe 1644	0.3	—	This study

[a] Cited by Boye et al and characterized here
[b] Polymorphisms characterized by Boye et al and confirmed in this study
[c] This study, first described here

In agreement with this prediction, the pattern of inheritance in these families is recessive, as mothers of PAR and NCL are asymptomatic, and the father of PAR has isolated microhematuria. A similar mild BFH phenotype in *COL4A3* heterozygotes has been reported by Heidet et al [10]. A different pathogenic mechanism is likely for two in-frame deletions in the collagenous domain (DIT and paternal mutation in NCL), predicted to result in the synthesis of shortened but NC domain-retaining alpha(IV) chains that may be incorporated in the collagen triple helix. This strongly supports their pathogenicity, and even suggests a dominant negative effect. The same mechanism can be ascribed, by analogy with the model of osteogenesis imperfecta [24, 25], to the two missense changes affecting glycine residues in the collagenous domain of the alpha3(IV) chain in GST, and the alpha4(IV) chain in ZIL. Since three of the four parents were already dead, inheritance could not be demonstrated. Even so, both mutations are probably dominant since no change in the second allele was detected by sequenc-

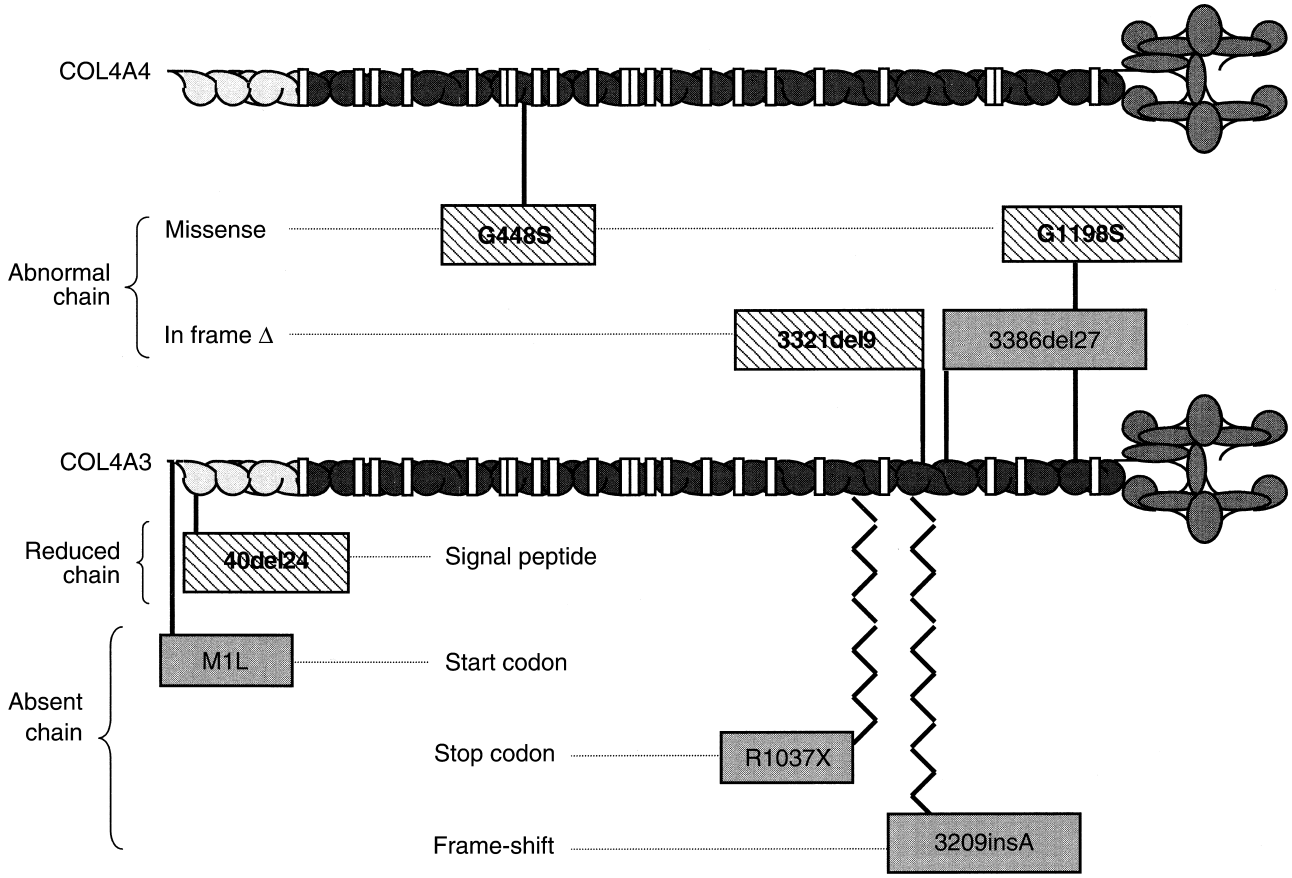

Fig. 3. Distribution of pathogenic mutations along the genes COL4A4 (top) and COL4A3 (bottom). Gray boxes denote mutations found in recessive phenotypes and striped boxes are mutations found in dominant phenotypes.

Amino acids		G	S	C	L	R	–	F	S	T
Homo sapiens	COL4A1	GGC	AGC	TGC	CTG	CGC	AAG	TTC	AGC	ACA
Homo sapiens	COL4A5	GGC	AGC	TGC	CTT	CGT	CGC	TTT	AGT	ACC
		***	***	***	**	*		**	*	**
Q1495R Variant	COL4A3					CGG				
Homo sapiens	COL4A3	GGC	AGC	TGC	CTG	**CAG**	CGA	TTT	ACC	ACA
		***	***	***	**	***	**	**	***	***
Macaca mulatta	COL4A3	GGC	AGC	TGC	CTG	CAG	CGA	TTT	ACC	ACA
Mus musculus	COL4A3	GGC	AGC	TGC	CTG	CAG	CGA	TTC	ACC	ACA
O. cuniculus	COL4A3	GGC	AGC	TGC	CTG	CAG	AGA	TTT	ACC	ACA
C. familiaris	COL4A3	GGC	AGC	TGC	CTG	CAG	CGA	TTT	ACC	ACA
Sus scrofa	COL4A3	GGC	AGC	TGC	CTA	CAG	CGA	TTT	ACC	ACA
Ovis aries	COL4A3	GGC	AGC	TGC	CTG	CAG	AGA	TTT	ACC	ACA
Amino acids		G	S	C	L	Q	R	F	T	T

Fig. 4. Evolutionary conservation at codon 1495 of the alpha3(IV) chain. In the box is the corresponding position in the human *COL4A1* and *COL4A5* genes and *COL4A3* of other species. The Q1495R variant substitutes an ancestral arginine residue for an alpha3(IV)-specific glutamine.

ing all PCR products of the relevant gene. Moreover, a diagnosis of BFH was ruled out in both patients by the presence of proteinuria, which was absent in the null mutation carriers. If our interpretation is correct, proband ZIL is the first example of autosomal-dominant ATS caused by a *COL4A4* mutation. His renal function is still preserved at age 35, in agreement with the very late progression to CRF reported in the family of autosomal-dominant ATS caused by a *COL4A3* mutation [13].

A tentative genotype-phenotype correlation is shown in Figure 3. Mutations are classified in two groups expected to either cause the production of an abnormal alpha-chain or reduce its synthesis, and operationally divided according to whether they are present alone (prob-

ably autosomal dominant), or belong to compound heterozygous genotypes (probably autosomal recessive). In recessive cases, both null- and abnormal-chain mutations are found, whereas the dominant inheritance seems limited to mutations predicted to result in the production of an abnormal chain. This is in agreement with the only reported patient with autosomal-dominant ATS, who had a large in-frame deletion of the alpha 3(IV) chain, resulting from the skipping of exon 21 [13]. Although no other family with a clear autosomal-dominant inheritance and a COL(IV) mutation has been reported, Heidet et al mentioned a father heterozygous for the G297E mutation who reached ESRD at the age of 40 [10].

It should be noted that the phenotype severity of carriers of mutations resulting in abnormal alpha(IV) chains did not clearly follow the biochemical predictions. Thus, the in-frame deletion in the signal peptide (40del24 in family COZ), predicted to strongly reduce secretion of the protein, was associated with a severe phenotype in both the heterozygous father and son, which indicates autosomal dominant inheritance. In contrast with the predicted dominant effect, one of the in-frame deletions (3386del27 in family NCL) was recessively inherited, as shown by the severe phenotype of the genetic compound daughter and the milder one of the heterozygous father. Testing larger series of patients will be necessary to ascertain whether penetrance in heterozygotes indeed may be confined to dominant negative mutations, and whether phenotype severity (as in osteogenesis imperfecta [24, 25]) correlates with the type and position of the change and/or with other interacting factor(s).

In view of these findings, we could reach a positive diagnosis only in six out of the 36 families. Our rate of identified pathogenic mutations (11% of the alleles) is much lower than the 53% reported by Heidet et al [10], though their patients were highly selected.

Identification of few mutations in this study may depend on both the relaxed selection criteria and limited analytical sensitivity, especially critical for autosomal ATS, where the immunohistochemical testing of skin, useful in X-linked ATS [26], is precluded by lack of expression of alpha3 (IV) and alpha4(IV) chains in normal skin.

The fact that direct sequencing revealed two additional mutations overlooked by SSCP demonstrates that limited analytical sensitivity is to some extent responsible. The same limitation also affected our previous COL4A5 screening: in a selected subset of X-linked cases the detection rate increased from <50% obtained by SSCP [3] to about 80% after re-analysis by direct sequencing (Zanelli, Neri and Savi, personal communication). Some kinds of mutation, that is, those in non-coding sequences such as introns and the promoter, will be systematically missed by any exon-scanning approach, while large deletions, which account for 5 to 15% of cases in X-linked ATS [15, 27], will be difficult to identify in heterozygous state and their fraction in *COL4A3/COL4A4* is largely unknown.

The high level of polymorphism in both autosomal genes is an additional difficulty in recognizing mutations by SSCP, which was not encountered in the analysis of *COL4A5* on the X chromosome. Under some circumstances, however, polymorphisms even contributed to the analysis, such as by excluding extended heterozygous deletions. The confounding occurrence of polymorphisms was partly reduced by using internal primers [9], though this excluded part of the coding sequence and some exon-intron junctions.

Several polymorphisms, previously noted but not sequenced in *COL4A4* by Boye et al [9], were further characterized. Of the 13 polymorphisms scattered over the entire gene, three were intronic and six resulted in synonymous changes. For some of them an adverse effect on splicing cannot be ruled out: for example, P1360P in exon 42 may have an influence by replacing the adenine two nucleotides upstream from the splice site. Similarly, the rare G975G variant identified in one patient may affect splicing by introducing a potential AG donor signal within exon 32 (GGGGAA→GGAGAA). The *COL4A3* gene was highly polymorphic as well, but displayed a different pattern: only one polymorphism was synonymous, while most resulted in amino acid substitutions.

Evaluating the pathogenic significance of missense changes was particularly difficult in this study. Substitutions of glycine residues in the collagenous domain of *COL4A5* are considered as pathogenic through a dominant negative mechanism, and are the most common missense mutations in that gene. Even they, however, showed less predictable effects in the autosomal collagen IV genes. Glycine changes in the alpha3(IV) and alpha4(IV) chains have been reported both as causative mutations in autosomal recessive ATS [9, 10] and BFH [11], and as common non-pathogenic polymorphisms: G43R in *COL4A3* [10] and G545A in *COL4A4* [9]. The high frequency of these variants in our healthy controls shows that they should be considered as common polymorphisms. G43R may exert a mild influence on chain flexibility since it affects the first glycine of the collagenous domain.

By contrast, six other COL4A3 variants could not be assumed to be pathogenic on biochemical grounds (Table 3), but were indirectly imputed by their absence in the controls and cosegregation with the disease, or by phylogenetic comparisons. Q1495R alters a glutamine not generally conserved, but apparently specific for the alpha3(IV) chain (Fig. 4), where it may have a particular structural significance related to the nearby Goodpasture epitope [28]. Its distribution in families also points to its partially penetrant pathogenicity. The two private non-glycine substitutions P1109S and D1347E affect conserved residues in the triple helical region. D1347E has been

reported in homozygous state in a French ATS patient [10]. Our finding that this change is ATS-associated in a different population further supports its pathogenicity.

If some of these variants, or even of the polymorphisms, are really pathogenic, our consideration of only the overtly pathogenic mutations would underestimate the role of *COL4A3* and *COL4A4*. Collagen IV rare variants or polymorphisms, both quantitative and qualitative, could perhaps modulate disease severity in the presence of a specific mutation in one of the other collagen IV chains, as in the case of the *CFTR* gene, where an intronic 5T polymorphism is able to modifies a complex spectrum of phenotypes [29]. Even changes in other structural proteins of the glomerular podocyte or slit diaphragm, including nephrin, podocin and β-actin [30], may interact with the COL(IV) network and affect the severity of GBM lesions. The role of other genes in BFH is supported by the lack of linkage to the *COL4A3/COL4A4* locus in different studies [31, 32].

Our results show that there are no clear-cut boundaries between the diagnoses of autosomal-dominant ATS, recessive ATS and BFH, but that these phenotypes belong to a continuous spectrum of graded severity. "Collagen IV diseases" would be a more appropriate expression. This is relevant both for classification and prognosis, since shifting the predicted effect of a mutation from recessive to dominant dramatically affects genetic counseling and risk ascertainment, and makes it very difficult to predict inheritance. In conclusion, molecular analysis on these collagen genes aids in diagnosis, but is of limited value in predicting prognosis, which is of major importance for patients and clinicians. More extensive study in type IV collagen and other interacting molecules is needed to take on this challenge.

ACKNOWLEDGMENTS

This study was supported by grants cofin98 (to MDM) and Italian Telethon (to AR). We are indebted to our patients and their families, and to our colleagues who provided clinical information.

Reprint requests to Dr. Ilaria Longo, Genetica Medica, Dipartimento di Biologia molecolare, Università degli Studi di Siena, Nuovo Policlinico Le Scotte, Viale Bracci 2, 53100 Siena, Italy.
E-mail: longo@unisi.it

APPENDIX

Abbreviations used in this article are: ATS, Alport syndrome; BFH, benign familial hematuria; CRF, chronic renal failure; GBM, glomerular basement membrane; ESRD, end-stage renal disease, PCR, polymerase chain reaction, RBC, red blood cell; SSCP, single-strand conformational polymorphism.

REFERENCES

1. Barker DF, Hostikka S, L, Zhou J, *et al*: Identification of mutations in the COL4A5 collagen gene in Alport syndrome. *Science* 248:1224–1227, 1990
2. Knebelmann B, Forestier L, Drouot L, *et al*: Splice-mediated insertion of an Alu sequence in the COL4A3 mRNA causing autosomal recessive Alport syndrome. *Hum Mol Genet* 4:675–679, 1995
3. Renieri A, Bruttini M, Galli L, *et al*: X-linked Alport syndrome: An SSCP-based mutation survey over all 51 exons of the COL4A5 gene. *Am J Hum Genet* 58:1192–1204, 1996
4. Lemmink HH, Schroder CH, Monnens LA, Smeets HJ: The clinical spectrum of type IV collagen mutations. *Hum Mutat* 9:477–499, 1997
5. Martin P, Heiskari N, Zhou J, *et al*: High mutation detection rate in the COL4A5 collagen gene in suspected Alport syndrome using PCR and direct DNA sequencing. *J Am Soc Nephrol* 9:2291–2301, 1998
6. Inoue Y, Nishio H, Shirikawa T, *et al*: Detection of mutations in the COL4A5 gene in over 90% of male patients with X-linked Alport's syndrome by RT-PCR and direct sequencing. *Am J Kidney Dis* 34:854–862, 1999
7. Lemmink HH, Mochizuki T, van den Heuvel LP, *et al*: Mutations in the type IV collagen alpha 3 (COL4A3) gene in autosomal recessive Alport syndrome. *Hum Mol Genet* 3:1269–1273, 1994
8. Mochizuki T, Lemmink HH, Mariyama M, *et al*: Identification of mutations in the alpha 3(IV) and alpha 4(IV) collagen genes in autosomal recessive Alport syndrome. *Nat Genet* 8:77–81, 1994
9. Boye E, Mollet G, Forestier L, *et al*: Determination of the genomic structure of the COL4A4 gene and of novel mutations causing autosomal recessive Alport syndrome. *Am J Hum Genet* 63:1329–1340, 1998
10. Heidet L, Arrondel C, Forestier L, *et al*: Structure of the human type IV collagen gene COL4A3 and mutations in autosomal Alport syndrome. *J Am Soc Nephrol* 12:97–106, 2001
11. Lemmink HH, Nillesen WN, Mochizuki T, *et al*: Benign familial hematuria due to mutation of the type IV collagen alpha4 gene. *J Clin Invest* 98:1114–1118, 1996
12. Jefferson JA, Lemmink HH, Hughes AE, *et al*: Autosomal dominant Alport syndrome linked to the type IV collage alpha 3 and alpha 4 genes (COL4A3 and COL4A4). *Nephrol Dial Transplant* 12:1595–1599, 1997
13. van Der Loop FT, Heidet L, Timmer ED, *et al*: Autosomal dominant Alport syndrome caused by a COL4A3 splice site mutation. *Kidney Int* 58:1870–1875, 2000
14. den Dunnen JT, Antonorakis SE: Mutation nomenclature extensions and suggestions to describe complex mutations: A discussion. *Hum Mutat* 15:7–12, 2000
15. Antignac C, Knebelmann B, Drouot L, *et al*: Deletions in the COL4A5 collagen gene in X-linked Alport syndrome. Characterization of the pathological transcripts in nonrenal cells and correlation with disease expression. *J Clin Invest* 93:1195–1207, 1994
16. Mazzucco G, Barsotti P, Onetti Muda A, *et al*: Ultrastructural and immunohistochemical findings in Alport's syndrome: A study of 108 patients from 97 Italian families with particular emphasis on COL4A5 gene mutation correlations. *J Am Soc Nephrol* 9:1023–1031, 1998
17. Jais JP, Knebelmann B, Giatras I, *et al*: X-linked Alport syndrome: Natural history in 195 families and genotype-phenotype correlations in males. *J Am Soc Nephrol* 17:649–657, 2000
18. Barsotti P, Onetti Muda A, Mazzucco G, *et al:* Distribution of alpha-chains of type IV collagen in glomerular basement membranes with ultrastructural alterations suggestive of Alport syndrome. *Nephrol Dial Transplant* 16:945–952, 2001
19. Gübler MC, Knebelmann B, Beziau A, *et al*: Autosomal recessive Alport syndrome: Immunohistochemical study of type IV collagen chain distribution. *Kidney Int* 47:1142–1147, 1995
20. Ding J, Stitzel J, Berry P, *et al*: Autosomal recessive Alport syndrome: Mutation in the COL4A3 gene in a woman with Alport syndrome and posttransplant antiglomerular basement membrane nephritis. *J Am Soc Nephrol* 5:1714–1717, 1995
21. Kalluri R, van den Heuvel LP, Smeets HJ, *et al*: A COL4A3 gene mutation and post-transplant anti-alpha 3(IV) collagen alloantibodies in Alport syndrome. *Kidney Int* 47:1199–1204, 1995
22. Finielz P, Cartault F, Chuet C, Genin R: Alport syndrome in Reunion Island: Phenotypic heterogeneity of the recessive-autosomal form. *Nephron* 79:237, 1998
23. Nomura S, Naito I, Fukushima T, *et al*: Molecular genetic and

immunohistochemical study of autosomal recessive Alport's syndrome. *Am J Kidney Dis* 31:E4, 1998
24. Byers PH: Brittle bones–fragile molecules: disorders of collagen gene structure and expression. *Trends Genet* 6:293–300, 1990
25. Wang Q, Orrison BM, Marini JC: Two additional cases of osteogenesis imperfecta with substitutions for glycine in the alpha 2(I) collagen chain. A regional model relating mutation location with phenotype. *J Biol Chem* 268:25162–25167, 1993
26. Kashtan CE, Michael AF: Alport syndrome. *Kidney Int* 50:1445–1463, 1996
27. Renieri A, Galli L, Grillo A, *et al*: Major COL4A5 gene rearrangements in patients with juvenile type Alport syndrome. *Am J Med Genet* 59:380–385, 1995
28. Netzer KO, Leinonen A, Boutard A, *et al*: The Goodpasture autoantigen. Mapping the major conformational epitopes of alpha3 (IV) collagen to residues 17-31 and 127-141 of the NC1 domain. *J Biol Chem* 274:11267–11274, 1999
29. Kerem B, Kerem E: The molecular basis for disease variability in cystic fibrosis. *Eur J Hum Genet* 4:65–73, 1996
30. Somlo S, Mundel P: Getting a foothold in nephrotic syndrome. *Nat Genet* 24:333–335, 2000
31. Yamazaki H, Nakagawa Y, Saito A, *et al*: No linkage to COL4A3 locus in Japanese thin basement membrane disease families. *Nephrol* 1:315–321, 1995
32. Piccini M, Casari G, Zhou J, *et al*: Evidence for genetic heterogeneity in benign familial hematuria. *Am J Nephrol* 19:464–467, 1999