U.PORTO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

ALPORT SYNDROME: CLINICAL AND MOLECULAR STUDY OF PORTUGUESE FAMILIES

MARIA JOÃO NABAIS SÁ

Tese de doutoramento em Patologia e Genética Molecular

2013

MARIA JOÃO NABAIS SÁ

ALPORT SYNDROME:

CLINICAL AND MOLECULAR STUDY OF PORTUGUESE FAMILIES

Tese de Candidatura ao grau de Doutor em Patologia e Genética Molecular submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientador – Doutor João Paulo Ferreira da Silva Oliveira

Categoria – Professor Associado

Afiliação – Faculdade de Medicina da Universidade do Porto

Co-orientador – Doutora Filipa Abreu Gomes de Carvalho

Categoria – Professora Auxiliar com Agregação

Afiliação – Faculdade de Medicina da Universidade do Porto

Apoios financeiros da Fundação para a Ciência e a Tecnologia (SFRH/SINTD/60054/2009; HMSP-ICJ/SAU-ICT/0003/2009; Unidade de Investigação e Desenvolvimento em Nefrologia – FCT-725) no âmbito do III Quadro Comunitário de Apoio e da Sociedade Portuguesa de Nefrologia através de bolsa de financiamento de Projecto de Investigação Científica de 2011.

Aos meus pais e irmão.

“Basic science could both inform and be informed by clinical genetics.”

Charles Joseph Epstein, 2011

A todos os que contribuem diariamente para esta ponte.

Agradecimentos

Acknowledgements

Aos doentes e familiares, que são “o vento que faz mover as minhas asas”.

Ao Professor Doutor João Paulo Oliveira e à Professora Doutora Filipa Carvalho, por supervisionarem o trabalho desta tese e por continuamente desenvolverem a minha autonomia e estimularem uma colaboração estreita com médicos e grupos de investigação dedicados às áreas científicas versadas desta tese. Ao Professor Doutor Alberto Barros, Diretor do Departamento de Genética da Faculdade de Medicina da Universidade do Porto (FMUP), por ter acolhido parte do trabalho experimental desta tese.

A todos os médicos que deram voz aos doentes neste projecto, em particular à Drª. Susana Sampaio, à Drª. Fernanda Carvalho, à Professora Doutora Ângela Carneiro e ao Dr. Sérgio Silva Estrela. À NephroCare – Portugal, por toda a colaboração. Ao Professor Doutor Manuel Pestana, Diretor do Serviço de Nefrologia do Hospital São João (HSJ) e Director da Unidade de I&D de Nefrologia e Doenças Infecciosas – INEB (I3S) da Universidade do Porto, pelo apoio incondicional.

A todos quem, no HSJ, nomeadamente nos Serviços de Arquivo Clínico, de Nefrologia, de Anatomia Patológica, de Oftalmologia, de Otorrinolaringologia e da consulta externa de Genética, assim como na FMUP, incluindo as Secretarias, a Biblioteca, o Laboratório de Imunohemoterapia, os Departamentos de Informática, de Estatística e de Genética, tão bem conhecem a importância do seu contributo para este projecto. À Susana Alves e Ana Paula Neto, por toda a dedicação.

À FCT, ao Programa Harvard Medical School-Portugal, à Sociedade Portuguesa de Nefrologia, ao Departamento de Genética e à Unidade de I&D de Nefrologia e Doenças Infecciosas – INEB (I3S) da Universidade do Porto, pela confiança depositada, assim como pela divulgação e financiamento do projecto.

Agradecimentos

Acknowledgements

À Professora Doutora Graça Porto, impulsionadora do desafio.

À Drª. Ana Maria Fortuna, Directora da Unidade de Genética Médica do Centro de Genética Médica Doutor Jacinto Magalhães/Centro Hospitalar do Porto, e à Drª. Gabriela Soares, Orientadora do Internato em Genética Médica, por confiarem no percurso pioneiro que escolhi.

À Professora Doutora Maria do Carmo Fonseca, por me ensinar a construir o caminho, do sonho à sua concretização.

To Guy Froyen and Nathalie Firemans, for their genuine collaboration.

To Frances Flinter, who inspires me and shows me that better has no limit. To Helen Storey and Mato Nagel, who committed their efforts with the aim of this project and continue to persue it, in a global perspective.

A todos os meus amigos, cuja coragem sempre espelharam.

To Arjan de Brouwer, for teaching me that treasure is everywhere.

À minha família, a quem sempre me acompanha e a quem já partiu para seguir com outro olhar esta jornada.

Resumo A análise genética molecular dos genes *COL4A5*, *COL4A4* e *COL4A3* genes é a melhor abordagem não invasiva disponível para estabelecer o diagnóstico, tratamento, prognóstico e aconselhamento genético de nefropatias associadas ao colagéneo tipo IV – síndrome de Alport e hematúria familiar benigna / nefropatia de membrana basal fina. Neste estudo multicêntrico nacional, foram estudadas 65 famílias não relacionadas com o diagnóstico clínico de síndrome de Alport ou de nefropatia de membrana basal fina. Foram identificadas mutações patogénicas no gene *COL4A5* em 22 famílias (22/60; 37%), que confirmaram o diagnóstico de síndrome de Alport ligada ao cromossoma X. Foi identificada uma nova correlação genótipo-fenótipo, dado ter sido detectada uma deleção que envolve o gene *COL4A5*, mas não o gene *COL4A6*, numa família com leiomiomatose difusa associada a síndrome de Alport ligada ao cromossoma X. Este resultado sugere que a deleção dos exões 5' do gene *COL4A6* e da região promotora comum aos genes *COL4A5_COL4A6* não é essencial para o desenvolvimento de leiomiomatose difusa nestes doentes. Foi efetuada sequenciação directa dos genes *COL4A3* e *COL4A4* em cinco famílias com padrão de hereditariedade autossómico, como abordagem de primeira linha, e em 35 famílias com padrão de hereditariedade incerto e nas quais não foram detectadas mutações patogénicas no gene *COL4A5*. Foram detectadas mutações patogénicas nos genes *COL4A3* ou *COL4A4* em 25 famílias (25/40; 63%), confirmando o diagnóstico de síndrome de Alport autossómica ou de nefropatia de membrana basal fina. Observou-se uma proporção semelhante de famílias com síndrome de Alport ligado ao cromossoma X e de famílias com síndrome de Alport autossómica ou com nefropatia de membrana basal fina, sendo as prevalências de 37% (22/62) e de 40% (25/62), respetivamente. No seu conjunto estes resultados sugerem que a abordagem laboratorial por método de sequenciação de ADN de próxima geração será mais eficaz e menos dispendiosa para diagnóstico genético molecular da síndrome de Alport na população portuguesa.

Palavras-chave: Síndrome de Alport; *COL4A3*; *COL4A4*; *COL4A5***.**

Abstract Molecular genetic analysis of the *COL4A5*, *COL4A4* and *COL4A3* genes is the best non-invasive available approach for diagnosis, treatment, prognosis, and genetic couseling of collagen IV-related nephropathies – Alport syndrome and thin basement membrane nephropathy. In this national multicenter study, 65 unrelated families with a clinical diagnosis of Alport syndrome or thin basement membrane nephropathy were studied. Pathogenic *COL4A5* mutations were identified by direct sequencing and *multiplex-ligation dependent amplification* in 22 families (22/65; 34%), confirming a diagnosis of X-linked AS. A novel genotype-phenotype correlation was disclosed, as a deletion involving *COL4A5*, but not *COL4A6*, was detected in a family with diffuse leiomyomatosis associated with X-linked AS. This result suggests that the deletion of the 5' exons of *COL4A6* and of the common *COL4A5_COL4A6* promoter region is not essential for the development of diffuse leiomyomatosis in patients with Alport syndrome and diffuse leiomyomatosis. *COL4A3* and *COL4A4* were analyzed by direct sequencing in five families with autosomal inheritance pattern, as a first-tier approach, and in 35 families with an unclear inheritance pattern and no pathogenic mutations in the *COL4A5* gene. Pathogenic mutations in *COL4A3* or *COL4A4* were detected in 25 unrelated families (25/40; 63%), confirming the diagnosis of autosomal Alport syndrome or thin basement membrane nephropathy. A similar proportion of families with X-linked Alport syndrome and of families with autosomal Alport syndrome or with thin basement membrane nephropathy was observed, with prevalences of 37% (22/62) and 40% (25/62), respectively. Altogether, these results suggest that Next-Generation Sequencing would be a more effective and less expensive approach to the genetic molecular study of Alport syndrome in the Portuguese population.

Keywords: Alport syndrome; *COL4A3*; *COL4A4*; *COL4A5*.

Publicações

Publications

Artigos em revistas internacionais

Papers in international journals

Nabais Sá MJ, Fieremans N, de Brouwer AP, Sousa R, Teixeira E Costa F, Brito MJ, Carvalho F, Rodrigues M, Teixeira de Sousa F, Felgueiras J, Neves F, Carvalho A, Ramos U, Vizcaíno JR, Alves S, Carvalho F, Froyen G, Oliveira JP. "Deletion of the 5'exons of *COL4A6* is not needed for the development of diffuse leiomyomatosis in patients with Alport syndrome." J Med Genet. 2013 Nov;**50**(11):745-53.

Nabais Sá MJ, Sampaio S, Oliveira A, Alves S, Pinto de Moura C, Estrela Silva S, Castro R, Augusto Araújo J, Rodrigues M, Neves F, Seabra J, Soares C, Gaspar MA, Tavares I, Freitas L, Couto Sousa T, Castro Henriques A, Teixeira e Costa F, Morgado E, Teixeira Sousa F, Pratas e Sousa J, Gomes da Costa A, Filipe R, Garrido J, Montalban J, Ponce P, Alves R, Faria B, Carvalho MF, Pestana M, Carvalho F, Oliveira JP. "Collagen type IV-related nephropathies in Portugal: spectrum of pathogenic *COL4A5* mutations and clinical characterization of 22 families." Manuscript submitted to Journal of Medical Genetics (Manuscript ID: jmedgenet-2014-102471).

Nabais Sá MJ, Storey H, Flinter F, Nagel M, Sampaio S, Castro R, Augusto Araújo J, Gaspar MA, Soares C, Oliveira A, Castro Henriques A, Gomes da Costa A, Pinto Abreu C, Ponce P, Alves R, Pinho L, Estrela Silva S, Pinto de Moura C, Mendonça L, Carvalho MF, Pestana M, Alves S, Carvalho F, Oliveira JP. "Collagen type IV-related nephropathies in Portugal: spectrum of pathogenic *COL4A3* and *COL4A4* mutations and clinical characterisation of 25 families." Manuscript submitted to Journal of Medical Genetics (Manuscript ID: jmedgenet-2014-102472).

Comunicacões em encontros cientificos internacionais

Communications in international meetings

Maria João Nabais Sá, Susana Alves, Helen Storey, Frances Flinter, Patrick J Willems,Filipa Carvalho, João P Oliveira on behalf of the investigators of the Portuguese Alport syndrome Collaborative Study. Molecular characterization of 65 Portuguese families with clinical diagnosis of collagen type IV glomerulopathies. 50th ERA-EDTA Congress, 18th-21st May 2013, Istanbul, Turkey. (Poster Presentation)

Maria João Nabais Sá, Rita Sousa, Fernando Teixeira e Costa, Fernanda Carvalho, Susana Alves, Filipa Carvalho, João Paulo Oliveira. A large deletion confined to COL4A5 causes Alport syndrome and diffuse leiomyomatosis. American Society of Human Genetics 62nd Annual Meeting, 6-10th November 2012, San Francisco, U.S.A. (Poster Presentation)

Maria João Nabais Sá, Susana Alves, Filipa Carvalho, João Paulo Oliveira on behalf of the Portuguese Alport Syndrome Collaborative Study. *COL4A5* mutational analysis of 51 unrelated Portuguese patients with Alport syndrome – preliminary report. European Society of Human Genetics 2012, 23rd-26th June 2012, Nürnberg, Germany. (Poster Presentation)

Comunicacões em encontros cientificos nacionais

Communications in national meetings

Maria João Nabais Sá, Susana Alves, Filipa Carvalho, João Paulo Oliveira on behalf of the Portuguese Alport Syndrome Collaborative Study. Alport syndrome: clinical and molecular study of Portuguese families – preliminary results. Encontro Renal 2012, 28th-31rst March 2012, Vilamoura, Portugal. (Poster presentation)

Maria João Nabais Sá, Susana Sampaio, Ana Oliveira, Isabel Tavares, Ana Paula Neto, Susana Alves, Filipa Carvalho, João Paulo Oliveira. X-linked Alport syndrome: a novel mutation shared by two apparently unrelated Portuguese families. 15ª Reunião da Sociedade Portuguesa de Genética Humana, 10th-12th November 2011, Lisboa, Portugal. (Poster presentation)

Table of Contents

Abbreviations

ACE: Angiotensin Converting Enzyme

ADAS: Autosomal Dominant Alport Syndrome

α1(IV) chain: alpha-1 chain of collagen type IV

α2(IV) chain: alpha-2 chain of collagen type IV

α3(IV) chain: alpha-3 chain of collagen type IV

α4(IV) chain: alpha-4 chain of collagen type IV

α5(IV) chain: alpha-5 chain of collagen type IV

α6(IV) chain: alpha-6 chain of collagen type IV

Anti-GBM: Anti-Glomerular Basement Membrane

ARAS: Autosomal Recessive Alport Syndrome

ARB: Angiotensin Receptor Blocker

AS: Alport Syndrome

ATS-DL: Alport Syndrome and Diffuse Leiomyomatosis

BFH: Benign Familial Hematuria

BM: Basement Membrane

CFHR5: Complement Factor H-related Protein 5

CKD: Chronic Kidney Disease

CRF: Chronic Renal Failure

DNA: Deoxyribonucleic Acid

EBM: Epidermal Basement Membrane

ECASCA: European Community Alport Syndrome Concerted Action

ESRD: End-Stage Renal Disease

GBM: Glomerular Basement Membrane

GFND: Glomerulopathy with Fibronectin Deposits

GFR: Glomerular Filtration Rate

HL: Hearing Loss

Hz: Hertz (audiometric measurement)

MPS: Massive Parallel Sequencing

NGS: Next-Generation Sequencing

OCT: Optical Coherence Tomography

pCr: Plasma Creatinine Measurement

PCR: Polymerase Chain Reaction

RBC: Red Blood Cell

RCE: Recurrent Corneal Erosion

RNA: Ribonucleic Acid

RRT: Renal Replacement Therapy

SNHL: Sensorineural Hearing Loss

TBMN: Thin Basement Membrane Nephropathy

UTR: Untranslated Region

XLAS: X-linked Alport Syndrome

Chapter 1 Introduction

1. Alport syndrome

1.1. Diagnosis: Historical perspective

In 1927, Arthur Cecil Alport recognized that the association of hematuria and sensorineural deafness, reported in a previously studied family, belong to a unique clinical entity (Alport 1927), which in time became known as Alport syndrome (AS) (Williamson 1961). This family had been the subject of several successive observations (Guthrie 1902; Kendall 1912; Hurst 1923; Alport 1927). Guthrie described the familial occurrence of hematuria of unknown cause, which he called hereditary hematuria, since it was observed in two consecutive generations of the family and it was congenital or presented during childhood (Guthrie 1902). Hematuria occurred in both genders, was microscopic and/or macroscopic, was persistent or recurrent and was prone to paroxysmal exacerbations concomitantly with flue-like symptoms. Phenotypic heterogeneity was emphasized by reporting that affected males manifested uremia and died at a younger age, while females were usually mildly affected (Hurst 1923; Eason J 1924; Alport 1927). Eason mentioned that, in this family, a non-affected female did not transmit the disorder, while affected females and another non-affected female did; and that infections by non-haemolytic streptococcus predisposed to the exacerbation of the kidney disease (Eason J 1924). Hurst had noted the occurrence of deafness in three patients of the family, but he did not make a comment on this (Hurst 1923). However, Authur C. Alport who, observing “nerve” deafness in both affected males and females, recognized auditory loss as an integral sign of this hereditary disease (Alport 1927). Perkoff and colleagues published a detailed study of a much larger Mormon family from Utah with similar clinical features and hypothesized for the first time a mechanism of inheritance for the disease – a partially sex-linked dominant inheritance with crossing-over between X and Y chromosomes, which was subsequently contested (Perkoff, Stephens et al. 1951; Stephens, Perkoff et al. 1951; Perkoff, Nugent et al. 1958). Ocular manifestations were first reported as part of the syndrome in 1954

(Reyersbach and Butler 1954; Sohar 1954; Sohar 1956). As there were no specific diagnostic criteria for the diagnosis of AS until 1988, a variety of clinically heterogeneous inherited nephritis were included in the AS phenotype for many years and a multiplicity of hypotheses were posed regarding the inheritance of AS (Crawfurd 1988).

The pathophysiology of AS was clarified by the observation of characteristic ultrastructural changes in the glomerular basement membrane (GBM) of the kidney (Kinoshita, Osawa et al. 1969; Hinglais, Grunfeld et al. 1972; Spear and Slusser 1972; Churg and Sherman 1973). In the late 60s, Kefalides had identified a unique collagen in the GBM (Kefalides 1966; Kefalides 1968). Three identical α-chains were suggested to form the scaffold of the GBM (Kefalides 1971). With advances in electron microscopy in the early 1970s, several laboratories reported typical ultrastructural changes in the GBM of AS patients, namely focal or diffuse thinning, thickening and splitting of the GBM containing granular material (Kinoshita, Osawa et al. 1969; Hinglais, Grunfeld et al. 1972; Spear and Slusser 1972; Churg and Sherman 1973). These findings lead Spear to hypothesize that mutations in the gene for this unique structural GBM component could cause AS (Spear 1973). Two additional observations were fundamental to the recognition of a possible type IV collagen defect as the cause of AS: (i) the Goodpasture (human anti-GBM) antigen, which was later known to be directed against the globular domain of collagen type IV (Wieslander, Barr et al. 1984), did not bind GBM of patients with AS (Olson, Anand et al. 1980; McCoy, Johnson et al. 1982); and (ii) the development of post-transplant glomerulonephritis in AS patients was due to the synthesis of antibodies directed against collagen type IV chain epitopes (Milliner, Pierides et al. 1982; Kashtan, Fish et al. 1986), more specifically antibodies directed against α3 and α5 collagen type IV chains (Fleming, Savage et al. 1988; Kleppel, Kashtan et al. 1989; Kleppel, Santi et al. 1989; Savage, Noel et al. 1989; vd Heuvel, Schroder et al. 1989; Kashtan, Butkowski et al. 1990).

Framing the reported findings, Flinter and colleagues (1988) proposed four criteria for the clinical diagnosis of classic AS, also known as X-linked Alport syndrome (XLAS). An individual with unexplained hematuria would probably have XLAS if, at least, three of the following four criteria were present: (i) Positive family history of hematuria with or without chronic renal failure (CRF); (ii) Typical

ultrastructural GBM changes in a renal biopsy specimen; (iii) High-tone sensorineural deafness; (iv) Characteristic ophthalmological signs (lenticonus and/or macular flecks) (Flinter, Cameron et al. 1988).

It was not until 1990 that molecular genetic techniques enabled the elucidation of the genetic basis of XLAS. By immunohistochemical analysis, α5(IV) chains were shown to localize to the GBM of the kidney and the *COL4A5* gene, coding for α5(IV) chain of collagen type IV, was found to be located to Xq22 (Hostikka, Eddy et al. 1990; Myers, Jones et al. 1990; Pihlajaniemi, Pohjolainen et al. 1990). The same locus had previously been identified by linkage analysis (Atkin, Hasstedt et al. 1988; Brunner, Schroder et al. 1988; Szpiro-Tapia, Bobrie et al. 1988). In addition, changes in the *COL4A5* sequence were identified in patients with AS, giving support to its X-linked inheritance pattern (Barker, Hostikka et al. 1990). In 1994, some families with AS were found to carry mutations in the *COL4A3* or *COL4A4* genes, located to the autosomal chromosome 2 (2q36-q37), molecularly demonstrating an autosomal recessive pattern of inheritance (Mochizuki, Lemmink et al. 1994) and proving the locus heterogeneity of AS. Diagnostic criteria were further extended by Gregory and colleagues (1996), who suggested two additional clinical criteria, namely the association of diffuse leiomyomatosis with AS (Cochat, Guibaud et al. 1988) and immunohistochemical absence of α3-5(IV) chains in the GBM or in the epidermal basement membrane (EBM) (Gregory, Terreros et al. 1996) (Table I). Finally, molecular genetic testing was also included as a major diagnostic criterion. Indeed, since the clinical diagnosis of AS is challenging due to the great spectrum of symptoms, the identification of pathogenic *COL4A5, COL4A4* or *COL4A3* mutation(s) became the most informative diagnostic criterion, also providing support for the prognosis, treatment and prevention of AS.

Table I. Diagnostic criteria of Alport syndrome.

	Flinter *et al*. (1988) *	**Gregory *et al*. (1996) ****
Hematuria	Unexplained hematuria (obligatory criterion)	Persistent hematuria, without evidence of another possibly inherited nephropathy such as thin GBM disease, polycystic kidney disease, or IgA nephropathy
Family history	Hematuria and /or CRF	Family history of nephritis or unexplained hematuria, in a first-degree relative of the index case or in a male relative linked through any number of females
ESRD	-	Gradual progression to ESRD, in the index case or at least two family members
Ultrastructural GBM abnormalities	Thinning of the GBM; Thickening of the GBM; Splitting of lamina densa; Electron-dense bodies in lacunae of lamina densa	Widespread GBM ultrastructural abnormalities, in particular thickening, thinning, and splitting
Immuno-histochemical GBM abnormalities	-	Immunohistochemical evidence of complete or partial lack of the α3(IV), α4(IV), and α5(IV)-chain, in the GBM, EBM or both
Hearing loss	Bilateral high-tone SNHL	Bilateral SNHL in the 2,000-8,000 Hz range: the hearing loss develops gradually, is not present in early infancy, and commonly presents before the age of 30 years
Ophthalmological abnormalities	Anterior lenticonus; Macular flecks	Ocular lesions, including anterior lenticonus, posterior subcapsular cataract, posterior polymorphous dystrophy, and retinal flecks
Leiomyomatosis	-	Diffuse leiomyomatosis of oesophagus, female genitalia, or both
[Absence of] Hematologic abnormalities †	-	Macrothrombocytopenia or granulocytic inclusions
Genetic molecular analysis	-	A *COL4A3*, *COL4A4*, or *COL4A5* mutation

* Combination of unexplained hematuria and at least three of the four criteria must be observed for clinical diagnosis of AS to be made. ** Combination of at least four of the ten criteria must be observed for diagnosis of AS to be made in an individual without family history; probable diagnosis of AS is made if family history is positive and one additional criterion is observed; definitive diagnosis of AS is made if family history is positive and two additional criteria are observed. CRF: Chornic renal failure. ESRD: End-stage renal disease. GBM: Glomerular basement membrane. EBM: epidermal basement membrane. SNHL: sensorineural hearing loss. † This criterion is now known to be diagnostic of Epstein and Fechtner syndromes, caused by a pathogenic *MYH9* mutation; thus excluding the diagnosis of Alport syndrome.

1.2. Prevalence

AS is one of the most frequent hereditary causes of end-stage renal disease (ESRD). The most widely used estimate of the prevalence of AS is 1:5,000 live births, based on the finding of about 300 cases in Utah and southern Idaho in a population of 1,500,000 people (Hasstedt and Atkin 1983). According to the United States Renal Data System (USRDS), approximately 0.2% of adults and 3% of children in the United States with ESRD carry a diagnosis of AS (United States Renal Data System. Annual Data Report 2008. USRDS - ADR. Available at http://www.usrds.org/adr.htm. Accessed June 19, 2013). In Europe, AS accounts for 0.56% of patients starting renal replacement therapy (RRT) (Rigden, Mehls et al. 1996; Levy and Feingold 2000; System, Health et al. 2010). The incidence of AS was found to be 1:53,000 in Finland (Pajari, Kaariainen et al. 1996), and 1:17,000 in southern Sweden (Persson, Hertz et al. 2005).

1.3. Clinicopathologic features

> "For decades, the syndrome was diagnosed solely based on its clinical manifestations."
> (Khoshnoodi, Pedchenko et al. 2008)

Collagen type IV-related nephropathies – a group of kidney diseases encompassing AS and Benign Familial Hematuria / Thin Basement Membrane Nephropathy (BFH/TBMN; MIM#141200) –, are hereditary glomerulopathies with three different inheritance patterns: X-linked, autosomal recessive and autosomal dominant. X-linked AS (XLAS; MIM#301050) explain approximately 80-85% of families with AS (Flinter, Cameron et al. 1988), while autosomal recessive AS (ARAS; MIM#203780) occurs in about 10-15% of families. An autosomal dominant inheritance pattern has been observed in around 5% of families with AS (ADAS; MIM#104200).

Since intrafamilial variability occurs, following examination of the proband and relatives, at least three out of four clinical criteria, as defined by Flinter and colleagues (1988) should be present (Flinter, Cameron et al. 1988) (Table I). Methodic investigation searching for a positive family history is important, as

patients are often not aware of the presence or severity of the disease in relatives. Urinalysis, audiological and ophthalmological examinations of parents and other first-degree relatives of a patient with unexplained hematuria will often be informative (Savige, Gregory et al. 2013). When diagnostic criteria are satisfied within a family, the pedigree should be studied to define the mode of inheritance of AS. When facing a patient with clinical diagnosis of AS and without evidence of family history of hematuria or CRF, both the occurrence of a *de novo* pathogenic *COL4A5* mutation, which is observed in 10-15% of males with XLAS, and ARAS should be considered (Kruegel, Rubel et al. 2013).

1.3.1. Impaired renal function

X-linked Alport syndrome in males

In XLAS, males often present with urinary abnormalities during childhood, the renal function gradually declines and ESRD is reached in late adolescence or young adulthood. In one third of males affected with XLAS, macrohematuria is observed during the course of an infectious event or a physical effort, at an average age of 3 years old (Flinter 1997) and all affected males have persistent microhematuria. These observations make hematuria the most prominent familial trait of the syndrome (Flinter, Cameron et al. 1988; Jais, Knebelmann et al. 2000). As a consequence, a 5 years old boy with consistent absence of hematuria in repeated urinalysis is very unlikely to develop AS (Flinter, Maher et al. 2003). Nearly all affected males develop proteinuria and become nephrotic, developing hypertension during adolescence (Grunfeld 1985; Kashtan and Michael 1996). Proteinuria does not precede hematuria (Firth and Hurst 2007). However, it may be the first sign to be noticed, as well as hearing loss, hypertension, end-stage renal disease (ESRD) or leiomyomatosis (Jais, Knebelmann et al. 2000). In the end, virtually all males with XLAS will develop ESRD. Two phenotypic types of AS were initially defined, according to the age of onset of ESRD: a “juvenile” type, in which male patients reach ESRD at an average age of 31 years or lower, and an “adult” type, in which ESRD develops only after 31 years (Schneider 1963). Extra-renal involvement was reported to be more frequently observed in “juvenile” type. However, with increasing knowledge in genotype-phenotype correlations, this

classification was rendered to more reliable diagnostic methods for genetic counselling, management and prognosis. Although inter- and intra-familial variability exists regarding the age at ESRD, 51%, 76.5% and 94% of males patients carrying a pathogenic *COL4A5* mutation reach ESRD before 20, 31 and 41 years old, respectively (Jais, Knebelmann et al. 2000).

X-linked Alport syndrome in females

Among females with XLAS, the clinical course is extraordinarily variable, presumably due to the X-chromosome inactivation (Rheault 2012). Somatic mosaicism in XLAS may also account for some cases of apparent non-penetrance in female carriers (Bruttini, Vitelli et al. 2000; Plant, Boye et al. 2000). More than 90% of females with molecularly confirmed XLAS manifested microscopic hematuria (Jais, Knebelmann et al. 2003). However, since it may be intermittent in the great majority of females, several urinalyses should be performed before excluding its occurrence. Macroscopic hematuria presents in around 35% of females with XLAS at an average age of 9 years and one-third develop hypertension (usually in middle-age) (Flinter 1997). Approximately 12% of females with XLAS develop ESRD before age 40 years, increasing to 30% by age 60 years and 40% by age 80 years (Jais, Knebelmann et al. 2003). Accordingly, whenever XLAS is molecularly confirmed in a female relative and she is being considered for living related kidney donation for transplant, the risk of the donor developing impaired renal function needs to be discussed (Kashtan 1990; Sessa, Pietrucci et al. 1995; Kashtan 2009; Rheault 2012). Proteinuria should be an exclusion criterion and even female donors with isolated microhematuria should be aware of an increased risk of CRF (Gross, Weber et al. 2009; Savige, Gregory et al. 2013).

Autosomal recessive Alport syndrome: patients and carriers

In ARAS, the course of the disease is similar to the affected males with XLAS, in both males and females. In particular, most patients who are affected with ARAS develop significant proteinuria in late childhood or early adolescence and ESRD before age 30 years (Kashtan and Michael 1996). An autosomal recessive inheritance pattern is suggested by parental consanguinity, by evidence

of equally severely affected male and female siblings at a young age and by occurrence of severe manifestations of the disorder in only one generation.

The parents of affected individuals with ARAS may develop isolated microscopic hematuria, although not manifesting hearing loss or ocular signs of the disease (Dagher, Buzza et al. 2001). Still, ARAS carriers show normal urinalysis in approximately half of the cases. A diagnosis of TBMN is established, in an ARAS carrier, when besides presenting with microscopic hematuria, associated or not with mild proteinuria, a thin GBM is visualized by electron microscopy on renal biopsy, an ultrastructural finding also observed as a normal variant in 5-10% of the population (Firth and Hurst 2007). TBMN is genetically heterogeneous, being caused by heterozygous pathogenic mutations in the *COL4A3* and *COL4A4* in 40-50% of the affected families and being transmitted in an autosomal dominant pattern, where male-to-male transmission may occur (Temme, Peters et al. 2012). Although TBMN was initially thought to be benign with excellent prognosis, the reason why it was previously called BFH, the work of two independent groups revealed that it was not indeed a benign disorder and that it should be the subject of annual long term follow-up (Voskarides, Damianou et al. 2007; Marcocci, Uliana et al. 2009). In some families with TBMN, it may be observed progression to CRF. In autosomal dominant AS (ADAS), age of onset of ESRD is later and progression may be more insidious than in ARAS (Kashtan and Michael 1996; Flinter, Maher et al. 2003). In summary, the spectrum of renal manifestations in ARAS carriers is wider and less benign than previously thought, as these individuals may express a gradient in severity of urinary abnormalities (hematuria with or without proteinuria) and renal function deterioration that may progress with age to ESRD. So, similarly to what is observed in XLAS females, due to the risk of progressive chronic kidney disease (CKD) with increasing age, long-term yearly follow-up of patients with BFH/TBMN by a nephrologist is advised (Temme, Peters et al. 2012).

Treatment of renal manifestations

Several therapeutic approaches proved effective in the treatment and prevention of renal manifestations. Pharmacologic inhibition of the renin-angiotensin-aldosterone system by using angiotensin-converting enzyme (ACE)

inhibition and angiotensin receptor blockade (ARB), tested in AS animal models and in clinical trials, has been shown to delay renal function impairment (Grodecki, Gains et al. 1997; Gross, Beirowski et al. 2003; Gross, Schulze-Lohoff et al. 2004). ACE inhibitors and/or ARBs should be administered to patients with AS who have proteinuria with or without hypertension. Both classes of drugs apparently help to reduce proteinuria by decreasing intraglomerular pressure. In children with AS and CKD, the decline in renal function is lower in those receiving ACEI and intensified blood-pressure control (Wuhl, Mehls et al. 2004; Wuhl, Trivelli et al. 2009). Moreover, the inhibition of angiotensin II, a growth factor that is implicated in glomerular sclerosis, can slow sclerotic progression. Furthermore, molecular genetic diagnosis before the onset of proteinuria would theoretically allow early effective medical intervention (Gross and Kashtan 2009). Also in females with XLAS who consider kidney donation, the risk of developing CRF might be minimized by post-transplant nephroprotective strategies (Gross, Weber et al. 2009). Cyclosporine treatment is not recommended due to its nephrotoxicity and the transient effect on proteinuria reduction (Charbit, Gubler et al. 2007; Massella, Muda et al. 2010).

Dialytic therapy and, ultimately, kidney transplantation, are usually offered to patients developing ESRD secondary to AS (Kashtan, McEnery et al. 1995; Haberal, Emiroglu et al. 2001; Byrne, Budisavljevic et al. 2002). This is due to several reasons: (i) renal transplantation significantly increases AS patients survival (Gobel, Olbricht et al. 1992; Karakayali, Pehlivan et al. 2008); (ii) recurrent disease does not occur in the transplanted kidney; (iii) the allograft survival rate in these patients is similar to that in patients with other renal diseases. However, post-transplant anti-glomerular basement membrane (anti-GBM) antibody nephritis, although rare, occurs in 1-5% of transplanted AS patients (Milliner, Pierides et al. 1982; Brainwood, Kashtan et al. 1998), increasing the risk of recurrence of this complication in this group of patients.

Biological treatment is under investigation and includes mainly two types: gene and cell based therapies. GBM reparation is the principal aim of biological treatment, since renal disease may be the only life-threatening complication of AS. Absence of life-threatening extra-renal manifestations turns the GBM of AS patients into the key target for gene therapy for several reasons: (i) the therapy

can be targeted solely at the kidney, with a specific circulatory system, decreasing systemic side effects; (ii) the turnover of type IV collagen is fairly slow, with a more than one year half-life. A lesson that was learnt from carriers is that concomitant expression of abnormal gene product is not an obstacle for gene therapy. Despite encouraging results in animal model studies, several issues need to be overtaken before translation of gene therapy into clinical practice (Heikkila, Tryggvason et al. 2000), including: (i) availability of an appropriate gene delivery system into cells of renal glomeruli; (ii) expression of the delivered type IV collagen gene in those cells; (iii) proper post-translational modifications and folding of the respective α(IV) chains which facilitate intracellular association into an α3α4α5(IV) triple helical molecule; (iv) incorporation of those heterotrimers, extracellulary, into the GBM, which could restore the deteriorated GBM structure (Heikkila, Tibell et al. 2001); (v) greater knowledge of the regulation of the genes requiring treatment (Tryggvason, Heikkila et al. 1997). Regarding cell based therapy, bone marrow-derived stem cells were shown to improve renal histology and function in *COL4A3* knockout mice (Prodromidi, Poulsom et al. 2006; Sugimoto, Mundel et al. 2006; LeBleu, Sugimoto et al. 2009). Mesenchymal stem cell therapy however was not proved to delay onset of end-stage renal disease (Ninichuk, Gross et al. 2006). The available experimental data on biological treatment antecipate a further therapeutic approach to AS patients, as soon as risks are addressed and prove to be minimal (Gross and Kashtan 2009). Furthermore, success of clinical trials will benefit from further efforts in gathering detailed information about the natural history of the disease, in order to show treatment efficacy by identifying a change in the disease natural course. The “European Alport Registry” and the “Alport Syndrome Treatments and Outcomes Registry (ASTOR)” aim to be large databases of AS patients followed in multiple centers, in Europe and North America, respectively. The active maintenance of both databases will be decisive for the development of future clinical trials (www.clinicaltrials.gov).

1.3.2. Impaired renal and skin structure

Renal biopsy from AS patients reveals characteristic, although not specific immunohistochemical and ultrastructural changes (Mazzucco, Barsotti et al. 1998; Mazzucco, De Marchi et al. 2002). Indications for renal biopsy are controversial;

however, when there is a clear familial segregation of microscopic hematuria, intermittent or persistent (>6 months), obtaining a kidney biopsy from an affected individual may be extremely informative, preventing the need of additional kidney biopsies in other affected family members.

Focal thinning of the *lamina densa* of the GBM in childhood is the first ultrastructural change, however it may be mistaken with TBMN, which may have a benign course (Grunfeld 1985). A renal biopsy of a patient with AS, performed at an older age, usually shows progressive widespread of characteristic ultrastructural findings of AS (Cangiotti, Sessa et al. 1996). Often indicated as pathognomonic, AS specific GBM changes include areas of GBM thinning and thickening, splitting and splintering of its lamina densa and entrapment of electron-dense granules within electron lucent areas of the GBM (Hinglais, Grunfeld et al. 1972; Spear and Slusser 1972) (Figure 1). The severity of GBM splitting correlates with degree of proteinuria and is indicative of progressive disease course (Rumpelt 1980; Basta-Jovanovic, Venkataseshan et al. 1990). Due to the diagnostic and prognostic information given by electron microscopy examination of a renal biopsy from a patient suspected of having AS, it is crucial that pathologists search for ultrastructural changes, specifically those typical of AS. Light microscopy is helpful solely in excluding diferential diagnosis (Pirson 1999).

Immunohistochemical analyses of collagen type IV expression in renal or skin biopsies are helpful diagnostic criteria. In the majority of males with XLAS, α5(IV) chains are lacking in the GBM. It is often found that α3(IV) and α4(IV) chain are simultaneously absent, since inability to synthesize α5 (IV) chains interferes with the assembling of α3α4α5 heterotrimers. In XLAS females, due to random inactivation of one X chromosome during the embryonic period, some cells will express α5(IV) chains necessary for the assembly of α3α4α5 triple helix of type IV collagen (Pirson 1999). As a result of the mosaic expression of α5(IV) chains, females (heterozygotes) show a patchy distribution of the collagen triple helix, in contrast with males (hemizygotes) who present a homogeneous co-absence of α-chains. Immunohistochemical analysis distinguishes XLAS from ARAS, because, in the latter, although there is absence of α3(IV) and α4(IV) chains, α5 (IV) chains are produced (Pirson 1999). Although not detectable in the GBM, they are visualized in Bowman’s capsule and tubuli basement membranes. The failure of

anti-GBM antibodies to bind to the GBM in patients with AS can be diagnostically useful, but has limited sensitivity and specificity (Mazzucco, Barsotti et al. 1997).

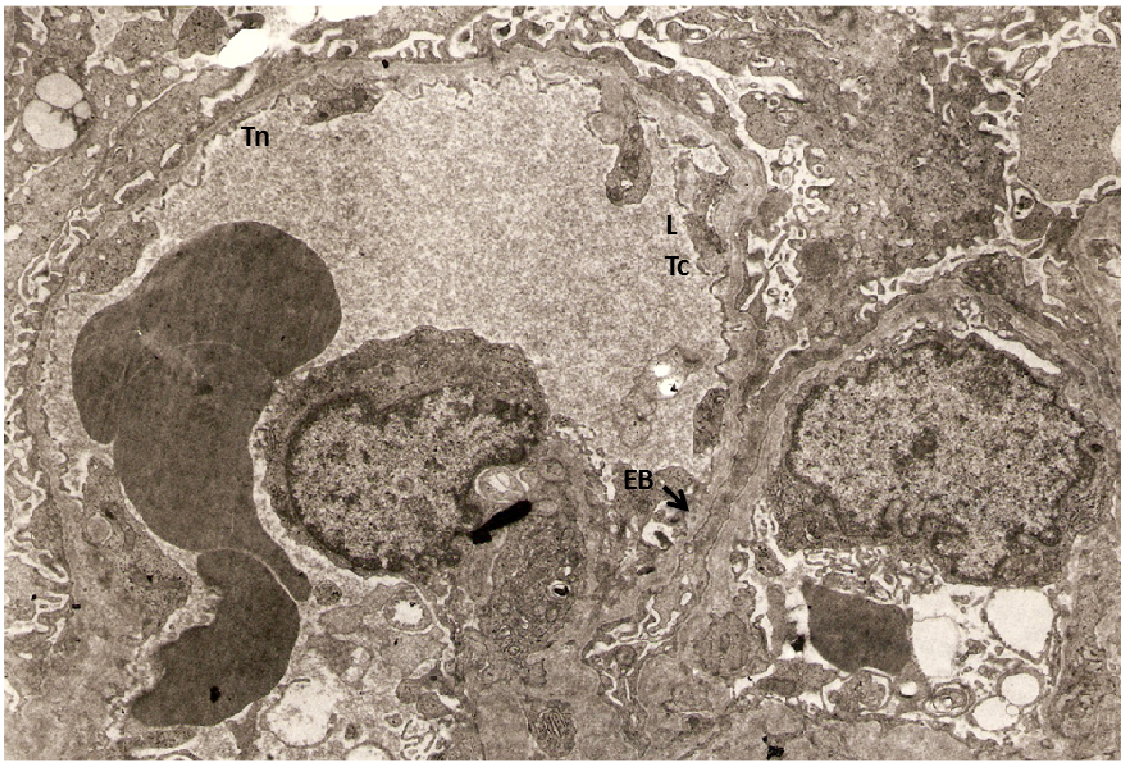


Figure 1. Photograph of the GBM ultrastructure of a patient with ARAS (25 years), showing alternating thinning (Tn), thickening (Tc) and lamellation (L) of the GBM with electrondense bodies (EB) surrounded by a halo [courtesy of Dr. Susana Sampaio, Unit of Renal Morphology, Department of Nephrology, São João Hospital Centre, Porto, Portugal].

In skin biopsies of individuals affected with XLAS, expression of α5(IV) chains in the EBM is similar to the observed in GBM. Absence of α5(IV) expression in the EBM of a male patient, or clearly mosaic α5(IV) expression in a female, suggests the diagnosis of XLAS. However, a discordant pattern with α5(IV) present in the EBM and absence in the GBM has been reported (Naito, Nomura et al. 1997). Importantly, normal expression of collagen type IV chains in the EBM does not exclude AS. Normal skin reactivity to α5(IV) chains antigens occurs in: (i) EBM of patients with ARAS, since only α3(IV) or α4(IV) chains are lacking (Gubler, Knebelmann et al. 1995); (ii) patients carrying a pathogenic *COL4A5* mutation which allows skin expression of α5(IV) chains. So, although

defective, all α3, α4 and α5(IV) chains may be present in X-linked and autosomal AS. Therefore, when α5(IV) chains are detected in the EBM, it is advisable to perform a renal biopsy to confirm whether the typical ultrastructural features of AS are observed (Kashtan 1999). On the other hand, immunohistochemical studies may be particularly helpful when standard electron microscopy examination of the GBM has given equivocal results (Flinter, Maher et al. 2003). The combination of electron microscopy and immunohistochemical studies enabled a diagnosis of AS in 92% of a cohort of 108 Italian patients (Mazzucco, Barsotti et al. 1998).

1.3.3. Sensorineural hearing loss

In the majority of XLAS and ARAS families, hematuria and CRF are accompanied by sensorineural hearing loss (SNHL) in the range of 2,000-8,000 Hz. According to Jais and colleagues, the proportion of XLAS males with SNHL is approximately 79% (Jais, Knebelmann et al. 2000). In XLAS females, SNHL is less frequent (Jais, Knebelmann et al. 2003). High-frequency SNHL may be asymmetrical, but is often bilateral. It is not present at birth and becomes apparent in late childhood or early adolescence, at an average age of 11 years old, in male patients with XLAS and patients with ARAS, and only exceptionally before 20 years old in females with XLAS (Flinter 1997; Jais, Knebelmann et al. 2003). SNHL may begin in adulthood in some families with autosomal dominant and X-linked inheritance patterns of AS (Barker, Pruchno et al. 1996).

SNHL is associated with the deterioration of renal function (Alves and Ribeiro Fde 2008). Consequently, its progression may be slowly and not easily perceptible during childhood, rapid in teens (especially in XLAS males and ARAS) and often reaches a plateau (Grunfeld 1985). SNHL may extend to lower frequencies and interfere with frequencies of conversational speech (Myers and Tyler 1972). However, patients do not usually become completely deaf. Hearing aids often show limited effectiveness (Flinter, Maher et al. 2003). Occasionally, there is some improvement in hearing post-transplant (Mitschke, Schmidt et al. 1975; McDonald, Zincke et al. 1978), but this may be secondary to the treatment of uremia (Mitschke, Schmidt et al. 1977). Since tonal audiometry detects SNHL before it is self-noticed, it should be performed in the evaluation of the affected patient and first degree relatives with suspected AS. Likewise, formal assessment

of hearing is indicated in any patient with unexplained hematuria. Conversely, a patient with SNHL with childhood or later onset should be investigated for urinary hematuria and CRF, as occasionally awareness of hearing problems may precede recognition of hematuria or other evidence of renal disease, and hearing impairment in members of families with AS is always accompanied by evidence of renal involvement (Flinter, Maher et al. 2003; Alves and de 2005; Alves and Ribeiro Fde 2008).

1.3.4. Ocular abnormalities

Ocular signs were observed in up to 72% of males and 38% of females with XLAS (Jais, Knebelmann et al. 2000; Dagher, Buzza et al. 2001; Jais, Knebelmann et al. 2003). Autosomal dominant inherited forms of AS are usually not associated with the characteristic ophthalmological signs, although ARAS may be (Colville, Savige et al. 1997; Colville and Savige 1997). The most frequent ocular findings associated with AS are central or peripheral dot-and-fleck retinopathy and anterior lenticonus (Sabates, Krachmer et al. 1983; Hentati, Sellami et al. 2008), which were observed in 30% and 15-20% of patients, respectively (Pirson 1999). They are not congenital, usually presenting in adolescents or young adults, in parallel with the decline in renal function (Jais, Knebelmann et al. 2000; Flinter, Maher et al. 2003). Lenticonus and retinopathy were first noted in 14 and 11 year olds, respectively (Zhang, Colville et al. 2008). Lesions of the cornea, as corneal endothelial vesicles (similar to posterior polymorphous dystrophy) and recurrent corneal erosion (RCE), occur more rarely (Snyers, Rhys et al. 1996; Bower, Edwards et al. 2009; Seymenoglu and Baser 2009). Patients with posterior polymorphous dystrophy should be examined for renal abnormalities and hearing loss (Teekhasaenee, Nimmanit et al. 1991). RCE has to be considered as an additional ocular manifestation of AS, with a prevalence of 15 to 21%. Ophthalmologists should also be aware of this association, when confronted with a patient suffering from non-traumatic RCE. (Snyers, Rhys et al. 1996; Rhys, Snyers et al. 1997). The demonstration of a dot-and-fleck retinopathy in any individual with a family history of AS or with end-stage renal disease is diagnostic of AS. The presence of anterior lenticonus or posterior polymorphous corneal dystrophy in any individual is highly suggestive of the

diagnosis of AS (Colville and Savige 1997). Additional ocular features described in XLAS include other corneal dystrophies, microcornea, arcus, iris atrophy, cataracts, spontaneous lens rupture, spherophakia, posterior lenticonus, poor macular reflex, fluorescein angiogram hyperfluorescence, electrooculogram and electroretinogram abnormalities, and retinal pigmentation (Colville and Savige 1997). These findings underline the great variation of ocular disorders related to AS (Herwig, Eter et al. 2011). Consequently, the assistance of an interested ophthalmologist is invaluable in the diagnosis of AS (Savige and Colville 2009).

The main lens abnormalities reported in patients with AS are anterior lenticonus, cataracts and spontaneous anterior capsular rupture. Anterior lenticonus was considered by some authors a pathognomonic sign of AS (Cheong, Kashtan et al. 1994), as more than 90% of cases of anterior lenticonus are associated with AS (Junk, Stefani et al. 2000). Consequently, any patient presenting with anterior lenticonus should be methodically investigated for microhematuria, CRF, SNHL and family history of any of these manifestations (Hild, Walter et al. 2009). Often bilateral and manifesting in early adulthood around the time the kidneys fail, anterior lenticonus is a conical or spherical protrusion of the anterior surface into the anterior chamber (Jacobs and Meire 2000; Junk, Stefani et al. 2000). Conversely, posterior lenticonus, often occurring in young children, unilaterally, sporadically and without association with systemic features, is not considered a specific sign of AS. Only rarely it is found bilaterally in patients with AS (Bhatnagar, Kumar et al. 1990; Sukhija, Saini et al. 2003; Vedantham, Rajagopal et al. 2005; Al-Mahmood, Al-Swailem et al. 2010). An “oil droplet” reflex is a typical sign of anterior lenticonus observed by direct ophthalmoscopy; slit lamp and biomicroscopy examination detect both anterior and posterior lenticonus (Jacobs and Meire 2000; Junk, Stefani et al. 2000). Anterior lenticonus may lead to visual impairment due to progressive myopia, anterior capsular cataract, or spontaneous rupture of the anterior lens (Olitsky, Waz et al. 1999; Flinter, Maher et al. 2003; Hild, Walter et al. 2009).

Retinal changes have been recognized for more than 60 years in patients with AS and may occur alone or in combination with anterior lenticonus or lens opacities (Fehmers and Croughs 1964; Reynolds 1964; Unger and Rother 1964). Retinal abnormalities include a perimacular dot-and-fleck retinopathy and a

peripheral fleck retinopathy, which might occur independently of each other; a 'dull macular reflex' or 'lozenge', when the perimacular flecks are confluent; and, rarely, a macular hole caused by retinal thinning (Savige and Colville 2009). The macular lesions observed in patients with AS are typically characterized by yellowish or whitish areas, which may vary from mainly pigmentary (“dots”) to mainly adjoining depigmented areas (“flecks”) within the macular and midperipheral regions of the retina (McCarthy and Maino 2000). In XLAS and ARAS the peripheral retinopathy occurs even when the central retinopathy is absent (Shaw, Colville et al. 2007). Clinicians must be aware that the "lozenge" or "dull macular reflex" described in AS is not a normal variant but reflects a severe, almost confluent perimacular dot and fleck retinopathy. This sign is useful diagnostically and also prognostically, since it is associated with early-onset renal failure (Colville, Wang et al. 2009). In contrary to the anterior lenticonus, retinal changes are rarely associated with visual impairment (Spraul and Lang 2000). Macular flecks in AS are not associated with demonstrable retinal dysfunction (Gehrs, Pollock et al. 1995).

Retinal abnormalities will be more evidently seen using a hand-held ophthalmoscope (fundoscopy), which may be complemented or replaced by bilateral ocular fundus photography (retinography), with the advantage of being analysed by several professionals at another location and/or time (for example, when monitoring the progression of the retinopathy). Careful retinal examination and photography that includes the periphery is a safe and inexpensive method that may help in the diagnosis of AS especially in carriers of X-linked disease (Shaw, Colville et al. 2007). Optical coherence tomography (OCT) has the powerful advantage that measurements can be performed over a period of time to longitudinally study disease progression (staging of ocular disease). Temporal macular thinning is a prominent sign associated with XLAS, suggesting that OCT measurements are essential in the diagnosis and prognosis of the disease. The pathological basis for the retinal abnormalities of XLAS remains to be established (Ahmed, Kamae et al. 2013). Hence, retinal OCT is a powerful technique to track the early stages of this disease in order to contribute to a better understanding of pathogenesis.

Although the identification of ocular signs of AS is not mandatory for the clinical diagnosis of the disease, eye examination (particularly of adult individuals)

may recognize abnormalities of the retina, the lens and the cornea which are typical of AS, enabling diagnosis and prognosis of the disease (Hentati, Sellami et al. 2008; Zhang, Colville et al. 2008; Xu, Zhang et al. 2010). Furthermore, diagnosis of ocular manifestations is crucial for an early surveillance and effective therapeutic intervention (Blaise, Delanaye et al. 2003). Regarding the treatment of the principal causes of visual impairment in patients with AS: (i) phacoemulsification and intraocular lens implantation, rather than correction of refractive errors, is a safe and efficient therapeutic choice for the management of anterior lenticonus secondary to AS, with or without associated cataract (Chung, Lin et al. 2007) (Zare, Rajabi et al. 2007) (Aslanzadeh, Gharabaghi et al. 2008) (Liu, Tan et al. 2008) (Seymenoglu and Baser 2009); (ii) routine treatment of cataracts is adequate (Kashtan 1993); (iii) assure cornea protection from minor trauma to avoid recurrent corneal erosions (Kashtan 1993).

1.3.5. Diffuse leiomyomatosis

Alport syndrome and diffuse leiomyomatosis (ATS-DL; MIM#308940) is a rare clinical association of benign nodular smooth muscle tumors of esophagus, tracheo-bronchial tree, and genital tract with XLAS manifestations (Miner 1999). This clinical entity was first described by Garcia-Torres in 1975 at the Sixth International Congress of Nephrology (Firenze, Italy), but only published in 1983 (Garcia Torres and Guarner 1983). It is likely that, at least, two formerly published reports already referred to this association (Johnston, Clagett et al. 1953; Kenney 1953). Diffuse leiomyomatosis, although rare, is a significant extra-renal cause of morbidity occurring in XLAS patients, both in children and in adults (Van Loo, Vanholder et al. 1997). ATS-DL has been regarded as a contiguous gene deletion syndrome, caused by deletions partially involving the *COL4A5* and *COL4A6* genes.

Leiomyomatosis is fully penetrant, both in males and females, while XLAS is fully penetrant in males, but has a lower penetrance in females (Dahan, Heidet et al. 1995). At least one female has already been reported who presented with DL but without signs of nephropathy, in whom a typical causative mutation of ATS-DL was found (Dahan, Heidet et al. 1995). Therefore, females with apparently isolated leiomyomatosis can be heterozygous for ATS-DL, their sons having a 50% risk of

developing ATS-DL, their daughters having a 50% risk of developing leiomyomatosis in association with variable degrees of severity of renal and extra-renal AS manifestations (Dahan, Heidet et al. 1995).

The clinical expression of ATS-DL is variable, even within the same family (Van Loo, Vanholder et al. 1997). The prognosis of ATS-DL depends as much on the location of the leiomyomas as on the nephropathy (Cochat, Guibaud et al. 1988). Typically, leiomyomatosis affects gastrointestinal, respiratory and female reproductive systems. The esophageal component is usually the first manifestation (Le Bras, David et al. 1998), being reported as earlier as 30 months of age (Cochat, Guibaud et al. 1988). Gastrointestinal symptoms can include dysphagia, odynophagia and retrosternal pain, regurgitation, and bleeding. (Lerone, Dodero et al. 1991). The diagnosis of esophageal leiomyomatosis may be delayed or may be mistaken as achalasia, based on the clinical presentation and the results of radiographic studies, endoscopy, and manometry (Leborgne, Le Neel et al. 1989; Legius, Proesmans et al. 1990; Rabushka, Fishman et al. 1991; Garcia-Torres and Orozco 1993). Leiomyoma of the esophagus must be included in the differential diagnosis of the mediastinal masses or when achalasia is considered as a diagnostic possibility in children (Bourque, Spigland et al. 1989).

Leiomyoma is the most common benign tumor of the esophagus (Bourque, Spigland et al. 1989). Leiomyomatosis associated with XLAS may be differentiated from sporadic esophageal leyomiomas, regarding the following clinical characteristics: (1) Age of onset: leiomyomas associated with AS are usually diagnosed in the pediatric population (mean age 14 years, range 4 to 20 years)(Bourque, Spigland et al. 1989), as opposed to isolated leiomyomas, usually diagnosed in the adult population (mean age 44 years; no patient was <12 years of age)(Seremetis, Lyons et al. 1976); (2) Sex ratio: in ATS-DL, leiomyomatosis appears 1.71 times more often in females than in males (Bourque, Spigland et al. 1989), while in the general population the female:male ratio is 1:1.9 (Seremetis, Lyons et al. 1976); (3) Extent of lesions: in ATS-DL, localized lesions are found in only 9%, whereas the diffuse form predominates in 91% (Bourque, Spigland et al. 1989), as opposed to the occurrence of only one tumor of variable size in the absence of ATS-DL (Seremetis, Lyons et al. 1976); (4) Location of lesions: in ATS-DL, the entire esophagus may be involved 35% of the time, and encroachment on

the cardia or upper stomach occurs in 70% (Bourque, Spigland et al. 1989), whereas, in a localized form of leiomyomatosis, the preferential location was the lower third of the esophagus, situated intramurally. Of note, the early age of onset of leiomyomatosis and its diffuse distribution are in keeping with its genetic etiology.

The genital manifestations are usually confined to women, typically involving the perineum. Because unusual, a combination of leiomyoma of vulva and esophagus (MIM#150700) should prompt search for other features of AS in the female patient herself or in male relatives (Cochat, Guibaud et al. 1988). Lesions can extend posteriorly to involve the rectum (Lerone, Dodero et al. 1991) and anteriorly to the vesical collar. Urethral leiomyomas have also been reported. In addition, uterine leiomyomas may occur in these patients (Cooper, Patterson et al. 1999).

Tracheobronchial leiomyomatosis is barely symptomatic. Histologic examination of relevant autopsy samples of three ATS-DL patients, disclosed extensive leiomyomas engulfing the trachea and stem bronchi (Cochat, Guibaud et al. 1988). Nonetheless, when symptomatic, it usually presents as paroxysmal dyspnea, mimicking asthma, and it may be lethal due to bronchospasm (Le Bras, David et al. 1998). When there is tracheobronchial involvement, careful endoscopic examination probably should be carried out because of the risk of sudden death (Cochat, Guibaud et al. 1988). Potentially fatal pulmonary complications have been observed in several patients and were attributed to tracheobronchial localization of leiomyomas (Garcia Torres and Guarner 1983; Van Loo, Vanholder et al. 1997). This diagnosis should be considered not only in the presence of esophageal leiomyomas in the child but also when hereditary nephritis is associated with digestive or respiratory signs (Cochat, Guibaud et al. 1988).

In summary, a timely diagnosis is important, in view of the morbidity and the inheritance pattern of the disease (Van Loo, Vanholder et al. 1997). Patients with AS should be examined for signs of leiomyomatosis and *vice versa* (Van Loo, Vanholder et al. 1997). Visceral leiomyomatosis is exceptional in children outside of the context of ATS-DL. In view of the important clinical and genetic implications, renal function and urinary status should be controlled in any patient

with esophageal leiomyomatosis. Conversely, the possibility of ATS-DL should be considered in AS patients with dysphagia (Van Loo, Vanholder et al. 1997).

1.3.6. Differential Diagnosis

The most common hereditary cause of hematuria progressing to ESRD is AS. However, clinical diagnosis of AS remain challenging due to age-dependent and gender-related expression of manifestations. In children, differentiation between TBMN and AS can be difficult, because both disorders manifest by persistent hematuria and thin GBM at that age (Lemmink, Schroder et al. 1997).

Additional etiologies of glomerular hematuria include a number of chronic glomerulopathies (Table II). In the child with no known family history of hematuria, the most likely diagnoses are IgA nephropathy, TBMN, AS, and membranoproliferative glomerulonephritis (Kashtan 1993). In addition, glomerular disorders including IgA nephropathy and focal segmental glomerulosclerosis may occur in individuals with TBMN (Norby and Cosio 2005). In a child with family history of hematuria and/or CKD, the causes of familial hematuria are TBMN and AS, *MYH9*-related disorders – previously known as Epstein syndrome, Fechtner syndrome, May-Hegglin anomaly, Sebastian syndrome (Arrondel, Vodovar et al. 2002) –, *CFHR5*-Related Dense Deposit Disease (DDD), glomerulopathy with fibronectin deposits and occasional familial cases of IgA nephropathy. Molecular genetics is crucial in discriminating the precise etiologic diagnosis when a patient presents with familial hematuria progressing to ESRD (Deltas, Pierides et al. 2013).

Table II. Differential diagnosis of hereditary nephritis syndromes of glomerular origin.

	Collagen IV-related nephropathies				***MYH9-Related Disorders***		***CFHR5 Nephropathy***	***Glomerulopathy with fibronectin deposits***
	Alport syndrome	**TBMN**	**ATS-DL**	**ATS-MR / AMME**	**Epstein syndrome**	**Fechtner syndrome**	**C3 Glomerulonephritis**	
Gene(s)	*COL4A3 COL4A4 COL4A5*	*COL4A3 COL4A4*	*Contiguous deletion of COL4A5 and COL4A6*	*Contiguous deletion including COL4A5*	*MYH9*	*MYH9*	*CFHR5*	*FN1*
Genetic loci	2q36.3 Xq22.3	2q36.3	Xq22.3	Xq22.3	22q12.3	22q12.3	1q31.3	2q35
Inheritance pattern	AR, AD, XL	AD	XL	XL	AD	AD	AD	AD
Hematuria progressing to ESRD	Frequent	Rare	Frequent	Frequent	+	+	+	+
Hearing loss	+	-	+	+	+	+	-	-
Ocular signs	+	-	+	+	+	+	+	-
Ultrastructural findings in the GBM	Alternate thinning and thickening, lamellation, electron-dense bodies	Thinning	Alternate thinning and thickening, lamellation, electron-dense bodies	Alternate thinning and thickening, lamellation, electron-dense bodies	Irregular thickness of the GBM	Irregular thickness of the GBM	Segmental, discontinuous, or diffuse pattern of dense deposits in the lamina densa	Mesangial and subendothelial deposits of fibrils or microtubules
Other features	-	-	Diffuse leiomyomatosis	Mental retardation Midface hypoplasia Elliptocytosis	Congenital macrothrombocytopenia	Congenital macrothrombocytopenia Leukocyte Inclusions	Immunohistochemical studies (kidney biopsy): C3-positive, usually in the absence of immunoglobulin deposition in the GBM	Immunohistochemical studies (kidney biopsy): fibronectin staining in the mesangium and GBM

TBMN: thin basement membrane nephropathy; ATS-DL: Alport syndrome and diffuse leiomyomatosis; ATS-MR / AMME: Alport syndrome with mental retardation, midface hypoplasia and elliptocytosis; XL: X-linked; AR: autosomal recessive; AD: autosomal dominat; + : Clinical sign present; - : Clinical sign absent; GBM: glomerular basementmembrane; ESRD: end-stage renal disease.

1.4. Molecular genetics and pathogenesis

"In 1990, the role of basement membranes in human disease was established by the identification of *COL4A5* mutations in Alport's syndrome."
(Van Agtmael and Bruckner-Tuderman 2010)

1.4.1. Type IV collagen coding genes

Genetic defects in genes that code for three of six collagen type IV α-chains – *COL4A3*, *COL4A4* and *COL4A5* – are the common genetic basis for a heterogeneous group of disorders, ranging from AS to TBMN (Table III). Pathogenic mutations in *COL4A5* result in X-linked inheritance of XLAS, ATS-DL and AMME complex. Pathogenic mutations in *COL4A3* and/or *COL4A4* are the cause of the autosomal recessive and the autosomal dominant forms of collagen IV-related nephropathies, while pathogenic mutations in *COL4A1* (MIM*120130) and *COL4A2* (MIM*120090) are not associated with any of those disorders.

Table III. Location of the human genes coding for α3, α4, α5 and α6 chains of type IV collagen.

Gene	MIM number	Genetic locus	Genomic coordinates (GRCh37)	Protein
COL4A3	MIM*120070	2q36.3	2: 228,029,280 - 228,179,507	α3(IV) chain
COL4A4	MIM*120131	2q36.3	2: 227,867,426 - 228,029,274	α4(IV) chain
COL4A5	MIM*303630	Xq22.3	X: 107,683,073 - 107,940,774	α5(IV) chain
COL4A6	MIM*303631	Xq22.3	X: 107,398,836 - 107,682,703	α6(IV) chain

These six evolutionarily related mammalian genes code for a family of six long, highly homologous polypeptides, designated α1(IV) to α6(IV), with three structurally different domains: (i) an amino-terminal 7S domain – thus named for being the sedimentation coefficient of the isolated crosslinked tetramer following bacterial collagenase treatment of a basement membrane (BM) (Timpl, Risteli et al. 1979; Risteli, Bachinger et al. 1980) –, containing a peptide signal and being essential for interchain crosslinking of four triple-helical molecules through dissulfide bonds and lysine-hydroxylysine crosslinks; (ii) a central collagenous domain, containing glycine-X-Y repetitions (where X and Y are proline or lysine residues, extensively hydroxy-

and glycosylated) separated by a variable number of interruptions, which are characteristic of each α-chain and which are responsible for their flexibility within the network and for serving as cell-binding sites and interchain crosslinking; and (iii) C-terminal noncollagenous (NC1) domain, where assembly of heterotrimers starts and dimerization of collagen type IV heterotrimers occurs (Table IV) (Hudson, Tryggvason et al. 2003; Khoshnoodi, Pedchenko et al. 2008). The crosslinked tetramer is heavily glycosylated making it resistant to collagenase activity (Risteli, Bachinger et al. 1980).

Table IV. Structure of the human α3, α4, α5 and α6 chains of type IV collagen.

Alpha (IV) chain	Chain size after translation*	Collagenous domain size*	NC1 domain size*	Number of interruptions in the collagenous domain	Cysteine residues (NC1 domain; 7S+collagenous domains)
α3(IV)	1670	1410	232	23	12; 12
α4(IV)	1690	1421	231	26	12; 20
α5(IV)	1685	1430	229	22	12; 8
α6(IV)	1691	1417	228	25	12; 9

* Number of amino acids. Adapted from (Khoshnoodi, Pedchenko et al. 2008)

The genes encoding the α1(IV) to α6(IV) chains are organized pairwise with head-to-head orientation, the *COL4A1 – COL4A2* located at chromosome position 13q34, the *COL4A3 – COL4A4* located at chromosome position 2q36.3, and the *COL4A5 – COL4A6* located at chromosome position Xq22.3. Due to the genes and respective α(IV)-chains similarity, *COL4A1*, *COL4A3* and *COL4A5* were grouped into α1-like group, while *COL4A2*, *COL4A4* and *COL4A6* belong to α2-like group. Based on sequence analysis from several mammalian species, Zhou et al. (1994) suggested that these gene pairs evolved from a unique gene through three consecutive gene duplications (Zhou, Ding et al. 1994). Another characteristic that makes type IV collagen unique in the large collagen superfamily is the fact that each gene pair shares a bidirectional promoter (Sugimoto, Oohashi et al. 1994; Heikkila and Soininen 1996). The bifunctional promoter of *COL4A5* and *COL4A6* regulates the expression of α5(IV) and α6(IV) chains in a distinct cell-specific fashion (Sund, Maeshima et al. 2005). Despite extensive studies and identification of the regulatory

elements in the promoter regions, the regulation of the differential expression collagen IV genes is not well understood (Khoshnoodi, Pedchenko et al. 2008).

1.4.2. Type IV collagen α-chains and protomers

The α(IV) chains associate in three different types of triple helical molecules – α1α1α2(IV), α3α4α5(IV), α5α5α6(IV) – that make up the building blocks of the collagen type IV network in BMs (Hudson, Kalluri et al. 1994; Boutaud, Borza et al. 2000; Borza, Bondar et al. 2001; Hudson 2004). The α(IV) heterotrimers (which are called protomers) assemble in the endoplasmic reticulum, a process starting as result of the specific interaction between three NC1 domains that form robust disulfide bridges and progressing toward the N terminus (Boutaud, Borza et al. 2000; Khoshnoodi, Cartailler et al. 2006). Finally, heterotrimers are secreted to the extracellular space, where they interact to form a basket weave collagen type IV network. Dimerization of protomers occurs by the interaction of C-terminal domains of two heterotrimers, while the N-terminal domain participates in tetramerization of the collagen type IV heterotrimers. Other interactions occur between protomer-protomer and protomer-extracellular matrix molecules (Vandenberg, Kern et al. 1991; Khoshnoodi, Pedchenko et al. 2008).

1.4.3. Type IV collagen networks

Collagen type IV is a major structural component of human BMs (Khoshnoodi, Pedchenko et al. 2008). The BM is a thin layer of specialized extracellular matrix found basolaterally to all epithelial and endothelial cell monolayers in the body, which compartmentalize tissues and provide important signals for the differentiation of the cells they support. BMs are important for the regulation of cell adhesion, migration and development and for tissue regeneration and wound healing; are a reservoir of growth factors and enzymes; and constitute a molecular sieve, which is a particularly important function in the GBM (Khoshnoodi, Pedchenko et al. 2008). In BMs, α(IV)-chains form a covalently-stabilized supramolecular network polymer, self-assembled through complex interactions between the triple helical α(IV)-chain monomers. Detailed analyses of high-resolution electron micrographs revealed that collagen molecules assemble to form irregular polygonal networks held together predominantly by overlapping lateral interactions along the triple-helical domain as

well as the N-terminal and C-terminal end-domains (Vandenberg, Kern et al. 1991). Supercoil formation by supramolecular twisting is a characteristic feature of lateral associations which provides further stability to the BM collagenous network (Vandenberg, Kern et al. 1991), important for its function in a complex molecular network. Through interactions with specific cellular receptors such as integrins, the BM collagen IV networks not only provide structural support to the cells and tissues, but they also affect their biological fate during and after development (Khoshnoodi, Pedchenko et al. 2008).

The expression of α(IV) chains is tightly regulated, both temporally and spatially (Butkowski, Wieslander et al. 1989; Miner and Sanes 1994; Sugimoto, Oohashi et al. 1994; Yoshioka, Hino et al. 1994; Gubler, Knebelmann et al. 1995; Ninomiya, Kagawa et al. 1995; Peissel, Geng et al. 1995). Unlike the 28 other different types of the large collagen superfamily, type IV collagen occurs only in the BMs. Networks composed of α1α1α2(IV)-α1α1α2(IV) hexamers are ubiquitously distributed in tissues and the first to be expressed during fetal life. It was observed that type IV collagen composition of the GBM changes during glomerular development and maturation (Abrahamson, Hudson et al. 2009). In some BMs, including the GBM and the tubular BM in the kidneys, the alveolar BM in the lungs, and ocular and the cochlear BMs, these α1α1α2(IV)-α1α1α2(IV) networks are gradually replaced, since early in the fetal development and through adult life, by networks composed of α3α4α5(IV)-α3α4α5(IV) hexamers (Butkowski, Wieslander et al. 1989). In the BMs of the skin, tracheobronchial tree, Bowman's capsule, and of smooth muscle cells in the esophagus and other parts of the gastrointestinal tract, bladder and uterus, a network of α1α1α2(IV)-α5α5α6(IV) hexamers is expressed after the embryonic period. In comparison to the α1α1α2(IV) network, the α3α4α5(IV) network is more highly crosslinked by inter- and intrachain disulfide bonds (Hudson, Kalluri et al. 1992) and the switch to a predominant expression of α3(IV), α4(IV) and α5(IV) chains confers to the GBM a greater resistance to proteolysis and to the age-related increase of the capillary glomerular pressure. In summary, while the α1(IV) and α2(IV) chains are present in the BM of all tissues, the remainder α(IV) chains are expressed in tissues where its greater structural resistance is required. The α3(IV), α4(IV), and α5(IV) chains are expressed in the GBM of the kidney, lung, testis, and eye, whereas the

α5(IV) and α6(IV) chains are found in the BM of skin, smooth muscle, and the kidney (Khoshnoodi, Pedchenko et al. 2008).

The developmental switch is abnormal in AS (Kalluri, Shield et al. 1997). In males with XLAS or ATS-DL, the α1(IV) and α2(IV) are the predominant α(IV)-chains identified by immunohistochemistry of the GBM, a pattern resembling the early stages of glomerular development. Most of these patients have no GBM expression of α3(IV), α4(IV) and α5(IV), but some *COL4A5* mutations allow partial formation of α3α4α5(IV) networks, resulting in less severe phenotypes. It has been suggested that pathogenic mutations of α3(IV), α4(IV) or α5(IV) might impair the folding and assembly of the α(IV) heterotrimers, resulting in their rapid degradation inside the cell and arrest of the normal developmental switch, leading to the persistence of α1α1α2(IV) networks in the mature GBM (Thorner, Zheng et al. 1996; Kalluri, Shield et al. 1997; Harvey, Zheng et al. 1998). Absence of α6(IV) has also been noted in renal BMs and the EBMs of XLAS patients, suggesting that the α6(IV) chain cannot assemble into these BMs without the α5(IV) chain (Hino, Takemura et al. 1996; Miner 1999). Irrespective of which gene is affected, mutations that alter normal expression, posttranslational modifications or assembly of the α3α4α5(IV) protomers will cause an arrest in the developmental switch and persistence of the α1α1α2(IV) network in the GBM (Gunwar, Ballester et al. 1998). The differences in their intrinsic physical and biological properties most likely underlie the inability of the α1α1α2(IV) network to compensate for the lack of the α3α4α5(IV) network in patients with AS (Gunwar, Ballester et al. 1998).

1.5. Genotype-phenotype correlations

The collagen IV-related nephropathies are caused by mutations in any of the *COL4A5*, *COL4A4* and *COL4A3* genes (Barker, Hostikka et al. 1990). The association of diffuse leiomyomatosis and Alport syndrome (ATS-DL) results from a contiguous gene deletion with partial loss of the 5’ end of *COL4A5* and the first two exons of the *COL4A6* gene (Cochat, Guibaud et al. 1988; Zhou, Mochizuki et al. 1993). The association of Alport syndrome with mental retardation, midface hypoplasia and elliptocytosis (ATS-MR or AMME; MIM#300194) is another contiguous gene deletion syndrome on the same chromosomal locus but, in contrast

to ATS-DL, extends in telomeric direction rather than centromerically (Jonsson, Renieri et al. 1998; Meloni, Vitelli et al. 2002; Rodriguez, Bhat et al. 2010).

More than 700, 60 and 90 pathogenic mutations have already been reported respectively in *COL4A5*, *COL4A4* and *COL4A3* [The Human Gene Mutation Database at the Institute of Medical Genetics in Cardiff, 2008. (Available at: http://www.hgmd.cf.ac.uk/ac/gene.php?geneCOL4A5. Accessed March 9, 2014)]. Other locus-specific databases are: (i) The Alport syndrome *COL4A5* variant database (http://arup.utah.edu/database/ALPORT/ALPORT_display.php) (Crockett, Pont-Kingdon et al. 2010); (ii) *COL4A5* homepage - Collagen, type IV, alpha (https://grenada.lumc.nl/LOVD2/COL4A/home.php?select_db=COL4A5) (Savige, Ars et al. 2013). Pathogenic *COL4A5* mutations are found in approximately 80-85% of families with clinical diagnosis of AS, while a minority carries pathogenic mutations in *COL4A3* or *COL4A4*. Pathogenic *COL4A5*, *COL4A4* and *COL4A3* mutations are often unique to each family. The nonprivate mutations are probably explained by common ancestors (Barker, Pruchno et al. 1996). Pathogenic mutations are distributed throughout these genes and include mostly small mutations, as missense, nonsense, frameshift and splice site mutations, but also large deletions in 10-15% of families. Pathogenic *COL4A5* mutations occurred *de novo* in 10-18% of families (Knebelmann, Breillat et al. 1996; Lemmink, Schroder et al. 1997; Inoue, Nishio et al. 1999; Plant, Green et al. 1999; Jais, Knebelmann et al. 2000). *COL4A5* molecular testing should be performed in any patient meeting at least two clinical diagnostic criteria of AS; *COL4A3* and *COL4A4* analysis should be considered primarily if autosomal inheritance is suspected and if a *COL4A5* mutation is not detected (Hanson, Storey et al. 2010). Initially, genealogical, clinical, histopathological data, and the relative frequencies of collagen IV-related nephropathies should guide molecular genetic analysis, which can be quite expensive and time consuming (Artuso, Fallerini et al. 2012). More recently, the possibility to simultaneously study *COL4A5*, *COL4A4* and *COL4A3* by Next Generation Sequencing (NGS) prior to a renal or skin biopsy, not only confirms information on the family history and clinical features, but it even adds reliable information on the natural history of the disease (Artuso, Fallerini et al. 2012).

Large genotype-phenotype correlations studies in European and North American populations showed that age of onset of ESRD, hearing loss detected by audiometry and typical ocular signs in males with XLAS is associated with the type of mutation in *COL4A5* gene. Although pathogenic *COL4A5* mutation types were grouped differently in different studies, genotype-phenotype correlations were similar. Age of onset of ESRD was earlier and hearing loss and ocular changes were more frequent in patients with mutations which lead to premature end of translation (large rearrangements, frameshift and premature stop mutations) than in patients with in-frame deletions/insertions or nucleotide substitutions (missense mutations) (Jais, Knebelmann et al. 2000; Gross, Netzer et al. 2002; Bekheirnia, Reed et al. 2010). Furthermore, independently of type of mutation, mutations occurring closer to the 5' end of *COL4A5* gene were suggested to be significantly correlated with younger age at onset of ESRD, presence of hearing loss and occurrence of ocular changes (Bekheirnia, Reed et al. 2010). A possible explanation may be the production of smaller proteins, in case of emergent truncated collagen IV α-chains, since translation begins in direction 5' → 3' of the gene. However, glycine substitutions in exons 1-20 were associated with a less severe phenotype (Gross, Netzer et al. 2002). In females heterozygous for *COL4A5* mutations no significant genotype-phenotype correlations were identified, but a high risk of progression to ESRD (50% at the age of 40) was emphasized, with hearing loss and proteinuria being adverse prognostic factors (Jais, Knebelmann et al. 2003).

Type of mutations in the *COL4A5* gene was shown to have great impact in management and genetic counseling of males with XLAS, since it influences development of ESRD, hearing loss and ocular abnormalities (Gross, Netzer et al. 2002). Gross and colleagues (2002) analyzed the genotype of XLAS patients by dividing mutations in seven types: 5' glycine substitutions (class 1), 3' glycine substitutions (class 2), in-frame mutations (class 3), splice donor mutations (class 4), splice acceptor mutations (class 5), frameshift and premature stop mutations (class 6), and large rearrangements (class 7) (Gross, Netzer et al. 2002). Genotype-phenotype analysis clustered patients with class 4, 6 and 7 mutations in a group with mean age of onset of ESRD of 19.8 ± 5.7 years (also referred as "truncated protein group", since these mutations are predicted to prevent the synthesis or lead to the

synthesis of a truncated protein); class 2, 3 and 5 mutations formed a group with mean age of onset of ESRD 25.7 ± 7.2 years (also called “altered protein structure group”); patients with class 1 mutations developed ESRD at an older age (30.1 ± 7.2 years) (Gross, Netzer et al. 2002). Percentages of patients with hearing loss and typical ocular signs were also higher in “truncated protein group” than in “altered protein structure group” (80% versus 65-70% and 40% versus 30%, respectively). Furthermore, in this meta-analysis, Gross and colleagues (2002) suggested a classification of the phenotype based on the genotype: (i) severe phenotype (type S: ESRD at ~20 years of age, 80% hearing loss, 40% ocular lesions), which included patients with class 4, 6 and 7 mutations; (ii) moderate-severe phenotype (type MS: ESRD at ~26 years of age, 65% hearing loss, 30% ocular lesions), which included patients with class 2, 3 and 5 mutations; (iii) moderate phenotype (type M: ESRF at ~30 years of age, 70% hearing loss, 30% ocular lesions), which included patients with class 1 mutations (Gross, Netzer et al. 2002).

If a mutation known to be the cause of AS was previously identified in the proband, a family member presenting with hematuria should be tested, independently of the need of further clinical tests (Flinter, Maher et al. 2003). Women carrying a pathogenic *COL4A5* gene mutation have 50% (1/2) probability of transmitting the same mutation either to female or male offspring. Affected males do not transmit a *COL4A5* mutation to their sons, but all his daughters will inherit the pathogenic mutation. Germline mosaicism of pathogenic *COL4A5* mutations was previously reported, and, apart from *de novo* mutations, it may explain the non-penetrance in female carriers without affected ancestors. Germline mosaicism increases the recurrence risk of XLAS in the offspring of the molecular mosaic male or female (Bruttini, Vitelli et al. 2000). When both parents are carriers of either *COL4A3* or *COL4A4* mutations, the couple has a 25% (1/4) probability of having an affected descendent. On the other hand, 50% of the progeny of such couples will be heterozygous for one the mutations carried by the parents and should receive appropriate long term clinical follow up (Frasca, Onetti-Muda et al. 2005; Marcocci, Uliana et al. 2009; Temme, Peters et al. 2012). Carrier status detection and specific reproductive options, as prenatal diagnosis (PND) and preimplantation genetic diagnosis (PGD), are available after gene mutation identification.

Almost one century has passed since the publication of the seminal Arthur C. Alport's clinical report and much has been learned from subsequent research, largely through molecular biology approaches in the last three decades. AS should be suspected when facing a young patient with hematuria or CRF, especially in the presence of family history of glomerulonephritis. Audiologic and ophthalmologic evaluations are not only diagnostic tools, but also predictive factors of general prognosis of kidney function (Bekheirnia, Reed et al. 2010). Since its availability, the molecular diagnosis of collagen IV-related nephropathies became a preferential method to confirm the etiologic diagnosis and the mode of transmission of these disorders in each family, offering a non-invasive diagnostic approach and making pre-symptomatic diagnosis available. Moreover, an earlier genetic diagnosis of a patient or a carrier has significant implications in the treatment, family genetic counselling, and reproductive options. The genetic diagnosis enables the treatment decision-making to be timely and standardized, allowing for earlier therapeutic interventions that may favorably impact on the clinical prognosis (Temme, Peters et al. 2012). Since intra- and inter-familial phenotypic variability is observed, the identification of pathogenic *COL4A5*, *COL4A4* or *COL4A3* mutations differentiates between X-linked and autosomal inheritance, enabling the recurrence risk to be accurately estimated and avoiding kidney donation by potential kidney donors who are at-risk of CRF (Gross, Weber et al. 2009). Finally, once the disease-causing mutation is identified, reproductive options became broader for the patients and their relatives.

2. Significance, Hypothesis and Objectives of the Study

The increasing number of publications describing effective therapeutic options for AS and the requirement of the underlying molecular defect characterization for precise diagnosis, genetic counseling and appropriate therapy (Savige, Gregory et al. 2013) determined the need for identifying pathogenic mutations in the genes known to cause collagen IV-related nephropathies in the Portuguese population of patients with AS. Although not being considered an innovative methodology, the molecular genetic analysis of those genes in a never previously genotyped population was

expected to reveal novel data on the molecular pathology of this spectrum of disorders, due to their *locus* and allelic heterogeneity. Concomitant clinical characterization of index cases and family members would be needed to confirm the pathogenicity of the novel DNA variants. Moreover, detailed knowledge on the molecular pathology of collagen IV-related nephropathy in the Portuguese population would guide the implementation of a molecular genetics study strategy adjusted to Portuguese patients with clinical diagnosis of AS, for use in clinical practice.

The detection of previously unreported as well as reported pathogenic variants in the *COL4A5*, *COL4A4* and *COL4A3* genes in a cohort of well clinically characterized patients with AS would enable the extension of genotype-phenotype correlations, clarifying the pathogenic mechanisms involved in collagen type IV related disorders, while having a direct utility in the clinical practice by (i) confirming the clinical diagnosis, (ii) reducing the time and invasiveness of the diagnostic process, (iii) assisting in choosing an adequate treatment and surveillance of patients, (iv) supporting the genetic counseling including the identification of at risk relatives, (v) preventing kidney donation between affected family members, essentially by estimating the risk of living-related female kidney donors developing CKD, and (vi) increasing the available reproductive options of couples at-risk of having affected offspring. Eventually, the identification of a more restricted group of patients with clinical features within the spectrum of collagen IV-related nephropathies but no detectable pathogenic variants in the *COL4A5*, *COL4A4* and *COL4A3* genes would be formed, in whom clinical and molecular genetic analyses should be persued with the aim to investigate new genotype-phenotype associations.

A first challenge of pioneering the study of families with AS in Portugal was the lack of knowledge regarding the number and phenotypic variation of Portuguese patients affected with collagen type IV glomerulopathy, which lead to the need of adopting broad inclusion criteria. The phenotypic variability, as well as the *locus* and allelic genetic heterogeneity underlying AS, would raise further challenges in distinguishing pathogenic from benign molecular changes. After identification of genetic defects in the *COL4A5*, *COL4A4* or *COL4A3* genes, molecular variations underlying AS in the population of Portuguese patients were further characterized by (i) studying first degree relatives of the proband and other at-risk, affected and

unaffected family members, (ii) investigating for the presence of the variation in Portuguese healthy controls and in described European-American control populations, (iii) predicting *in silico* the effect of novel genetic variants in the structure and function of the corresponding protein using bioinformatic tools. In the end, the relevance of the implementation of the clinical and molecular study of Portuguese families with AS would be (i) to determine the *COL4A5*, *COL4A4* and *COL4A3* mutation detection rate in Portuguese AS patients and (ii) to broadly describe the genetic epidemiology of AS in Portugal.

Consequently, a largely exploratory research project was undetaken, based on the following research questions (Figure 2):

Research question 1: Are the type of pathogenic mutations in the genes known to cause collagen IV-related nephropathies – *COL4A3*, *COL4A4*, *COL4A5* and *COL4A6* – and their frequency among Portuguese families with AS similar to other cohorts?

Hypothesis 1: The type and frequency of pathogenic *COL4A3*, *COL4A4*, *COL4A5* and *COL4A6* mutations in Portuguese families are different from mutations found in other countries.

Aim 1: To identify pathogenic mutations in the *COL4A3*, *COL4A4*, *COL4A5* and *COL4A6* genes in Portuguese families with clinical diagnosis of AS and compare them with other previously identified pathogenic mutations.

Research question 2: Are clinical features in Portuguese individuals with pathogenic mutations in the genes known to cause collagen IV-related nephropathy similar to clinical features seen in other cohorts of patients with collagen IV-related nephropathy?

Hypothesis 2: The phenotype of Portuguese individuals with pathogenic mutations identified in the genes known to cause collagen IV-related nephropathy is similar to the phenotype observed in other cohorts.

Aim 2: To perform the clinical characterization of Portuguese individuals with pathogenic mutations identified in the genes known to cause collagen IV-related nephropathy and compare their phenotype with other cohorts.

Research question 3: Is the association between genotype and phenotype in Portuguese individuals with pathogenic mutations identified in the genes known to cause collagen IV-related nephropathy similar to the genotype-phenotype association seen in other cohorts of patients with collagen IV-related nephropathy?

Hypothesis 3: The association between the genotype and the phenotype in Portuguese individuals with mutations identified in the genes known to cause collagen IV-related nephropathy is similar to the genotype-phenotype association observed in other cohorts.

Aim 3: To establish genotype-phenotype correlations in a Portuguese cohort of individuals with mutations identified in the genes known to cause collagen IV-related nephropathy and compare it with other cohorts.

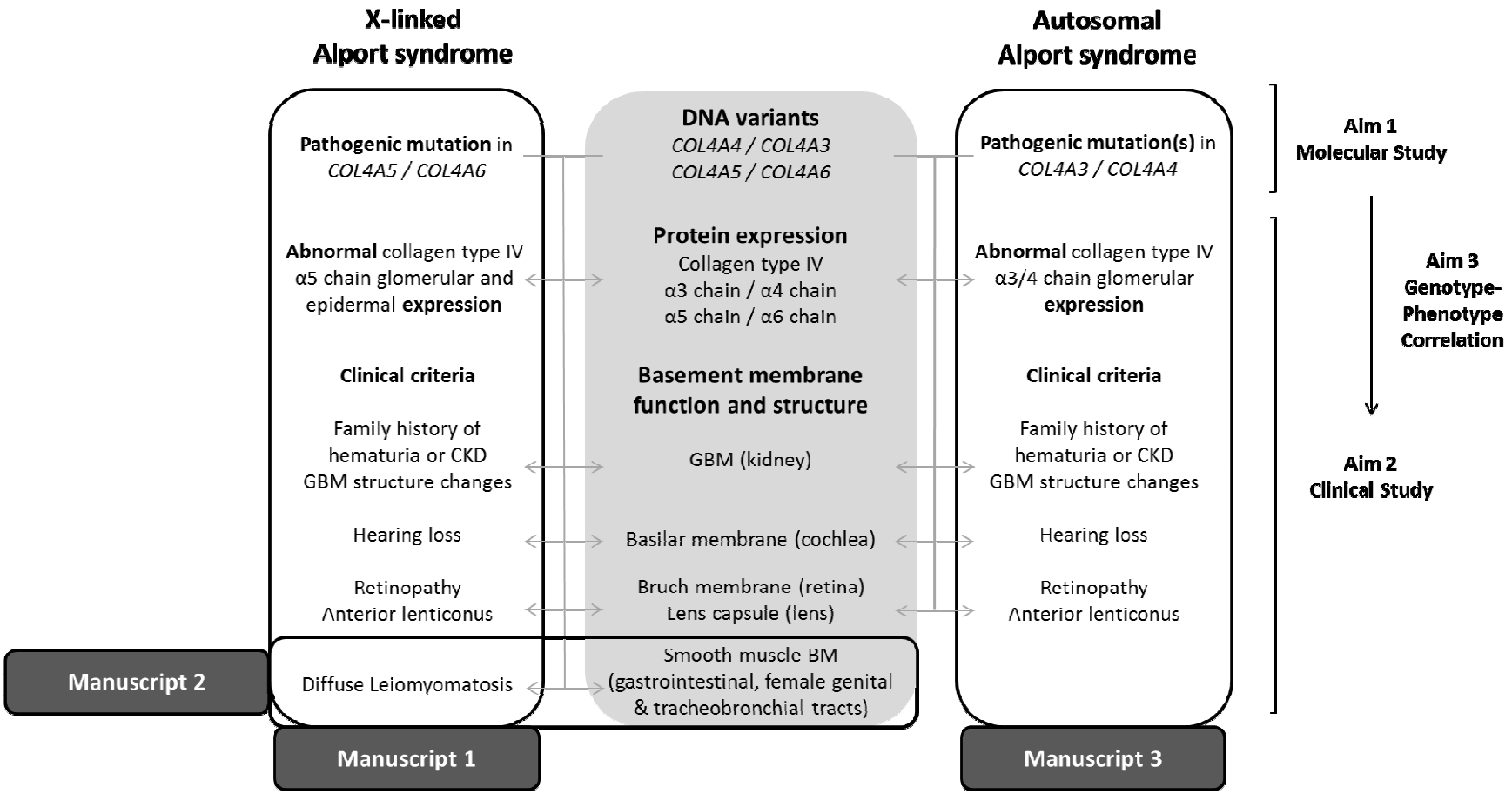


Figure 2. Overview of the research project, showing the aims 1, 2 and 3 and the research areas investigated in each manuscript. GBM: Glomerular basement membrane.

Chapter 2 Patients and Methods

1. Study cohort

The study cohort consisted of patients with a clinical diagnosis of AS, treated or followed at clinical nephrology, transplantation or dialysis clinics of public hospitals or treated in private hemodialysis clinics in Portugal. Since few mutations in Portuguese patients with AS had already been reported (namely patients included in the European Community Alport Syndrome Concerted Action – ECASCA) (Jais, Knebelmann et al. 2000; Jais, Knebelmann et al. 2003), the cohort assembly was also expected to be a largely exploratory process of making an inventory of Portuguese families with AS and of recognition of the disease causing mutations in each family. The research project was approved by the Health Ethics Committee of São João Hospital Centre (CHSJ), Porto, Portugal.

A closed cohort was gathered between the 1st January 2009 and the 30th June 2012. Index cases were assembled by simple random sampling, while family members were identified by chain-referral sampling (Table V). Probands were referred for molecular study of the *COL4A5*, *COL4A4* and *COL4A3* genes based on the occurrence of at least one of the four clinical criteria proposed by Flinter (Flinter, Cameron et al. 1988). This cut-off inclusion criterion would enable the constitution of a cohort of patients with hereditary nephropathy, facilitating the enrollment of carriers of XLAS or ARAS and patients with early manifestations of XLAS or ARAS, besides affected probands with well-founded clinical diagnosis of XLAS or ARAS. Informed consent for participation in this research project, and subsequent collection of clinical data and biological samples was obtained from 178 individuals who fulfilled the eligibility criteria, belonging to 65 unrelated Portuguese families.

Table V. Eligibility criteria for selection of the study sample.

Subjects eligible for the research study:
▪ Index cases with history of hematuria of unknown etiology and one or more of the clinical criteria of AS proposed by Flinter *et al*. (1988) ▪ Family members of index cases with clinical signs of AS or ≥ 18 years without clinical signs of AS

The specific characteristics of the sample that made it reliable as a set of one or more examples of the unit of observation are:

- The gross or less specific (versus fine or with highly specificity) selection of families with collagen IV-related nephropathy (AS or TBMN);
- The study of *COL4A5*, *COL4A4* and *COL4A3* genes in all, but three, of the 65 selected families (Figure 3).

In three probands without a pathogenic *COL4A5* mutation, it was decided not to perform the molecular analysis of *COL4A3* and *COL4A4* within this research project, as: (i) in one family, there was insufficient available clinical data (family 59); (ii) in two families, clinical manifestations were not characteristic of AS (family 27 and 58).

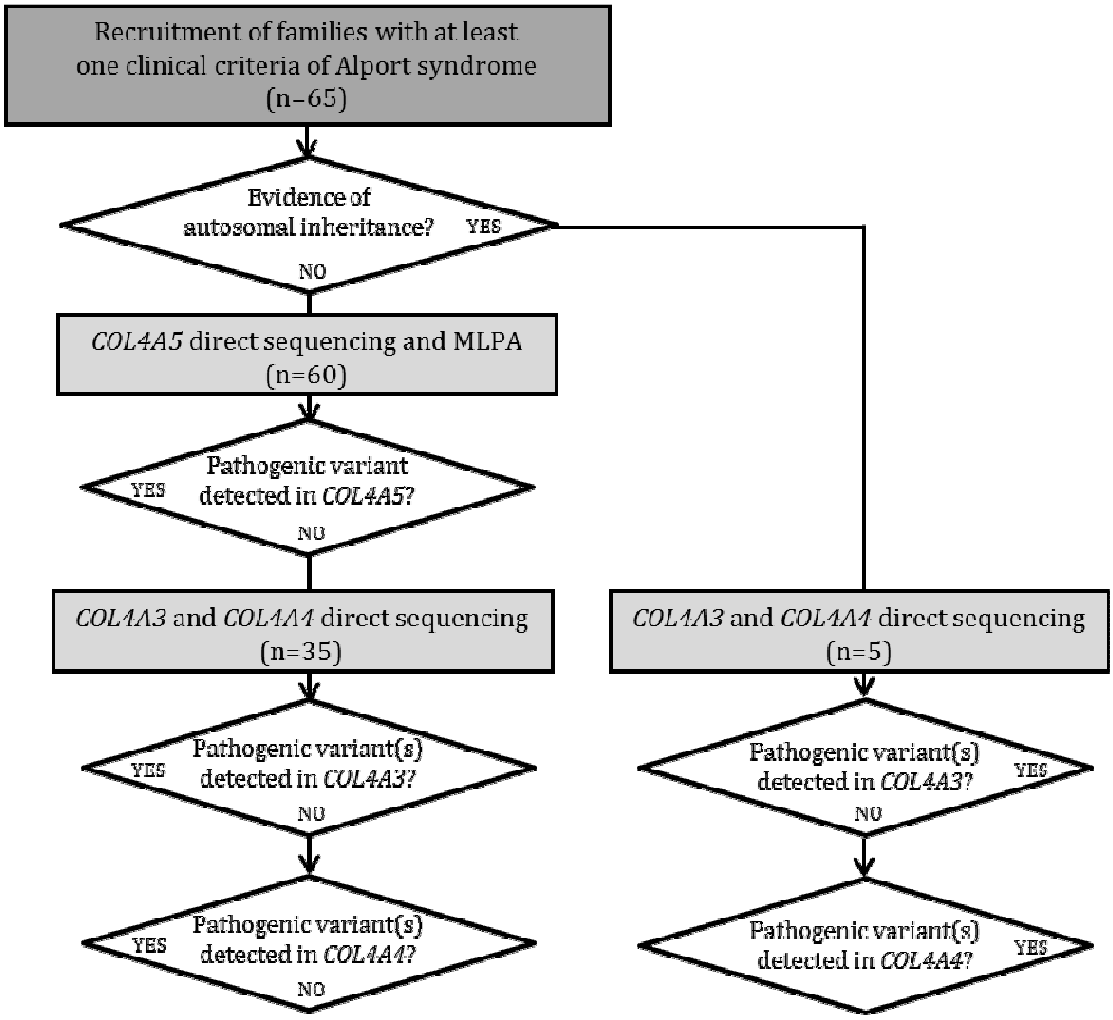


Figure 3. Research study flowchart.

1.1. Units of analysis

To operationalize the impact of the genotype (exposure) on the phenotype (outcome) of Portuguese families with collagen IV-related nephropathies, one unit of analysis of the phenotype, one unit of analysis of the genotype and one unit of analysis of the genotype-phenotype correlation were determined.

The elected phenotypic unit of analysis was the individual. The levels of analysis of the phenotype are the function and, when available, the structure of three main organs affected by pathogenic *COL4A5*, *COL4A4* and *COL4A3* mutations – kidney, ear and eye.

The selected genotypic unit of analysis was the disease causing mutation. The levels of analysis chosen, or the scales of molecular unit of analysis, are the mutation status and the type of the mutation. Mutation statuses are hemizygous, homozygous, compound heterozygous and heterozygous. Mutations types were previously proposed to be aggregated in two or more groups: type M – moderate – versus type MS – moderate/severe – versus type S – severe (Gross, 2002), truncating versus non-truncating (Hertz, 2009). Other levels of analysis may be the exon, intron or region (5' versus 3') in the gene where the mutation occurred, the class of aminoacid that was changed, the domain of the protein the mutation affected, etc.

The selected unit of analysis to characterize the phenotypic effect of each mutation was the group of patients with similar genetic background, i.e. identical mutation status or identical mutation type, as mutations in the three know genes that cause AS are usually private, i.e. unique to a single family.

The major benefit of using a group of patients with identical mutation status or identical mutation type as a unit of analysis is because it corresponds to the cohort of patients that express the phenotypic effect of a specific mutation. Additionally, by identifying a first patient in a family, and usually the more severely affected, it is relatively easy to localize, contact and characterize other family members carrying the same pathogenic mutations. Family members are generally motivated to have the cause of the disease confirmed by a genetic test and to understand the inheritance pattern and transmission risks of the disease in the family. The limitation of choosing families as units of analysis is the expected intra-familial and inter-familial variability

and the possibility that it may not be feasible to enroll several patients by family to increase the number of observations of the phenotype of a particular pathogenic mutation. Consequently, there is the risk that, if only one member of a family is studied, an outlier may be characterized, not being sufficient to adequately characterize the phenotypic effect of the mutation.

The levels of analysis of the genotype-phenotype correlation are the function and, when available, the structure of three main organs affected by pathogenic *COL4A5*, *COL4A4* and *COL4A3* mutations – kidney, ear and eye – by mutation status or mutation type.

1.2. Units of observation

The unit of observation about which data were collected to characterize the unit of analysis was the individual, including both affected and non-affected family members of each family. At a molecular level, the affected individual is the “perfect” proxy for the family, because it is expected that all affected family members will carry the same mutation(s). At the phenotypic level, the benefit of collecting data on the individual as a proxy for the clinical characterization of the family is to identify and increase the number of observed outcomes occurring as the result of a single mutation (and many times unique, i.e. not previously identified or reported). Consequently, it will enable:

- Observation of the phenotypic effect of hemizygous or heterozygous pathogenic *COL4A5* or *COL4A5_COL4A6* mutations;
- Observation of the phenotypic effect of truncating and non-truncating pathogenic *COL4A5* mutations;
- Observation of the phenotypic effect of homozygous, compound heterozygous or heterozygous *COL4A3* or *COL4A4* mutations.

In each individual, data were collected on renal function and ultrastructure, audiologic function and ophthalmologic structure and function.

1.3. Nature of data

The qualitative data collected in the research study included words (interviews, hospital clinical files and questionnaires), drawings (pedigrees), images (audiograms)

and photographs (renal biopsies, skin biopsies, OCT). Molecular data was also classified qualitatively (mutation type, mutation status).

The quantitative data collected in the research project included age of participants, biochemical analytes in blood (plasma creatinine concentration, pCr) and urinalysis results (red blood cells (RBC) count and urine, protein excretion), and the estimated glomerular filtration rate (eGFR).

Ultimately, the research was quantitative, because, independently of the nature of the data collected, the analyses involved statistic treatment of data.

1.4. Origin of data

The primary data of the research project included (Figure 4):

- results of blood and urine laboratory analyses;
- images of audiologic exams (tonal and high frequency audiograms);
- photographs obtained during ophthalmologic examination;
- photographs of the GBM ultrastructure obtained on electron microscopy examination of kidney biopsies;
- mutations identified in each patient, i.e. data that were collected by implementation of the molecular genetics analysis of *COL4A5* gene and by purchasing the molecular genetic testing of *COL4A3* and *COL4A4* from DNA diagnostic laboratories in the United Kingdom and Germany.

Expected benefits of the primary data were to review, catalogue and organize the clinical data according to the criteria chosen in the research project, and to describe novel pathogenic *COL4A5*, *COL4A4* and *COL4A3* mutations in a population that had never been genetically screened for collagen IV-related nephropathies.

The secondary data (data that were originally collected elsewhere) included:

- data collected on patient's interviews;
- data collected from the patient's medical records, namely clinical diaries, pre-dialysis medical reports, referral information;
- data provided through a two-phase questionnaire filled by each patient's assistant physician.

Expected benefits of secondarily collected data were the reduction of costs and time savings; limitations are those inherent to retrospective data collection, including incompleteness and dispersal.

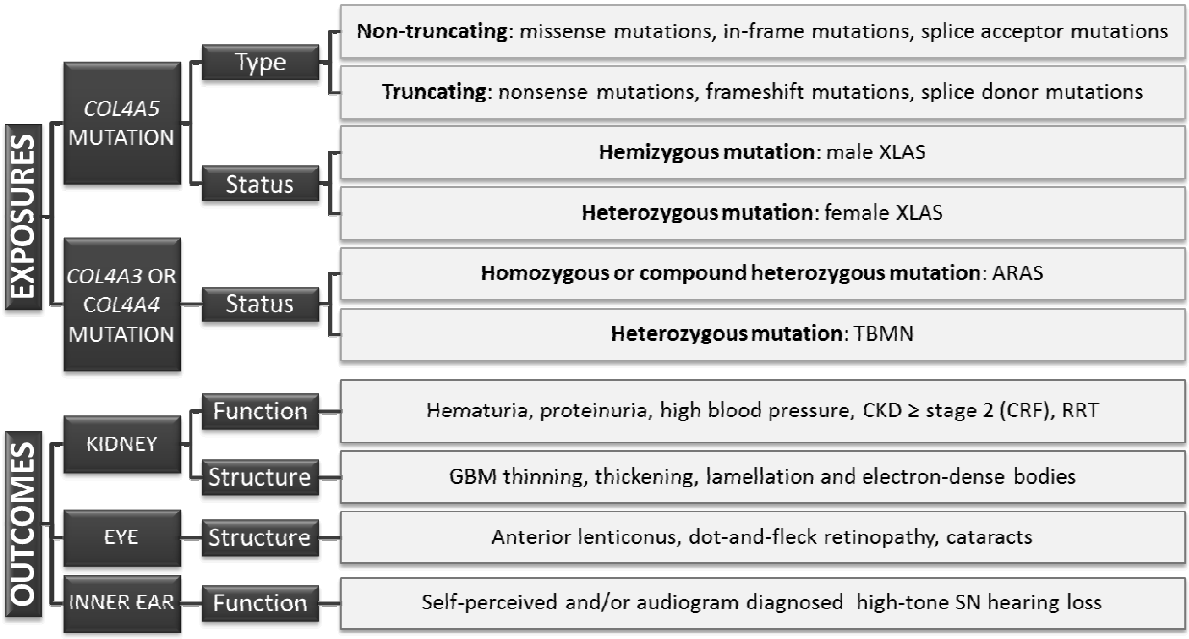


Figure 4. Overview of the molecular (exposures) and clinical (outcomes) variables.

2. Study design

An observational (descriptive) study followed by an analytic study was the selected research strategy. A retrospective multicentric cohort study was considered the most efficient to satisfactorily reach the established aims, which were to identify pathogenic *COL4A5*, *COL4A4* and *COL4A3* mutations, to perform the clinical characterization and to investigate the genotype-phenotype correlations in Portuguese individuals with clinical diagnosis of AS or TBMN. Multicentric enrolment, enabling the study of a random sample of patients from the whole country (or at least from several districts and different medical care settings) would enhance the generalizability (external validity) of the study. The retrospective nature of the study design was appropriate for a disease of low prevalence (Levy and Feingold 2000), with outcomes that progress during decades. Furthermore, follow-up time, effort and economic resources were effectively saved.

2.1. Clinical data collection and analysis

Both clinical data collection techniques and data analyses techniques were selected for this specific research strategy to meet the required scientific quality criteria.

Clinical data: Collection techniques

The clinical data collection techniques adopted in this research were: (i) clinical interviews with the patient and, when required, with family members (CHSJ, Porto, and University Hospitals, Coimbra); (ii) review of medical records (CHSJ, Porto, and University Hospitals, Coimbra); (iii) questionnaires with data retrieved from interviews and clinical files (Protocol P1); (iv) photographs of ultrastructure of GBM obtained by electron microscopy and of immunohistochemistry analysis of glomerular and EBM obtained by light microscopy.

The photographic and clinical archives of the kidney biopsy databases of the Pathology Department of CHSJ (Porto) and of the Unit of Renal Morphology of Curry Cabral Hospital (Lisboa) – which are the only two sites performing electron microscopy of kidney biopsies in Portugal –, were systematically reviewed in collaboration with their curator nephrologists, Dr. Susana Sampaio and Dr. Fernanda Carvalho, respectively. By this approach, a total of 26 patients from 22 unrelated families were eventually enrolled in the study cohort through the CHSJ. From 22 patients there were electron microscopy photographs of kidney biopsies available for review, allowing to characterize in detail the ultrastructural GBM pathology. In all biopsied patients, light microscopy examination had disclosed only non-specific features. In addition, electron microscopy photographs of the kidney biopsies of two patients enrolled through sites in the Lisbon region were reviewed at the Curry Cabral Hospital.

Questionnaires were requested at two different time points: (i) a first standardized questionnaire was fulfilled at enrollment; (ii) a customized second questionnaire, containing questions regarding incomplete answers or fields left blank in the first questionnaire, was fulfilled at a later time, with the objective of updating or detailing the clinical evaluation of participants, and reduce the missing data, since

some patients had renal, audiological and/or ophthalmical investigation requested after the enrollment in the research project. The structured questionnaires on each patient's phenotype was completed and retrieved by the participant's nephrologist and/or geneticist.

Clinical data: Analysis techniques

The data selected to answer the research question #2, i.e. to define the natural history of the disease in the study cohort were: (i) pedigree analysis (including family history of hematuria, ESRD or SNHL), which enabled inference of the inheritance pattern in each family; (ii) clinical and laboratory manifestations of renal disease, CKD progression and need for RRT, as evaluated by urinalysis (for ascertainement of occurrence and age at diagnosis of hematuria and proteinuria), blood analysis (serial pCr measurements for ascertainement of occurrence and age at diagnosis of CKD), assessement of blood pressure (for ascertainement of occurrence and age at diagnosis of hypertension), as well as age at onset of RRT; (iii) renal pathology, as specifically described by ultrastructural examination of kidney biopsies (ascertainement of occurrence and age at diagnosis of the typical GBM features of AS – thinning, thickening, lamellation and electron-dense bodies); (iv) history of hearing loss and of specialized otolaryngology investigation, namely the occurrence of self-perceived hearing loss and age at self-perception of hearing loss, and whether confirmation and characterization of SNHL by audiogram was obtained (following bilaterally normal otoscopic and tympanogram examinations); (v) presence of typical ocular signs, as specifically diagnosed by ophthalmoscopy, fundoscopy and OCT (ascertainement of occurrence and age at diagnosis of anterior lenticonus, dot-and-fleck maculopathy and cataracts) (Supplementary table S1 of manuscript 1).

The CKD-EPI equation (Levey, Stevens et al. 2009) or the four variable MDRD equation (Levey, Bosch et al. 1999) (calculators accessed online at http://mdrd.com/) were used, as appropriate, to compute eGFR from each available pCr value.

2.2. Molecular data collection and analysis

Molecular data: Collection techniques

The molecular data obtained in this study were generated by the following laboratory approaches: (i) polymerase chain reaction (PCR); (ii) direct DNA Sanger sequencing by capillary electrophoresis; (iii) multiplex ligation-dependent probe amplification (MLPA); (vi) fragment length analysis; (v) high-resolution X chromosome-specific microarray-based comparative genomic hybridization (array-CGH); (vi) quantitative PCR (qPCR). The molecular genetic analysis of the *COL4A5* gene was developed in the Department of Genetics of the Faculty of Medicine of the University of Porto (Porto, Portugal) (Protocol P2). Distinct methods were applied for identification of different types of mutation in *COL4A5* gene: PCR amplification and direct sequencing of genomic DNA for detection of point mutations and MLPA for detection of large rearrangements. The molecular genetic analyses of the *COL4A3* and *COL4A4* genes were performed in the DNA Laboratory, GSTS Pathology, Guy's & St Thomas' Hospital Foundation Trust (London, UK), in collaboration of with Hellen Storey and Prof. Frances Flinter, and in the Center for Nephrology and Metabolic Diseases (Weisswasser, Germany), in collaboration of Dr. Mato Nagel. The high-resolution array-CGH and qPCR were performed at the Human Genome Laboratory from the Department of Human Genetics of the VIB Center for the Biology of Disease (Leuven, Belgium), under the supervision of Prof. Guy Froyen.

Molecular data: Analysis techniques

The molecular data analysis techniques adopted to answer the research question #1 are the following theoretical knowledge and bioinformatic tools: (i) the pathogenicity of previously described pathogenic variants was supported by clinical and molecular data from the literature; (ii) the pathogenicity of the previously unreported variants was presumed from the type of mutation, and was also supported by intrafamilial genotype-phenotype correlations, following genotyping of at-risk first degree relatives of the probands, as well as the absence of the DNA sequence variant in 150 X-chromosomes of Portuguese healthy controls or in more

than 8,000 alleles from more than 4000 European-Americans unrelated individuals listed on the Exome Variant Server (Exome Variant Server); (iii) the pathogenicity of the previously unreported nonsense, splicing and frameshift variants was presumed from the type of mutation, with such mutations regarded as disease-causing; (iv) the possible impact of novel point mutations upon the structure and function of the corresponding protein was predicted *in silico* with the bioinformatic tools "MutationTaster" (http://www.mutationtaster.org/), "MutPred" (http://mutpred.mutdb.org/), "PolyPhen-2" (http://genetics.bwh.harvard.edu/pph2/) and SNPs&GO (http://snps.biofold.org/snps-and-go/snps-and-go.html) (Calabrese, Capriotti et al. 2009; Adzhubei, Schmidt et al. 2010; Thusberg, Olatubosun et al. 2011; Adzhubei, Jordan et al. 2013); (v) the NNSPLICE software (http://www.fruitfly.org/seq_tools/splice.html) was used to predict the effect of intronic variants on mRNA splicing; (vi) to confirm whether the mutation identified in the proband was inherited or *de novo*, molecular testing of the proband's parents was performed, whenever possible; (vii) genotyping using four microsatellite polymorphic markers flanking the *COL4A5* gene (DXS1120, DXS1105, DXS1210, DXS456) was performed to track its inheritance across the generations of the pedigree available for study and to test whether two apparently unrelated families with a same *COL4A5* mutation shared a common haplotype.

Although it had been initially planned to perform mRNA extraction and cDNA synthesis, as a confirmatory method of the presumably pathogenic splicing effect of variants in the nucleotide sequence on the aminoacid sequence of the COL4A5, COL4A4 and COL4A3 proteins, this method was replaced by bioinformatic tools analysis (Thusberg, Olatubosun et al. 2011). Besides the consistency of prediction pathogenic mutations in the genomic DNA into proteins, the use of bioinformatic tools was also justified to additionally overcome the refusal of patients to undergo skin or kidney biopsy and save time and financial resources required to implement and perform the analysis of mRNA.

2.3. Genotype-phenotype correlation: Analysis techniques

The analysis techniques adopted to determine genotype-phenotype associations were: (i) quasi-statistics (interpretation of frequencies); (ii) bivariate

statistical analysis (two variables). The predictor (independent) variable selected for the genotype-phenotype correlation analyses was the mutation status: hemizygosity, homozygosity, compound heterozygosity and heterozygosity. Pathogenic *COL4A5* mutations were also divided into two groups, according to its type: non-truncating mutations (glycine substitutions, in-frame mutations, splice acceptor mutations) and truncating mutations (nonsense mutations, frameshift mutations, splice donor mutations, and large rearrangements). Clinical data were censored at the date the last information collection. Nominal and numeric dependent variables included in the four main groups of outcomes – (i) of renal function; (ii) of renal structure; (iii) audiological; (iv) ophthalmological – were analyzed using non-parametric tests: Mann-Whitney test for comparison of continuous unpaired data, Fisher's exact test and chi-square test for comparison of proportions between the two independent groups and log rank test for comparison of time to an event between two groups. Chi-square for trend was used to test for the presence of a trend in the *COL4A5*/*COL4A4*/*COL4A3* mutation detection rate as related to the number of diagnostic criteria identified in each patient. The 95% confidence intervals (95%CI) for prevalence estimates were computed by the modified Wald method (Agresti and Coull 1998). The data were analyzed using the following statistics software packages: IBM® SPSS® Statistics version 21 (SPSS IBM; New York, NY, U.S.A) and GraphPad Prism® version 5 (GraphPad Software, Inc.; La Jolla, CA, U.S.A.).

Chapter 3 Results

In the first part of this chapter, the results of the patient enrollment process are presented and discussed. The second part consists of the scientific manuscripts that have been generated by this research project. Manuscript #1 reports the results of the comprehensive mutational screening of *COL4A5* that was carried out as first tier approach to genetic diagnosis in all probands with a clinical diagnosis or suspicion of AS, who belonged to families in whom X-linked inheritance could not be excluded by pedigree analysis. Manuscript #2 describes a novel genotype-phenotype correlation identified in a family with ATS-DL, raising the question that this disorder may not be a contiguous gene deletion syndrome, as it is usually described. Manuscript #3 reports the results of the mutational analyses of *COL4A3* and *COL4A4* genes by Sanger sequencing, in patients without a pathologic *COL4A5* mutation identified on the first tier genetic study, or who belonged to families with unequivocal evidence of autosomal inheritance of kidney disease.

1. Cohort assembly: Results of the recruitment process

Nationwide patient enrollment through public hospitals

In order to guarantee the largest possible nationwide enrollment of patients, 28 nephrology departments and two medical genetics departments of 29 public hospitals – located in fifteen of the eighteen districts of mainland Portugal and in the two autonomous Atlantic islands regions of Azores and Madeira –, were formally invited to participate in this study. Kidney patients living in the three mainland districts where the local hospitals do not have nephrology services are referred to hospitals in neighboring districts but, for the purpose of this study, the geographic origin of each enrolled patient was defined as the district of residence.

Overall, 51 family probands and 102 at-risk relatives were eventually enrolled in this study through 11 nephrology departments and two medical

genetics departments of 12 public hospitals, located in 11 of the mainland districts and in the autonomous region of Madeira. (These data are summarized in Table VI; see also the Supplementary figure S3 of manuscript #1). Of note, the two medical genetics departments contributed with 25 probands and 53 at-risk relatives, an observation that underscores the importance of the collaboration between nephrologists and clinical geneticists for effective delivery of comprehensive medical care to patients and families with hereditary kidney diseases, such as the collagen IV-related nephropathies.

As shown in Table VI, the major single institutional contributors to the enrollment of cases in the study cohort were the Department of Human Genetics and the Department of Nephrology of CHSJ. The CHSJ is a large tertiary care and university hospital whose catchment area for those two specialties is the northwest region of Portugal, comprising part of the district of Porto and the districts of Braga and Viana do Castelo, with a global resident population of about 2,500.000 people. Twenty three probands were enrolled through the clinic of hereditary kidney diseases of the Department of Human Genetics, and respectively the peritoneal dialysis and the kidney transplant clinics of the Department of Nephrology further contributed with 2 and 3 probands.

As summarized in Table VII, the nephrology departments of three major hospitals – in Lisbon (Curry Cabral Hospital), Porto and Azores –, originally expressed their interest in participating in this study but ultimately did not refer any patients for genetic analyses. Some patients with AS followed at the Curry Cabral Hospital had been previously enrolled in the ECASCA study and did not consent to participate in this research project. Eight of the invited nephrology departments (29%) replied that they were not currently caring for any patients with a diagnosis of AS and other 5 (18%) did not formally reply to the invitation, three of them having later informed the research team that they had no patients diagnosed with AS. In the southern districts of Portalegre, Beja and Évora there were no patients diagnosed with AS known to nephrologists of the local public hospitals.

Table VI. Number of participants enrolled in public hospitals, by geographic district (n=153).

District	***Hospital***	***Local researcher***	***Probands***	***Family members***	***Total of participants***
Vila Real	Hospital São Pedro, Vila Real	Dr. Rui Castro	5	12	**17**
Braga	Hospital de Braga	Dr. Carlos Soares	3	4	**7**
Porto	Centro Hospitalar de São João, Porto	Dr. João Paulo Oliveira	28	48	**79**
	Hospitala de Vila Nova de Gaia	Dr. Joaquim Seabra	3	0	
Coimbra	Hospitais da Universidade de Coimbra	Dr. Jorge Pratas e Sousa, Dr. Luís Freitas	2	10	**12**
Viseu	Hospital de São Teotónio, Viseu	Dr.ª Tânia Couto Sousa, Dr. Jesus Garrido	1	0	**1**
Castelo Branco	Hospital Amato Lusitano, Castelo Branco	Dr. Rui Alves Filipe	1	0	**1**
Lisboa	Hospital Santa Maria, Lisboa	Dr. António Gomes da Costa	1	1	**18**
	Hospital Dona Estefânia, Lisboa	Dr.ª Márcia Rodrigues	2	14	
Setúbal	Hospital Garcia de Orta, Almada	Dr. Fernando Teixeira e Costa	1	3	**4**
Faro	Hospital de Faro	Dr.ª Elsa Morgado	1	2	**3**
R.A. Madeira	Hospital dos Marmeleiros, Funchal	Dr. José Augusto Araújo	3	8	**11**
Total			51	102	**153**

R.A.: autonomous region.

Table VII. Invited non-participating hospitals in the research study, by geographic district (n=17).

District	***Hospital***	***Local Contact***	***Patients with Alport syndrome***
Bragança	Hospital de Bragança	Dr.ª Fátima Ramos	Unknown (no reply)
Braga	Hospital de Fafe	Dr. Joaquim Pinheiro	Unknown (no reply)
Porto	Hospital de Santo António, Porto	Dr.ª Manuela Almeida	Yes (number not determined)
	Instituto Português de Oncologia, Porto	Dr. Alfredo Loureiro	None
	Hospital Pedro Hispano, Matosinhos	Dr. José Maximino	None
Aveiro	Hospital Infante D. Pedro, Aveiro	Dr.ª Carmen do Carmo	None
Coimbra	Hospital dos Covões, Coimbra	Dr. Armando Carreira	None
Santarém	Hospital de Santarém	Dr. João Bispo	Unknown (no reply)
	Hospital Rainha Santa Isabel, Torres Novas	Dr. Sequeira Andrade	Unknown (no reply)
Lisboa	Hospital Curry Cabral, Lisboa	Dr. Fernando Nolasco	Yes (number not determined)
	Hospital Santa Cruz, Carnaxide	Dr. José Diogo Barata	Unknown (no reply)
	Hospital Professor Doutor Fernando Fonseca, Amadora/Sintra	Dr. Luís Incháustegui	None
Setúbal	Hospital de São Bernardo, Setúbal	Dr. José Vinhas	Unknown (no reply)
Portalegre	Hospital Dr. José Maria Grande, Portalegre	Dr. António Sousa	None
Beja	Hospital José Joaquim Fernandes, Beja	Dr. Carlos Pires	None
Évora	Hospital Espírito Santo, Évora	Dr. João Aniceto	None
R.A. Açores	Hospital de Santo Espírito, Angra do Heroísmo	Dr.ª Lurdes Dias	Yes (number not determined)

R.A.: autonomous region.

Enrollment of patients from hemodialysis clinics

For a disease like AS, progressing to ESRD at relatively early ages in a substantial proportion of patients, it is expected that a large fraction of them be on RRT at any point in time. In Portugal, about 64% of the more than 17,500 patients on RRT are on hemodialysis; only about 6.6% of the patients on dialytic RRT are on peritoneal dialysis (http://www.spnefro.pt/comissoes_gabinetes/Gabinete_registo_2012/registo_2012.pdf); and only 10.7% of the patients on chronic hemodialysis are cared for in public hospitals, the remainder being treated at private for-profit dialysis facilities. Therefore, despite the risk of ascertainement bias towards more severe AS clinical phenotypes, case identification through private hemodialysis clinics would be an important step in the research plan, to ensure a thorough nationwide ascertaiment of families. To this end, the two major networks of hemodialysis clinics operating in Portugal – Diaverum and NephroCare / Fresenius Medical Care – were formally invited to collaborate. In addition, the Portuguese Society of Nephrology (SPN) publicized this research project on their internet site. Of the two hemodialysis networks, only NephroCare ultimately approved their participation in the study. Taking into consideration the geographic distribution of the hemodialysis clinics of the two networks, the non-participation of Diaverum may have affected a thorough ascertainment of AS patients in the central-western region of the country. On the other hand, an AS patient on chronic hemodialysis at a Diaverum facility in Lisbon, self-referred to the research team after having read the announcement posted at the SPN website and was enrolled as the proband of her family.

The results of patient enrollment through hemodialysis clinics are summarized in Table VIII: a total of 14 family probands were identified in 12 dialysis clinics, from northern and central mainland districts. The importance of having involved dialysis companies in this study is evident in the following results: (i) the probands of all the families identified in the districts of Santarém and Guarda were enrolled through NephroCare dialysis centers; (ii) even in districts where affected families had been identified through public hospitals, additional probands were enrolled through private hemodialyis clinics; (iii)

several affected relatives of probands enrolled through public hospitals were recruited in private hemodialysis clinics; (iv) no patients diagnosed with AS were receiving dialysis treatment in the southern districts of Portalegre, Beja and Évora, confirming the findings in the local public hospitals.

Table VIII. Number of participants enrolled in hemodialysis clinics, by geographic district (n=25).

District	***Dialysis Clinic***	***Local researcher***	***Probands***	***Family members***	***Total***
Braga	Nephrocare, Braga	Dr. António Castro Henriques	1	1	**2**
Porto	Uninefro, Santo Tirso	Dr.ª Isabel Tavares	2	0	**4**
	Diaverum, Paredes	Dr.ª Ana Oliveira / Dr.ª Liliana Pinho	0	2	
Coimbra	Nephrocare, Coimbra	Dr. Jorge Pratas e Sousa	0	1	**1**
Viseu	Nephrocare, Viseu	Dr. Rui Alves / Dr. Carlos Miguel Almeida Botelho	1	0	**1**
Guarda	Nephrocare, Guarda	Dr.ª Tânia Couto Sousa / Dr. Bernardo Faria	1	4	**5**
Castelo Branco	Nephrocare, Covilhã	Dr. José Montalban / Dr.ª Ana Bernardo	1	0	**1**
Santarém	Nephrocare, Santarém	Dr. Fernando Neves / Dr. Adelino Carvalho	3	1	**6**
	Nephrocare, Montijo	Dr. Francisco Teixeira de Sousa / Dr.ª Joana Felgueiras	1	1	
Lisboa	Nephrocare, Lumiar	Dr. Pedro Ponce	1	0	**5**
	Nephrocare, Restelo	Dr.ª Maria Augusta Gaspar / Dr. Luís Filipe Carvalho	2	1	
	Diaverum, Lumiar	Dr.ª Cristina Pinto Abreu	1	0	
Total			14	11	**25**

2. *Clinical and molecular characterization of patients with clinical diagnosis or suspicion of Alport syndrome in Portugal*

2.1. Patients with pathogenic COL4A5 *and* COL4A6 *mutations*

2.1.1. Collagen type IV-related nephropathies in Portugal: spectrum of pathogenic *COL4A5* mutations and clinical characterization of 22 families (Manuscript 1)

Manuscripts 1 and 3 were submitted simultaneously to the Journal of Medical Genetics, on the 11th April 2014 (Manuscript 1 ID: jmedgenet-2014-102471).

Collagen type IV-related nephropathies in Portugal: spectrum of pathogenic *COL4A5* mutations and clinical characterization of 22 families

Maria João Nabais Sá[1,2], Susana Sampaio[2,3], Ana Oliveira[3], Susana Alves[1], Carla Pinto de Moura [4,5], Sérgio Estrela Silva[6], Rui Castro[7], José Augusto Araújo[8], Márcia Rodrigues[9], Fernando Neves[10], Joaquim Seabra[11], Carlos Soares[12], Maria Augusta Gaspar[13], Isabel Tavares[2,14], Luís Freitas[15], Tânia Couto Sousa[16,17], Castro Henriques[18], Fernando Teixeira e Costa[19], Elsa Morgado[20], Francisco Teixeira Sousa[21], Jorge Pratas e Sousa[15,22], António Gomes da Costa[23], Rui Filipe[24], Jesus Garrido[16], José Montalban[25], Pedro Ponce[26], Rui Alves[27], Bernardo Faria[17], Fernanda Carvalho[28], Manuel Pestana[3,29], Filipa Carvalho[1], João Paulo Oliveira[1,2,5]

[1] Department of Genetics, Faculty of Medicine, University of Porto, Porto, Portugal

[2] Unit of Research and Development of Nephrology (FCT-725), Faculty of Medicine, University of Porto, Porto, Portugal

[3] Department of Nephrology, Hospital de São João, Porto, Portugal

[4] Department of Otolaryngology, Hospital de São João, Porto, Portugal

[5] Medical Genetics Outpatient Clinic, Hospital de São João, Porto, Portugal

[6] Department of Ophthalmology, Hospital de São João, Porto, Portugal

[7] Department of Nephrology, Centro Hospitalar de Trás-os-Montes e Alto Douro, Vila Real, Portugal

[8] Department of Nephrology, Hospital dos Marmeleiros, Funchal, Portugal

[9] Department of Genetics, Hospital Dona Estefânia, Lisboa, Portugal

[10] Dialysis Clinic of Santarém, NephroCare-Portugal, Santarém, Portugal

[11] Department of Nephrology, Centro Hospitalar Vila Nova de Gaia/Espinho, Vila Nova de Gaia, Portugal

[12] Department of Nephrology, Hospital de Braga, Braga, Portugal

[13] Dialysis Clinic of Restelo, NephroCare-Portugal, Lisboa, Portugal

[14] Dialysis Clinic of Santo Tirso, Uninefro, Santo Tirso, Portugal

[15] Department of Nephrology, Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal
[16] Department of Nephrology, Hospital de São Teotónio, Viseu, Portugal
[17] Dialysis Clinic of Guarda, NephroCare-Portugal, Guarda, Portugal
[18] Dialysis Clinic of Braga, NephroCare-Portugal, Braga, Portugal
[19] Department of Nephrology, Hospital Garcia de Orta, Almada, Portugal
[20] Department of Nephrology, Hospital de Faro, Faro, Portugal
[21] Dialysis Clinic of Montijo, NephroCare-Portugal, Montijo, Portugal
[22] Dialysis Clinic of Coimbra, NephroCare-Portugal, Coimbra, Portugal
[23] Department of Nephrology, Hospital de Santa Maria, Lisboa, Portugal
[24] Department of Nephrology, Hospital Amato Lusitano, Castelo Branco, Portugal
[25] Dialysis Clinic of Covilhã, NephroCare-Portugal, Covilhã, Portugal
[26] Dialysis Clinic of Lumiar, NephroCare-Portugal, Lisboa, Portugal
[27] Dialysis Clinic of Viseu, NephroCare-Portugal, Viseu, Portugal
[28] Unit of Renal Morphology, Department of Nephrology, Hospital Curry Cabral, Lisboa, Portugal
[29] Nephrology and Infectious Diseases Research and Development Group – INEB, University of Porto, Porto, Portugal

Corresponding author: Maria João Nabais Sá, Department of Genetics, Faculty of Medicine, University of Porto, Alameda Prof. Hernâni Monteiro, 4200 - 319 Porto, Portugal. E-mail: mjs.jano@gmail.com. Telephone: +351965546341. Fax: +351225513648.

Key words: X-linked Alport syndrome, *COL4A5*, molecular genetics, genotype-phenotype correlations.

ABSTRACT

Background

Alport syndrome (AS) is a disorder of basement membranes caused by mutations in the genes encoding the α3 (*COL4A3*), α4 (*COL4A4*) or α5 (*COL4A5*) chains of collagen IV. Organ-selective expression of those α(IV)-chains explains the AS phenotype including hematuria, chronic renal failure (CRF), sensorineural hearing loss (SNHL) and ocular abnormalities. Mutations in the X-linked *COL4A5* gene have been identified in 80-85% of the families (XLAS).

Patients and Methods

Sixty-five apparently unrelated families with clinical diagnosis of AS were enrolled in this study, aiming to describe the genetic pathology of AS in Portugal. *COL4A5* molecular analysis was carried out by Sanger sequencing and multiplex-ligation probe amplification in 60 probands who showed X-linked or unclear inheritance patterns.

Results

Twenty-two out of the 60 probands (37%) had a pathogenic *COL4A5* mutation, of which 12 (57%) are novel and 9 (43%) were previously described. Two families shared a missense mutation, with identical microsatellite haplotypes. In two probands (9%), the *COL4A5* mutation proved to be *de novo*. Males had more severe and earlier renal and extrarenal complications, but microscopic hematuria was a constant finding irrespective of patient gender. Truncating mutations were associated with a younger age of onset of SNHL in males, and with a higher risk of CRF and SNHL in females.

Discussion

Diagnosis of XLAS was genetically confirmed in a lower than expected proportion of Portuguese families. Conclusive interpretation of these data requires reliable information about the prevalence of autosomal AS, due to pathogenic *COL4A3* or *COL4A4* mutations, in the remaining families.

INTRODUCTION

The eponym Alport syndrome (AS) refers to a heterogeneous group of progressive familial hematuric nephropathies caused by pathogenic mutations in any of the α5, α4 or α3 type IV collagen chains.(Kashtan 1993-2013; Kashtan 1999; Kashtan 2001; Tryggvason and Patrakka 2009; Kruegel, Rubel et al. 2013) Since these α(IV)-chains are the major components of the protein scaffold of the mature glomerular basement membrane (GBM) in the human kidney,(Hudson, Tryggvason et al. 2003; Khoshnoodi, Pedchenko et al. 2008) genetic defects that either impair their synthesis or extracellular assembly in specific collagenous networks ultimately lead to chronic kidney disease (CKD),(Cosgrove 2012; Kruegel, Rubel et al. 2013) by compromising the normal role of the GBM as a specialized plasma filtration barrier. The structural abnormalities of the GBM observed on electron microscopy (EM), particularly the thickening, splitting and fragmentation of the lamina densa, are an important diagnostic clue of AS.(Haas 2009) Overall, about 0.6% of the patients with end-stage renal failure (ESRF) starting renal replacement therapy (RRT) in Europe had AS, of which 22.8% were females and 6.5% were children aged less than 15 years.(Rigden, Mehls et al. 1996) Besides CKD, high-tone sensorineural hearing loss (SNHL) and a spectrum of ocular lesions affecting the lens, the retina, and the cornea are additional, though less common manifestations of AS.(Crawfurd 1988; Flinter 1997; Kashtan 2001; Savige and Colville 2009; Tryggvason and Patrakka 2009; Kruegel, Rubel et al. 2013) In patients with hematuria of uncertain etiology, the following set of criteria enable the diagnosis AS:(Flinter, Cameron et al. 1988) (*i*) family history of macro/microscopic hematuria and/or of chronic renal failure (CRF); (*ii*) EM evidence of AS on renal biopsy; (*iii*) high tone SNHL; (*iv*) characteristic ophthalmological signs (anterior lenticonus and/or white macular flecks). When the patient, or the proband and other affected family members between them, fulfill at least three of those criteria, the clinical diagnosis of AS can be confidently established.

Type IV collagen is exclusive of basement membranes (BM) and comprises six genetically distinct isoforms of α(IV)-chains, designated α1(IV) through α6(IV). The α(IV)-chains form three different types of triple-helical protomers – α1:α1:α2(IV), α3:α4:α5(IV) and α5:α5:α6(IV) – that combine with each other, giving rise to complex collagenous networks that provide structural support to BM and influence cell

adhesion, migration, and differentiation.(Hudson, Tryggvason et al. 2003; Khoshnoodi, Pedchenko et al. 2008) Whereas the α1:α1:α2(IV) are ubiquitously present in all mammalian BM, the two other heterotrimers are selectively expressed in different tissues.(Hudson, Tryggvason et al. 2003; Khoshnoodi, Pedchenko et al. 2008) The presence of α3:α4:α5(IV) in the GBM and in BM of the ear and in the eye explains the phenotype of AS. Furthermore, the α3:α4:α5(IV) protomers, which are more resistant to increased intraglomerular pressure and to proteolytic degradation, gradually replace the α1:α1:α2(IV) heterotrimers in the GBM, after early embryonic development. This is a critical step of GBM adaption to the postnatal glomerular hemodynamic conditions and to the increased exposure to serum proteases.(Hudson, Tryggvason et al. 2003; Khoshnoodi, Pedchenko et al. 2008) The human α(IV)-chain genes (*COL4A1* through *COL4A6*) have similar genomic structures and are arranged pairwise in head-to-head orientation, *COL4A1-COL4A2* on the long arm of chromosome 13, *COL4A3-COL4A4* on the long arm of chromosome 2 and *COL4A5-COL4A6* on the long arm of chromosome X.(Khoshnoodi, Pedchenko et al. 2008)

About 80-85% of AS families reportedly have the classical X-linked form of the disease (XLAS; MIM#301050), caused by *COL4A5* mutations.(Kashtan 1999; Kashtan 2001; Tryggvason and Patrakka 2009; Kruegel, Rubel et al. 2013) XLAS is characterized by more severe phenotype and worse prognosis in males.(Jais, Knebelmann et al. 2000; Jais, Knebelmann et al. 2003) The remainder of cases, which are due to mutations in *COL4A43* and/or *COL4A4*, are inherited as autosomal recessive (ARAS; MIM#2013780) or autosomal dominant (ADAS; MIM#104200) disorders. In families with ARAS, males and females are equally affected and the kidney disease is similar to that observed in males with XLAS.(Tryggvason and Patrakka 2009) ADAS is usually quoted as the rarest form of AS,(Rumpelt 1980; Kashtan 1999; Kashtan 2001; Tryggvason and Patrakka 2009) comprising less than 5% of the patients, but this concept has been recently challenged.(Fallerini, Dosa et al. 2013) Besides XLAS, mutations affecting *COL4A5* also occur in AS with diffuse leiomyomatosis (ATS-DL; MIM#308940)(Garcia Torres and Guarner 1983) and in AS with mental retardation, midface hypoplasia and elliptocytosis (AMME;

MIM#300194),(Jonsson, Renieri et al. 1998) two rare syndromes caused by microdeletions involving *COL4A5* and adjacent genomic regions.

More than 700 disease-causing *COL4A5* mutations, most of them unique to single families, have already been described, and about 400 more are estimated to be unpublished.(Savige, Ars et al. 2013) In "The Human Gene Mutation Database"(Stenson, Mort et al. 2013) (HGMD®; http://www.hgmd.cf.ac.uk/ac/gene.php?gene=COL4A5, last accessed on March 1, 2014), missense and nonsense substitutions respectively account for 41.7% and 6.7% of all the reported *COL4A5* mutations; substitutions that affect the normal splicing of *COL4A5* mRNA account for 16.8%; small insertions, deletions and insertions/deletions, involving 20 base pairs (bp) or less, account for 22.2%; deletions and insertions/duplications greater than 20 bp and gross complex rearrangements account for 12.6%. Glycine substitutions in the conserved Gly-Xaa-Yaa repeat sequence of the α5(IV) collagenous domain are the most frequent type of pathogenic *COL4A5* mutation, variably accounting for 56.2-95.2% of the missense mutations identified in several large European(Jais, Knebelmann et al. 2000; Hanson, Storey et al. 2011) and North-American(Bekheirnia, Reed et al. 2010) cohorts. Between 14.3-26.1% of the pathogenic *COL4A5* mutations are nonsense substitutions or small frameshifting deletions or insertions that result in premature stop codons. In hemizygous males, large deletions and truncating mutations are associated with earlier progression to ESRF, higher risk of developing hearing loss before age 30 years and higher probability of expressing anterior lenticonus.(Jais, Knebelmann et al. 2000; Gross, Netzer et al. 2002; Bekheirnia, Reed et al. 2010) Mutations located closer to the 5' end of the gene are also associated with more severe renal phenotype.(Bekheirnia, Reed et al. 2010) Contrastingly, no significant genotype-phenotype correlations could be identified in heterozygous females.(Jais, Knebelmann et al. 2003)

In males with XLAS,(Jais, Knebelmann et al. 2000) the presenting clinical manifestation was hematuria in 81% and isolated proteinuria in 12.5%; during the course of the disease, a single or recurrent episodes of gross hematuria occurred in 62% of the patients and proteinuria was found in 95%; and all patients who had a kidney biopsy examined by EM showed either typical or suggestive ultrastructural

GBM abnormalities. Between 88.5-94% of the patients had positive family history, with 76-84% of the families reporting history of ESRF.(Jais, Knebelmann et al. 2000; Bekheirnia, Reed et al. 2010; Hanson, Storey et al. 2011) The prevalence of extrarenal diagnostic criteria of AS ranged between 67-89% for clinical hypoacusia and/or audiologically confirmed SNHL and 30-40% for ocular lesions (lenticonus and/or maculopathy and/or congenital or early onset cataract).(Jais, Knebelmann et al. 2000; Bekheirnia, Reed et al. 2010; Hanson, Storey et al. 2011) At least 85% of the females with XLAS have microscopic hematuria.(Crawfurd 1988; Flinter 1997; Kashtan 2001) Since microscopic hematuria may be intermittent, repeat testing may be necessary before expression of the disease can be excluded, especially in a female. At age 40 years, the probability of having ESRF or deafness is, respectively, 90% and 80% in males and 12% and 10% in females.(Jais, Knebelmann et al. 2000; Jais, Knebelmann et al. 2003) Overall, 30% of the heterozygous females develop ESRD by age 60 years and an additional 10% progress to ESRD during the next two decades of life.(Jais, Knebelmann et al. 2003)

Data about the genetics of AS in Portugal are scarce and result of a small contribution to an European Concerted Action.(Jais, Knebelmann et al. 2000; Jais, Knebelmann et al. 2003) Our major goals were to describe the genetic pathology and genotype-phenotype correlations in a large cohort of Portuguese families with AS, identified through systematic ascertainment all over the country.

PATIENTS AND METHODS

Enrollment of Patients and Families

Between 1/January/2009 and 30/June/2012, patients with clinical diagnosis of AS who were residing in Portugal and consented to participate, as well as their affected or at-risk relatives who accepted or requested genetic evaluation, were enrolled in this study. In order to ensure the broadest possible geographic catchment area, all nephrology departments and two major medical genetics departments from 29 public hospitals throughout the country, as well as the two biggest private dialysis clinics networks operating in Portugal, were formally invited to participate, and the study was publicized through the internet site of the Portuguese Society of Nephrology.

As operationally defined herein, “genetic probands” were the first patient of each apparently unrelated family who was enrolled in this study and underwent molecular genetics analysis. The clinical diagnosis of AS in genetic probands was based on the set of criteria proposed by Flinter and colleagues,(Flinter, Cameron et al. 1988) for patients with hematuria of unknown etiology. The diagnosis of AS was considered clinically “established” if the genetic proband, or the genetic proband and other affected family members between them, fulfilled at least three of the diagnostic criteria. Genetic probands presenting with less than three diagnostic criteria were considered to have a “probable” diagnosis of AS, being also eligible for mutational analysis. The observation of typical ultrastructural GBM abnormalities on kidney biopsy was regarded as highly suggestive of the diagnosis of AS.(Hanson, Storey et al. 2011)

The study protocol and informed consent documents were reviewed and approved by the Health Ethics Commission of São João Hospital Centre (CHSJ, Porto; Portugal).

Clinical questionnaire and definitions

Specific clinical data from each participant, as well as the relevant details of their family histories, were collected by means of a standardized questionnaire. Clinical outcomes assessed in the questionnaire and the corresponding definition and measurement method are described in Supplementary table S1. Baseline clinical assessment was performed at enrollment by the referring physician, either a nephrologist or a clinical geneticist. Additional data were retrospectively collected by review of archive medical records and pathology reports. In some cases, information on the family history was completed by interviewing patient’s relatives.

Molecular genetics analyses

Genomic DNA of the participants was extracted from whole blood samples using a commercial kit (Citomed; Lisbon, Portugal), and aliquots kept refrigerated at 4ºC until used for molecular analysis. Blood sampling from patients on hemodialysis was done at the beginning of a dialysis session, before the administration of heparin. Blood samples drawn at the central laboratory (Genetics Department; Faculty of Medicine, University of Porto – FMUP; Portugal) were processed for DNA extraction within 24 hours; blood samples drawn at other sites were sent to the central laboratory by

priority mail, at room temperature, and processed for DNA extraction no later than 72 hours following blood collection.

In the genetic probands, all 53 exons of *COL4A5* as well as their corresponding intronic flanking regions were polymerase chain reaction (PCR)-amplified using previously described intronic primers and conditions (see Supplementary table S2).(Martin, Heiskari et al. 1998) PCR products were automatically sequenced in forward and reverse directions in an ABI Prism® 310 Genetic Analyzer (Applied Biosystems; Foster City, CA, USA), using the BigDye® Terminator v3.1 Cycle Sequencing kit (Applied Biosystems), and the corresponding electropherograms were visually checked for sequence variants. Subsequent screening for a known mutation in affected or at-risk probands' relatives was performed by direct analysis of relevant exon.

The GenBank® identifier NM_000495.3 (National Center for Biotechnology Information; Bethesda, MD, USA; http://www.ncbi.nlm.nih.gov/nuccore/), corresponding to transcript variant 1 of *COL4A5*, that lacks the two alternatively spliced exons located within intron 41, was used for reporting the *COL4A5* variants. In order to facilitate the comparison of our data with those of previous reports, the alternatively spliced exons are respectively designated as 41A and 41B, according to their original nomenclature.(Guo, Van Damme et al. 1993; Martin and Tryggvason 2001) Therefore, in comparison to the GenBank® *COL4A5* genomic reference sequence NG_011977.1, where all the exons are numbered sequentially, the exon identifiers beyond exon 41 differ by two, the last one being numbered 51 instead of 53.

The SALSA P191-B1/192-B1 Alport Multiplex Ligation-dependent Probe Amplification (MLPA) commercial kit (MRC-Holland; Amsterdam, The Netherlands), which includes probes for 48 of the 53 exons of *COL4A5* and for the first three exons of *COL4A6* (originally designated 1', 1 and 2(Sugimoto, Oohashi et al. 1994)), was used to detect large deletions/duplications involving *COL4A5* and the 5' region of *COL4A6*. The MLPA assay was performed according to the recommendations of the manufacturer (available at http://www.mrc-holland.com).

"The Human Gene Mutation Database" (HGMD®; accessed at http://www.hgmd.cf.ac.uk/ac/gene.php?gene=COL4A5), The "Leiden Open Variation

Database" (accessed at https://grenada.lumc.nl/LOVD2/COL4A/variants.php?action=search_unique&select_db=COL4A5), and the Alport (*COL4A5*) Mutation Database of the ARUP Scientific Resource for Research and Education, University of Utah, USA (accessed at http://www.arup.utah.edu/database/Alport/Alport_display.php) were last checked on the March 1, 2014 for previous reports of the *COL4A5* sequence variants identified in this study.

Novel DNA sequence variants were considered pathogenic if: (*i*) the same DNA variant was carried by affected relatives of the proband but was not identified in first degree at-risk relatives with no evidence of AS; (*ii*) the DNA sequence variant was not detected in a control sample of X-chromosomes (n=235); (*iii*) exonic single nucleotide substitutions were predicted to be disease-causing by *in silico* analyses (see Table I for details); (iv) intronic or exonic single nucleotide substitutions that might affect mRNA splicing were predicted to be pathogenic by *in silico* analyses (see Table I for details).

In apparently sporadic cases, parental genotyping was carried out, whenever possible, to confirm whether the *COL4A5* mutation identified in a proband was inherited or *de novo*. In cases where the same mutation was identified in apparently unrelated probands, the microsatellite polymorphic markers DXS1120, DXS1105, DXS1210, DXS456, flanking the *COL4A5-COL4A6* genes, were used for haplotype analysis as previously described.(Srivastava, McMillan et al. 1999; Tazon-Vega, Ars et al. 2007)

Whenever the comprehensive mutational analyses of *COL4A5* did not reveal a known or presumed pathogenic mutation, the patient was enrolled in a second tier study screening for *COL4A3* and *COL4A4* mutations. Scanning for mutations in the latter two genes was additionally performed in patients with novel *COL4A5* variants of uncertain pathogenicity.

Gender comparisons and genotype-phenotype correlation analysis

For this purpose, the *COL4A5* mutations were classified as non-truncating or truncating, as formerly proposed.(Gross, Netzer et al. 2002) Demographic and clinical data were censored at the time they were collected for each patient and the relevant outcomes were compared between genders and between carriers of

truncating or non-truncating mutations within each gender. In order to minimize the ascertainment bias, comparisons were made in a larger “phenotyping cohort”, including not only the genetic probands (“genotyping cohort”) but also their affected first or second degree relatives either previously identified or prospectively diagnosed on family screening. Parametric and non-parametric statistics were used, as appropriate, for comparisons of demographic variables and clinical outcomes. The data were analyzed with the IBM SPSS Statistics software version 21 (IBM Corporation; Armonk, NY, USA).

RESULTS

Sixty-five unrelated genetic probands with clinical diagnosis of AS were enrolled in this study from all over the country (Supplementary figure S3). Comprehensive molecular analyses of *COL4A5* was carried out in 33 males (median age: 38 years) and 26 females (median age: 36 years) but was limited to the relevant exon in a 5-year-old boy in whom a pathogenic missense *COL4A5* had already been identified as part of his clinical diagnostic workup. In the remainder five cases, the family history was indicative of autosomal inheritance and mutational analysis of *COL4A5* was skipped [manuscript submitted to the JMG simultaneously]. In addition to hematuria, respectively 5 (8%), 19 (29%), 28 (43%) and 13 (20%) probands fulfilled four, three, two and one diagnostic criteria.

COL4A5 gene mutations

Twenty-two out of the 60 probands (37%) had a pathogenic mutation in *COL4A5*, of which 12 (57%) are novel and 9 (43%) have been previously described (Table I).(Knebelmann, Breillat et al. 1996; Renieri, Bruttini et al. 1996; Martin, Heiskari et al. 1998; Plant, Green et al. 1999; Wang, Wang et al. 2005; Bekheirnia, Reed et al. 2010; Wang, Zhao et al. 2012; Sa, Fieremans et al. 2013) Eighty percent of the probands who presented with all four diagnostic criteria, but only 1/3 of those who presented with two or three criteria, had a pathogenic mutation in *COL4A5* (Supplementary tables S4 and S5). Missense substitutions (7/21; 33%), all of them involving glycine residues, were the most common type of mutation. The c.4342G>C (p.Gly1448Arg) missense mutation was identified in two apparently unrelated families, living in the same geographic area. As the microsatellite risk haplotype was

the same in both families (Supplementary figure S6), and additionally included an intragenic single nucleotide polymorphism (SNP) in exon 23 (c.1584A>G), the two families probably share a common ancestor. Other types of pathogenic *COL4A5* mutations identified in our cohort were nonsense mutations (n=2), acceptor splice site mutations (n=2), donor splice site mutations (n=4), small deletions (n=2) and large deletions (n=4). It is of note that exon 47 harbored two distinct point mutations; that the acceptor splice site of intron 30 harbored 2 distinct small deletions; and that the breakpoint of two deletions confined to *COL4A5* were located to intron 1. The *COL4A5* mutation was proven to be *de novo* by parental genotyping in 2 out of 13 (15%) genetic probands. Since the remainder 9 probands had family history of major AS manifestations, we assume that pathogenic *COL4A5* mutations occurred de novo in 9% of probands from this cohort.

Identification of the disease-causing mutation in the genetic probands allowed confirmation of the diagnosis of XLAS in an additional 43 patients of the 22 families, increasing the size of the phenotyping cohort to 65 patients – 30 males (46%) and 35 females (54%), with mean ages at genotyping respectively of 36±16.8 and 42±14.8 years (Supplementary tables S7, S8, S9 and S10).

Table I. Pathogenic *COL4A5* mutations identified in 22 families with clinical diagnosis of collagen type IV-related nephropathy.

Type of mutation	Exon / / Intron	Nucleotide change	Predicted effect on the protein	Family (ID#)	[Reference] or original report // Annotations and comments
Missense					
	E8	c.458G>A	p.(Gly153Asp)	53	[(Bekheirnia, Reed et al. 2010)] // (a) (b) (c)
	E13	c.715G>A	p.(Gly239Arg)	23	[(Wang, Zhao et al. 2012)] // (a) (b) (c) (§)
	E18	c.1009G>A	p.(Gly337Ser)	48	**Novel** // (a) (b) (c) (§) (^)
	E24	c.1718G>A	p.(Gly573Asp)	57	[(Martin, Heiskari et al. 1998)] // (a) (b) (c)
	E25	c.1844G>A	p.(Gly615Glu)	49	**Novel** // (a) (b) (c) (§) (^)
	E31	c.2633G>T	p.(Gly878Val)	50	[(Plant, Green et al. 1999)] // (a) (b) (c)
	E47	c.4342G>C	p.(Gly1448Arg)	7, 11	**Novel** // (a) (b) (c) (§) (^)
Nonsense					
	E33	c.2815G>T	p.(Glu939*)	26	**Novel** // (a) (§)
	E47	c.4444C>T	p.(Gln1482*)	6	LOVD[@] // (a) (§)
Splicing					
	I19	c.1165+1G>A		60	[(Renieri, Bruttini et al. 1996)] // (a)
	I20	c.1339+6C>G		3	**Novel** // (a) (d) (§)
	I46	c.4297+1G>A		31	[(Wang, Wang et al. 2005)] // (a)
	I49	c.4803+1G>A		45	[(Knebelmann, Breillat et al. 1996)] // (a)
Small deletions					
	E10	c.590delC	p.(Pro197Glnfs*6)	15	**Novel** // (a) (§)
	E30	c.2423delG	p.(Gly808Aspfs*11)	1	**Novel** // (a) (§)

	I30/E31	c.2510(-1)_2525del(15bp)[Δ]	34	**Novel**
Large deletions				
	E1-13	delE1_13 (+delE1_3,*COL4A6*)	33	[(Sa, Fieremans et al. 2013)] (**Novel**)
	E2-29	delE2_29	41	[(Sa, Fieremans et al. 2013)] (**Novel**)
	E2-51	delE2_51	19	[(Plant, Green et al. 1999; Sa, Fieremans et al. 2013)] // (^^)
	E31	c.2510_2554del(45bp)[ΔΔ]	28	**Novel**
	E43-45	delE43_45	18	**Novel**

Mutations were classified according to the Human Gene Mutation Database nomenclature [www.hgmd.cf.ac.uk/]. Nonsense point mutations, frameshifting mutations and exonic deletions were considered pathogenic.

[@] Mutation reported at the LOVD – Leiden Open Variation Database [accessed at https://grenada.lumc.nl/LOVD2/COL4A/variants.php?select_db=COL4A5&action=search_all&search_Variant%2FDNA=c.4444C>T].

Bioinformatic predictions of pathogenicity for missense non-synonymous mutations, splice site intronic mutations and short (<12 nucleotides) microdeletions:

(a) MutationTaster [www.mutationtaster.org/];
(b) PolyPhen-2 [genetics.bwh.harvard.edu/pph2/];
(c) SIFT [sift.jcvi.org];
(d) Human Splicing Finder [http://www.umd.be/HSF/].

(§) Sequence variation not identified in 234 control X-chromosomes.

(^) Pathogenic mutations reported at the same codons:

(*i*) p.(Gly337Asp) / c.1010G>A, reported at the Leiden Open Variation Database [LOVD; accessed at https://grenada.lumc.nl/LOVD2/COL4A/variants.php?select_db=COL4A5&action=search_all&search_Variant%2FDNA=c.1010G>A];
(*ii*) p.(Gly615Arg) / c.1843G>C, reported by Wang et al, 2012(Wang, Zhao et al. 2012);
(*iii*) p.(Gly1448Ser) / c.4342G>A, reported by Wang et al. 2005(Wang, Wang et al. 2005).

(^^) Plant *et al.* 1999(Plant, Green et al. 1999) also reported large deletions encompassing exons 2_51 but their exact breakpoints were not determined.

[Δ] GGTTTACATCATGGAATA.

[ΔΔ] GTTTACATGGAATACCAGGAGAGAAGGGGGATCCAGGACCTCCTG (p.delGLHGIPGEKGDPGPP): the first nucleotide involved in this deletion is the first nucleotide of exon 31.

Inter-gender comparisons of clinical outcomes

Microscopic hematuria was detected in all the 65 patients of the XLAS phenotyping cohort (Table II). History of macroscopic hematuria was significantly more common in males than females (62% vs 17%; p=0.002). Proteinuria was more frequently reported in males than females, but the difference did nor reach statistical significance (96.2% vs 79%; p=0.067). Although the proportion of patients with history of hypertension also did not differ significantly between males and females (73% vs 53%; p=0.112), the median age at onset of hypertension was significantly lower in the male patients (20 vs 34 years; p=0.002). Eighty-six percent of the male patients and 62% of the females developed CRF (p=0.029). CRF was diagnosed in males at a significantly lower age than in females (21 vs 32 years; p=0.038). At diagnosis of CRF, eGFR was significantly lower in males (40 vs 61 ml/min/1.73m^2; p=0.003). While 77% of male patients with CRF had already started RRT, only 17% of females with CRF were on RRT (p<0.001). Six males and three females in the XLAS phenotyping cohort had EM photomicrographs of kidney biopsies available for review. All of them showed the typical ultrastructural features of AS, i.e. alternate thinning and thickening of the GBM with splitting and/or reticulation of the lamina densa, enclosing electron-lucent areas that contain round electron-dense bodies with diameters of up to 90 nm. Subjective hearing loss was more often reported (87% vs 46%; p=0.001) in males than in females. Likewise, anterior lenticonus and dot-and-fleck retinopathy were more frequently diagnosed in the males (p=0.013 and p=0.129, respectively). ATS-DL was diagnosed in a single family, in association with a large deletion involving *COL4A5*, but not the *COL4A6* gene.(Sa, Fieremans et al. 2013)

Table II. Phenotypic comparisons between hemizygous males and heterozygous females for pathogenic *COL4A5* mutations (n=65).

	Hemizygotes (n=30)		Heterozygotes (n=35)		p-value
		N		N	
PHENOTYPE					
Age at enrolment (median (interquartile range))	33.5 (18.5)	30	42 (14.8)	35	0.088
History of renal abnormalities					
History of macroscopic hematuria (%) Age at diagnosis (median (interquartile range))	61.9 5 (4)	13/21 11	17.4 3 (11.5*)	4/23 3	0.002 0.501
History of microscopic hematuria (%) Age at diagnosis (median (interquartile range))	100.0 12.5 (20.3)	25/25 22	100.0 21 (26.5)	35/35 28	a 0.219
History of proteinuria (%) Age at diagnosis (median (interquartile range))	96.2 18 (22.5)	25/26 21	78.8 24 (19)	26/33 23	0.067 0.508
History of hypertension (%) Age at diagnosis (median (interquartile range))	73.1 20 (11.5)	19/26 12	52.9 34 (22.3)	18/34 16	0.112 0.002
History of CKD stage 2 or higher (%) Age at diagnosis (median (interquartile range)) eGFR at diagnosis (median (interquartile range))	86.2 21 (10) 40 (35)	25/29 16 13	61.8 32 (17.5) 61 (26.3)	21/34 20 20	0.029 0.038 0.003
History of renal replacement therapy (%) Age at onset (median (interquartile range)) eGFR at onset (median (interquartile range))	76.7 23 (19.5) 9 (6)	23/30 21 12	17.1 40 (14) 6 (1)	6/35 6 5	<0.001 0.272 0.019
GBM ultrastructural abnormalities					
Age at kidney biopsy (median (interquartile range))	20 (14.5)	6	24 (11*)	3	0.933
Thinning (%)	100.0	6/6	100.0	3/3	a
Thickening (%)	100.0	6/6	100.0	3/3	a
Lamellation (%)	100.0	6/6	100.0	3/3	a
Electrondense bodies (%)	100.0	6/6	100.0	3/3	a
Hearing loss					
Self-noticed or subjective (%) Age at self-noticed or subjective (median (interquartile range))	86.7 15 (30)	26/30 15	45.5 39 (19)	15/33 9	0.001 0.142
Audiogram (%) Age at diagnosis of hearing loss by audiogram (median (interquartile range))	94.1 27.5 (18)	16/17 16	91.3 41 (10)	21/23 13	0.053 0.030
Ocular abnormalities					
Anterior lenticonus (%) Age at diagnosis (median (interquartile range))	33.3 19 (12)	5/15 5	0.0 -	0/18 -	0.013 -
Maculopathy (%) Age at diagnosis (median (interquartile range))	50.0 25 (19)	10/20 8	26.3 44 (22)	5/19 5	0.129 0.180
Cataracts (%) Age at diagnosis (median (interquartile range))	46.2 30 (35.8)	6/13 6	15.0 56 (56*)	3/20 3	0.107 0.263
Other abnormalities					
Leiomyomatosis **(%)** Age at diagnosis (median (interquartile range))	3.3 24 (0*)	1/30 1	5.7 18 (0*)	2/35 1	1.000 0.317
GENOTYPE					
Truncating mutation (%)	47	14/30	37.1	13/35	0.437

SD: standard deviation. CKD: chronic kidney disease. eGFR: estimated glomerular filtration rate. GBM: glomerular basement membrane. Age is expressed in years. eGFR is expressed in ml/min/1.73m^2. * The range is indicated instead of the interquartile range. [a] Statistics not computed because the proportions in the comparison groups are the same.

Genotype-phenotype correlations

Among males, no statistically significant differences were found between patients with truncating and non-truncating *COL4A5* mutations in the prevalence of the major renal manifestations of AS, or of their natural histories as captured by age or severity of CRF at diagnosis (Supplementary table S11). Although the prevalence of SNHL was not statistically different in the two groups, subjective hearing loss was reported by patients with truncating mutations at a much younger age than by those with non-truncating mutations (7 vs 36 years; $p<0.001$). While the difference in the prevalences of anterior lenticonus in patients with truncating and non-truncating mutations who underwent formal ophthalmological examination did not reach statistical significance (50% vs 14%; $p=0.282$), dot-and-fleck retinopathy was identified in a significantly higher proportion of patients with truncating mutations (75% vs 13%; $p=0.020$).

Among females, CRF was diagnosed in 92.3% of those with truncating *COL4A5* mutations but only in 43% of those with non-truncating mutations ($p=0.005$) (Supplementary table S12). However, the age at diagnosis of CRF did not differ significantly (32 vs 31 years; $p=0.684$) and the proportion of patients who needed RRT was similar in the two groups (15% vs 18%; $p=1.000$). Patients with truncating mutations reported subjective hearing loss more frequently than those with non-truncating mutations (69% vs 30%; $p=0.027$) and SNHL was also confirmed by tonal audiogram at a significantly younger age in the former group of patients (38 vs 47 years; $p=0.003$). Irrespective of the type of mutation, ocular involvement was diagnosed in only a small minority of females with AS.

DISCUSSION

Herein, we report the clinical characterization, the molecular pathology and the genotype-phenotype correlations observed in 22 Portuguese families with genetically confirmed XLAS, identified by mutational screening of a nationwide cohort of 65 non-related probands in whom the diagnosis of AS was considered clinically established or probable, according to standard clinical and pathologic criteria.(Flinter, Cameron et al. 1988)

A pathogenic *COL4A5* mutation was identified in only 34% (22/65) of families. This prevalence is within the 30-67% range reported in several large studies that have used single-strand conformation polymorphism (SSCP) analysis to screen for *COL4A5* mutations,(Hertz 2009) but is significantly lower than the 82% rate reported with PCR amplification and direct DNA sequencing of the promoter and all the exons of *COL4A5*.(Martin, Heiskari et al. 1998) Such a result cannot be attributed to limitations of the laboratory methods that we have used to scan for mutations, as the sensitivity of MLPA combined with direct sequencing to identify *COL4A5* mutations is higher than 95%(Hanson, Storey et al. 2011) while that of SSCP is estimated to be around 80%.(Grompe 1993) Although we did not scan for promoter and deep intronic mutations, missing *COL4A5* mutations at those locations is an unlikely explanation(King, Flinter et al. 2006) for the lower than expected prevalence of XLAS observed in the Portuguese families.

The wide variety of different types of pathogenic *COL4A5* mutations identified in this study is further evidence of the effectiveness of our laboratory approach, and their relative frequencies did not significantly differ from those reported to the HGMD®. Surprisingly, a large deletion confined to *COL4A5* was identified in a family with ATS-DL, challenging an established concept of genotype-phenotype correlation.(Sa, Fieremans et al. 2013) Pathogenic *COL4A5* mutations were novel in 57% of families (12/21) and occurred *de novo* in 15% (2/13) of the probands, which is within the expected 10-15% range.(Kashtan 2001) With the exception of a glycine substitution (p.Gly1448Arg) shared by two apparently unrelated families from neighboring villages, all other pathogenic *COL4A5* mutations identified in our cohort were exclusive of single families. These findings show that mutation panels designed for the genetic diagnosis of AS should be based on population-specific data.[4]

The detection rates of pathogenic *COL4A5* mutations in probands with complete clinical assessment (n=22) that fulfilled one, two, three, or four diagnostic criteria were, respectively, 0.0%, 25.0%, 14.3% and 80.0%. The corresponding figures reported for British patients (n=101) were 9.5%, 48.5%, 82.1%, and 81.3%,(Hanson, Storey et al. 2011) which differ significantly from our own results only for patients fulfilling three diagnostic criteria. Since the extrarenal manifestations of AS seem to be relatively rare in patients with autosomal forms of the

disease,(Tryggvason and Patrakka 2009) a possible explanation for that difference might be the higher prevalence of ADAS or ARAS in Portuguese families. Indeed, preliminary results of second-tier mutational screening of *COL4A3* and *COL4A4* [manuscript submitted to the JMG simultaneously] suggest that the proportion of XLAS in Portuguese families with AS is significantly lower than the 80-85% generally referred in the literature.(Kashtan 1999; Tryggvason and Patrakka 2009; Kruegel, Rubel et al. 2013) These data are much in agreement with those of a recent Italian study(Fallerini, Dosa et al. 2013) reporting the diagnosis of XLAS in only 65% of 48 families with genetically confirmed AS, by unbiased simultaneous next generation sequencing (NGS) of *COL4A5, COL4A3* and *COL4A4*. In keeping with the above argument, more than half of the Italian patients that presented with only one or two diagnostic criteria of AS had a pathogenic mutation identified by NGS.(Fallerini, Dosa et al. 2013)

The clinical manifestations of renal disease observed in our phenotyping cohort did not substantially differ from the classical descriptions of the XLAS.(Kashtan 1999; Tryggvason and Patrakka 2009) Microscopic hematuria was the most consistent diagnostic clue, being present in all cases, irrespective of gender. Although history of macroscopic hematuria was much more frequent in males, the first episode typically occurred before school age in both boys and girls. Proteinuria was also highly prevalent in males and females but its diagnosis usually followed the detection of hematuria by several years. The risks of developing progressive CRF and of needing RRT were considerably higher in males; nevertheless, CRF was diagnosed in 62% of the heterozygous females, at young adult age in many cases. Only two-thirds of the patients in our genotyping cohort had undergone audiological assessment and only approximately 58% had been referred for specialized ophthalmological examination. These data show that formal screening for the extrarenal signs of AS is frequently overlooked in clinical practice.

Because of the genetic heterogeneity and variable clinical expression of AS, and the incomplete penetrance of XLAS in females, clinicians should take a detailed three-generation family history in order to try to recognize a pattern of inheritance, before initiating genetic testing for AS.(Hanson, Storey et al. 2011) Particularly in females, the diagnosis of microscopic hematuria requires repeated urinalysis. Not

screening for SNHL by audiogram, and for the typical ocular lesions by slit-lamp ophthalmological examination, are common errors in clinical practice that should be avoided. Indeed, failure to comprehensively assess patients for the clinical diagnostic criteria of AS may be one of the reasons for missing the diagnosis in some cases, for uninformative family histories and erroneous estimates of genetic risks, and for false assumptions of *de novo* mutations.

In contrast with results of previous studies,(Jais, Knebelmann et al. 2000; Gross, Netzer et al. 2002; Bekheirnia, Reed et al. 2010) we did not identify any significant differences on the prevalence or severity of expression of renal disease between males carrying truncating or non-truncating *COL4A5* mutations, although they had a more severe extrarenal phenotypes. On the other hand, the higher prevalence of CRF and of hearing loss in females with truncating *COL4A5* mutations observed in our cohort was not apparent in the largest study of females with XLAS reported so far.(Jais, Knebelmann et al. 2003) However, the relative small size of our study cohort should caution about the generalization of the present findings.

Genetic testing in AS is useful for confirming a clinical diagnosis, informing genetic counseling, and facilitating both prenatal and pre-implantation diagnosis.(Bekheirnia, Reed et al. 2010) The identification of a pathogenic *COL4A5* mutation is the most specific diagnostic criterion of XLAS, avoiding more invasive and complex diagnostic procedures as kidney or skin biopsy.(Kashtan 1993-2013) In addition, the type of mutation is of help in the estimation of the clinical prognosis.(Jais, Knebelmann et al. 2000; Gross, Netzer et al. 2002; Bekheirnia, Reed et al. 2010) However, molecular genetic analysis by Sanger sequencing of all 53 *COL4A5* exons and their flanking intronic sequences is an expensive, time-consuming process and several weeks are usually needed to complete a diagnosis.(Artuso, Fallerini et al. 2012) The recently available, high-throughput NGS methods for DNA sequencing are cheaper and have much lower turn-around times than the conventional Sanger method. High-throughput DNA sequencing methods are most useful for the molecular diagnosis of genetically heterogeneous disorders, and NGS may soon become the first-tier genetic test in cases where AS is clinically suspected, particularly in populations where the prevalence of XLAS is not much higher than that of the autosomal forms.(Fallerini, Dosa et al. 2013) In those families

where the causative *COL4A5* mutation is already known, genetic testing of at-risk persons is much less expensive, allows the accurate assessment of individual genetic risks and provides a tool for pre-symptomatic diagnosis, which is particularly valuable in females, and for the selection of potential family donors for kidney transplantation.(Gross, Weber et al. 2009; Kashtan 2009) Moreover, early confirmation of the diagnosis of XLAS offers the opportunity to optimize treatment outcomes, since pharmacologic blockade of the renin–angiotensin–aldosterone system with angiotensin converting enzyme inhibitors or angiotensin receptor blockers proved effective in delaying the progression of CRF in both males and females.(Gross, Licht et al. 2012; Temme, Peters et al. 2012)

ACKNOWLEDGEMENTS

The authors thank the patients and their relatives for the invaluable collaboration in this study; NephroCare – Portugal, for having allowed the nationwide ascertainment of patients with Alport syndrome in their dialysis centers; and Dr.s Adelino Carvalho (NephroCare, Santarém), Ana Bernardo (NephroCare, Covilhã), Carlos Miguel Botelho (NephroCare, Viseu), Joana Felgueiras (NephroCare, Montijo), Liliana Pinho (Dialysis clinic, Paredes) and Luís Filipe Carvalho (NephroCare, Restelo), for their contributions to the clinical data collection of dialysis patients enrolled in this study.

The authors also thank Dr. Francis Earl Cook, from the Department of Epidemiology, Harvard School of Public Health, Boston, MA, for his advice and supervision of the epidemiological and statistical analyses.

This work is part of the PhD project of Maria João Nabais Sá (supervisor: João Paulo Oliveira), to be presented at the ICBAS – "Instituto de Ciências Biomédicas Abel Salazar" (Abel Salazar Biomedical Sciences Institute), University of Porto, Portugal.

These studies were partially supported by FEDER funds through the COMPETE – "Programa Operacional Factores de Competitividade", and by Portuguese national funds through the FCT – "Fundação para a Ciência e a Tecnologia" (Foundation for Science and Technology) within the framework of the Cooperation Agreement between Portugal and Harvard Medical School (FCOMP-01-0124-FEDER-HMSP-ICJ/SAU-ICT/0003/2009); by a research grant from the Portuguese Society of Nephrology; by the Unit of Research & Development in Nephrology (FCT-725) at the Department of Nephrology of the Faculty of Medicine, University of Porto, Portugal; and by the São João Hospital Centre.

Parts of these data were presented as posters at the European Human Genetics Conference 2012, Nürnberg, Germany, June 23-26, 2012 (P12.007) and at the 62nd Annual Meeting of the American Society of Human Genetics, November 6-10, 2012 San Francisco, California (3122F).

Supplementary table S1. Outcome variables, definitions, measurement of outcomes, diagnostic methods, degree of evidence and time of collection of clinical variables.

Outcomes	Definition of outcomes	Measurement of outcomes	Diagnostic evaluation	Degree of evidence	Time of diagnosis
Birth District	District of Portugal where the patient was born	Portuguese districts	None	I, CF, Q	–
Residence District	District of Portugal where the patient lived at the date of enrollment	Portuguese districts	None	I, CF, Q	–
Affected relatives	Family history of hematuria, CRF or hearing loss	Exams required to evaluate the renal, audiologic and ocular involvement	Pedigree	I, CF, Q	At enrollment Retrospectively
Renal anomalies					
History of macroscopic hematuria	Presence of red blood cell (RBC) in urine visible to the eye, as “cola-colored” urine.	Occurrence of outcome (Yes/No) Age at first event (in years)	Urinalysis	I, CF, Q	Retrospectively
History of microscopic hematuria	≥3 RBC visible per high power field under light microscopy examination of the centrifuged urine sediment or an equivalent value per microliter when assessed by flow cytometry	Occurrence of outcome (Yes/No) Age at first event (in years)	Urinalysis: - RBC count	Urinalysis report I, CF, Q	Retrospectively At enrollment During the study
History of proteinuria	Urinary protein excretion ≥300 mg/day or a urinary protein to creatinine ratio ≥300 mg/g, in morning spot samples. The highest recorded level of 24-hour proteinuria was graded as follows: ≤500 mg/day; 500–1000 mg/day; 1000–3500 mg/day; ≥3500 mg/day.	Occurrence of outcome (Yes/No) Age at first event (in years) Degree of proteinuria (mg/day or mg/g) at diagnosis/most recent Highest recorded level of proteinuria	Urinalysis: - Protein count	Urinalysis report I, CF, Q	Retrospectively At enrollment During the study

Outcomes	Definition of outcomes	Measurement of outcomes	Diagnostic evaluation	Degree of evidence	Time of diagnosis
History of hypertension	Systolic blood pressure ≥140 mmHg and/or diastolic blood pressure ≥90 mmHg. Patients undergoing chronic anti-hypertensive medication were also considered to have hypertension, irrespective of their blood pressure levels.	Occurrence of outcome (Yes/No) Age at first event (in years) Therapy with anti-hypertensive medication (Yes/No) Age at onset of anti-hypertensive medication (in years)	Blood pressure measurement	I, CF, Q	Retrospectively At enrollment During the study
History of CRF (≥ stage 2 CKD)	Plasma creatinine (pCr) level ≥1.2 mg/dl in males and ≥0.9 mg/dl in females.*	Plasma creatinine level (mg/dl) CKD stage (≥2, according to the KDOQI clinical guidelines)	Blood analysis: - Plasma creatinine level	Blood analysis report I, CF, Q	Retrospectively At enrollment During the study
History of advanced CRF (≥ stage 3 CKD)	Plasma creatinine level >1.5 mg/dl in males and >1.2 mg/dl in females.	Plasma creatinine level (mg/dl) CKD stage (3 to 5, according to the KDOQI clinical guidelines)	Blood analysis: - Plasma creatinine level	Blood analysis report I, CF, Q	Retrospectively At enrollment During the study
Estimated glomerular filtration rate (eGFR)	eGFR computed from pCr values, namely at the diagnosis of CRF and advanced CRF (eGFR at diagnosis of stage 2 and stage 3 CKD).**	eGFR (ml/min/1.73 m^2) CKD stage (≥2, according to the KDOQI clinical guidelines)	Blood analysis: - Plasma creatinine level	Blood analysis report I, CF, Q	Retrospectively At enrollment During the study
History of renal replacement therapy (RRT)	First procedure of renal replacement function (hemodialysis, peritoneal dialysis or renal transplantation), performed when the GFR reached <15 ml/min	Need of RRT (Yes/No) Age at onset of first RRT (in years) Plasma creatinine level at onset of first RRT (mg/dl) eGFR at onset of first RRT (ml/min/1.73 m^2)	Blood analysis: - Plasma creatinine level	Blood analysis report I, CF, Q	Retrospectively At enrollment During the study
Time between onset of first RRT and first kidney transplant	Time from the date of onset of first RRT (dialysis) and first kidney transplant	Time (in months)	None	I, CF, Q	Retrospectively At enrollment During the study
Last available pCr (eGFR)	eGFR calculated from the last measurement of plasma creatinine level of the participant, to evaluate his/her CKD stage**	eGFR (ml/min/1.73 m^2) CKD stage (≥2, according to the KDOQI clinical guidelines)	Blood analysis: - Plasma creatinine level	Blood analysis report I, CF, Q	Retrospectively At enrollment During the study

Outcomes	Definition of outcomes	Measurement of outcomes	Diagnostic evaluation	Degree of evidence	Time of diagnosis
Typical ultrastuctural GBM abnormalities	Characteristic ultrastructural GBM abnormalities of Alport syndrome, namely thinning, thickening, lamellation and electron-dense particles in GBM lacunae	Age at kidney biopsy (in years) GBM thinning (Yes/No) GBM thickening (Yes/No) GBM lamellation (Yes/No) Electron-dense particules in GBM lacunae (Yes/No)	Kidney biopsy with electron microscopy examination	Pathology database: - Electron microscopy photographies Kidney biopsy report (CF), Q	Retrospectively
Typical hearing loss (HL)					
History of self-noticed HL	Self-perception of HL and approximated age at onset of self-noticed HL	Self-noticed HL (Yes/No) Age at onset of self-noticed HL (in years)	None	I, CF, Q	Retrospectively
Diagnosis of HL	Diagnosis of HL and its characterization by audiogram, including uni- or bilateral HL, type of HL and degree of HL. The degree of HL was classified using standardized hearing loss categories, according to the WHO classification (Danermark, Cieza *et al*. 2010), as follows: dB HL < 26 (normal); 26 ≤ dB HL < 40 (Mild HL); 41 ≤ dB HL < 60 (Moderate HL); 61 ≤ dB HL < 80 (Severe HL); 80 < dB HL (Profound HL).	Diagnosis of HL (Yes/No) Type of HL Degree of HL: - Pure tone average (PTA): average of the three thresholds (in decibels, dB) at 500 Hz, 1000 Hz and 2000 Hz - Average of high frequencies: average of the two thresholds (in decibels, dB) at 4000 Hz and 8000 Hz Age at the first and the most recent audiometric evaluations (in years)	Pure tone audiogram High frequency audiogram	Audiogram	Retrospectively At enrollment During the study
Typical ocular abnormalities					
History of anterior lenticonus, retinopathy and	History of and age at diagnosis of anterior lenticonus, dot-and-fleck maculopathy, retinal atrophy and cataracts	Anterior lenticonus (Yes/No) Dot-and-fleck maculopathy (Yes/No)	Biomicroscopy Funduscopy OCT	Biomicroscopy photography Fundus	Retrospectively At enrollment During the study

Outcomes	Definition of outcomes	Measurement of outcomes	Diagnostic evaluation	Degree of evidence	Time of diagnosis
cataracts		Retinal atrophy (Yes/No) Cataracts (Yes/No) Age at the first and the most recent ophthalmological assessment (in years)		photography I, CF, Q	
Other clinical outcomes					
Abnormal type IV collagen expression in EBM	Abnormal epidermal basement membrane (EBM) reactivity to a monoclonal antibody directed against the α5(IV) chain	Complete absence, discontinuous or presence of staining of the EBM with a monoclonal antibody directed against the collagen α5(IV) chain	Skin biopsy with immunohistochemical analysis of the type IV collagen expression in the EBM	EBM photography	Retrospectively During the study
History of diffuse leiomyomatosis	Gastrointestinal, tracheobronchial and female genital manifestations	Histologic diagnosis	Endoscopic ultrasonography with esophageal biopsy	Pathology report I, CF, Q	Retrospectively At enrolment During the study
History of intellectual impairment	Level of education and employment	Number of year in school Degree completed Job	None	I, CF, Q	Retrospectively At enrolment During the study
Typical dysmorphic face	Craneofacial dysmorphy described in patients with AMME complex (Meloni et al., 2002)	Mild midface hypoplasia Anteverted nares Flat nasal bridge Downward slanting palpebral fissures Peripalpebral fullness Facial hypotonia	Physical examination	Facial photography I, CF, Q	Retrospectively At enrollment

I: Patient interview.
CF: Patient clinical file: consultation registry, pre-dialysis medical report, referral information.
Q: Questionnaire filled by assistant physician.
OCT: Optical Coherence Tomography.
*These pCr cut-off values were selected to be roughly equivalent to the upper limits of eGFR for the diagnosis respectively of stage 2 and stage 3 chronic kidney disease (CKD), in patients with kidney damage, over the age of 18 years, according to the Kidney Disease Outcomes Quality Initiative (KDOQI) clinical practice guidelines for CKD evaluation, classification and stratification of the National Kidney Foundation (NKF; New York, NY, USA) (http://www.kidney.org/professionals/KDOQI/guidelines_ckd/p4_class_g1.htm).
**The eGFR was computed from pCr values using the calculators available online at the NKF website (http://www.kidney.org/professionals/kdoqi/gfr_calculator.cfm), as appropriate. [Online calculator used at the NKF website: CKD-EPI Creatinine equation 2009]

Supplementary table S2. Primer sequences, annealing temperatures and product size for PCR amplification of the *COL4A5* gene.[†]

Exons	Foward Primers (5' to 3')	Reverse Primers (5' to 3')	Annealing temperature	PCR Product size
1	AAGCCTCACTGTCCCTCTC	AAAGGAAGATAAAGGGACCC	60ºC	296 bp
2	GATTGTTGATTTCAGTTGAGCTGT	CTTAAGTACTGAGATAGAAGCTC	51ºC	248 bp
3	TCTCAACCATGCCTGTGCTTG	TGATGTGACACCTAGTCCCAC	60ºC	229 bp
4	TCACAGATGTTTACAGTAGTTTAAA	GGTCTTTTCCAATTGTCTCAT	57ºC	237 bp
5(F)/6(R)	GAACAATGAGTAGTAATAAATAGTG	TTGTCAAGTATACTACCCAAGAT	52ºC	373 bp
7(F)/8(R)	GGAAAGTGAAGGCTAATGAAC	ATTATCCTATTGAAGTTGCCAGC	52ºC	644 bp
9	GCCTAATCTTTTAGTACATCTC	GTACTAAAGGTTGAGGGAT	52ºC	275 bp
10	GGCGACACAAGTGAGACTTT	GAATGTTGAGAATGCATTATGTTTTC	52ºC	248 bp
11(F)/13(R)	AATACTATTTTGATGGGCTTT	TATACATTGTGATGTGATTAC	52ºC	628 bp
14(F)/16(R)	CTCCAGCTCTAACCATGTTG	TTTTGTCATACTGCTTCTCT	55ºC	726 bp
17	GAAGACAATCTTTGGAG	TAACTCAAGCCTGGGAGAA	52ºC	201 bp
18	TGGAAAGTTTCTCTTATATTC	TGTGTTCTATCAAGAAAGAG	48ºC	224 bp
19	CAGGAGAACAAGGCTTTTCTTCTTTGCATT	ACATGGATTAGTAAGGATGC	52ºC	248 bp
20(F)/21(R)	GAAGATCTTATCATTATCTAATG	GTCTTGGGGTATGACCAT	51ºC	793 bp
22	GTGGAAATGCTGTCCCTTAG	CATTCTAAAAGTATAAGCTCAAC	52ºC	266 bp
23(F)/24(R)	TCAAAGCTTACGTTATTGTGT	GGTGTTTATCTCAGCATCAG	55ºC	726 bp
25	GAAGAACTATTTATGGCTATATCC	TTGGCTACTCATGGCTTCC	57ºC	350 bp

Exons	Foward Primers (5’ to 3’)	Reverse Primers (5’ to 3’)	Annealing temperature	PCR Product size
26(F)/27(R)	AGACCTTTAGTTGAGTAAATAC	CATGATGTTATTCTCTCATAAA	52ºC	757 bp
28	ACATCTTACTGTTGTCACTAAGC	GTCTTTACTATAAATATACAGATAAC	52ºC	334 bp
29	CGGCATTAAATTCTCTGT	GTTTCCAGTTAATTGGGC	57ºC	333 bp
30	TTCTTGCTGAATGAATGC	TCACTTTATTGATGAGCTAAC	55ºC	300 bp
31	TTAGGTCTGTTATCTACAGGG	GAATTATCTACCAGAGTCGTATTAG	48ºC	347 bp
32	AGTTTTCTGGTTGACATCTTAAAA	TATTCTGTACTGACATAAAGC	55ºC	249 bp
33	ATATGCATTAATCTTTGATGGA	ATAAGTCACTTTTCATGC	48ºC	293 bp
34	CTTATAGTTTAACACTTGAGTAG	GAATTCAGTGTCAGCTAAGCA	52ºC	243 bp
35(F)/36(R)	GACTATCCATTCCCATGAAACCAG	AAAACTTAGGAATGAGCATAG	50ºC	800 bp
37	ATTTACATCAAGTACTTACTGGAG	AGTCTGCCAAATAAGAAGCTGC	53ºC	337 bp
38	GTTCTTCACTGTTTCTATGCT	AATGTTCACAGCTGAACATGA	48ºC	279 bp
39	GAAGGGAGCATATGGAAG	CAAGTTAAATTCAACACAG	50ºC	233 bp
40	AGTTTGTATTATCCACTTGAGT	GTAAAAGGTGGAGATGGAAAA	48ºC	191 bp
41	ATTGCCCTAATGTATGTGAATAGC	CAAGCTATTTACTTGTAAGAAGTTAT	50ºC	441 bp
41A	CTTTTTGTTAATGATGACAT	ACAGAAACACTGGGTTCTACA	55ºC	232 bp
41B	CTTCTGTATGGTTCTGTTTGC	TTGCATTTCTCTTATCACACAC	55ºC	269 bp
42	GAAGATGACTGATATTTTAAAAG	GAATTCACATTTTGAATACA	50ºC	286 bp
43(F)/44(R)	AGTTTGACTCTAGAAATAGTGC	GCATCTAAGTATCAGGTATAAC	50ºC	433 bp
45	GTCCTGAACTTAGGTCACTTT	GCACATGTATCCCAGAAAT	55ºC	228 bp
46	CTTATGTCTCCTAGATCTGTC	ATCTCACAGCTCTAAGTGGAG	55ºC	263 bp

Exons	Foward Primers (5' to 3')	Reverse Primers (5' to 3')	Annealing temperature	PCR Product size
47	ATGAGGTCATAATGTTTTGTC	CACTTTGGCCAAGGCTACTC	55ºC	386 bp
48	TTGAAAGGCTGTTTGCTATTG	ACCTAGGAGGAATATCATCAG	55ºC	364 bp
49	ATCGGCTTCCATACTAAGAAG	GCAAATGACAGGGATTCCTC	55ºC	304 bp
50	GATTTGAATTTGGCCAAGCTC	CATTTGTTGAGGATAAACCAT	55ºC	366 bp
51	GTCACCAAGAGAGCTACTTAACAC	CATTGACGGCAGCAGTAGTAAAG	55ºC	303 bp

† Cycling conditions (adapted from (Martin, Heiskari et al. 1998)) were: initial denaturation for 5 minutes at 94°C, followed by 35 cycles of denaturation at 94°C for 60 seconds, annealing for 60 seconds, and extension at 72°C for 90 seconds, ending with a final extension at 72°C for 10 minutes.

Supplementary figure S3

Supplementary figure S3. Map of Portugal showing the boundaries of each of the 14 continental districts (white background) and the two autonomous Atlantic islands regions. The names and geographic locations (black squares) of the respective capital cities are shown only for those where probands enrolled in this study were resident (numbers in parenthesis). Fifty-one (78.5%) of the 65 probands were enrolled through 11 nephrology and two genetics services of public hospitals, and 14 (21.5%) were enrolled through 12 outpatient dialysis clinics. Of the 16 nephrology departments of public hospitals that did not enroll patients, four did not reply to the invitation to participate, two declined to participate and ten responded they had no patients with the clinical diagnosis of Alport syndrome. Nationwide case ascertainment through public hospitals and private dialysis clinics minimized the negative impact that the nonparticipation in this study of some public hospitals might have had upon the catchment of affected families in some regions of the country. Indeed, several families could have been redundantly identified by different patients, at different stages of kidney disease, through different sites. The average proband catchment, per 100,000 inhabitants, was 0.62 in the continental districts and 1.12 in the Madeira Islands. The geographic distribution of the 65 probands by area of residence broadly paralleled the Portuguese population density, when comparing the coastal, densely populated, and the inland, scarcely populated, regions. Possible explanations for the lower than expected catchment in a few of the highly populated central-western districts, including in the capital Lisbon district, are the nonparticipation in this study of a private dialysis clinics network that is the major provider of dialysis care in some of those districts, and the exclusion of four families with X-linked Alport syndrome that had been previously enrolled in the European Community Alport Syndrome Concerted Action. In contrast, the disproportionately higher number of families identified in the northwestern districts of Braga and Porto most probably reflects the long-term activity of an outpatient clinic dedicated to hereditary kidney disorders operating in a major academic hospital in Porto.

Supplementary table S4. *COL4A5* mutation detection rate per number of diagnostic criteria met, among probands tested for *COL4A5*, with complete and incomplete clinical evaluation (n=60).

		Diagnostic criteria				***COL4A5* molecular study**				
		FH	AC	OC	UC	Probands with a pathogenic mutation identified	Probands without a pathogenic mutation identified	Total number of probands	Mutation detection rate	Mutation detection rate per number of diagnostic criteria
1 criterion	FH	9				4	5	9	44%	33%
	AC		1			0	1	1	0%	(4/12)
	OC			0				0		
	UC				2	0	2	2	0%	
2 criteria	FH		14			8	6	14	57%	38%
				1		0	1	1	0%	(10/26)
					7	1	6	7	14%	
	AC			1		0	1	1	0%	
					3	1	2	3	33%	
	OC				0			0		
3 criteria	FH AC			9		2	7	9	22%	24%
					5	1	4	5	20%	(4/17)
	FH OC				0			0		
	AC OC				3	1	2	3	33%	
4 criteria	FH AC OC				5	4	1	5	80%	80% (4/5)

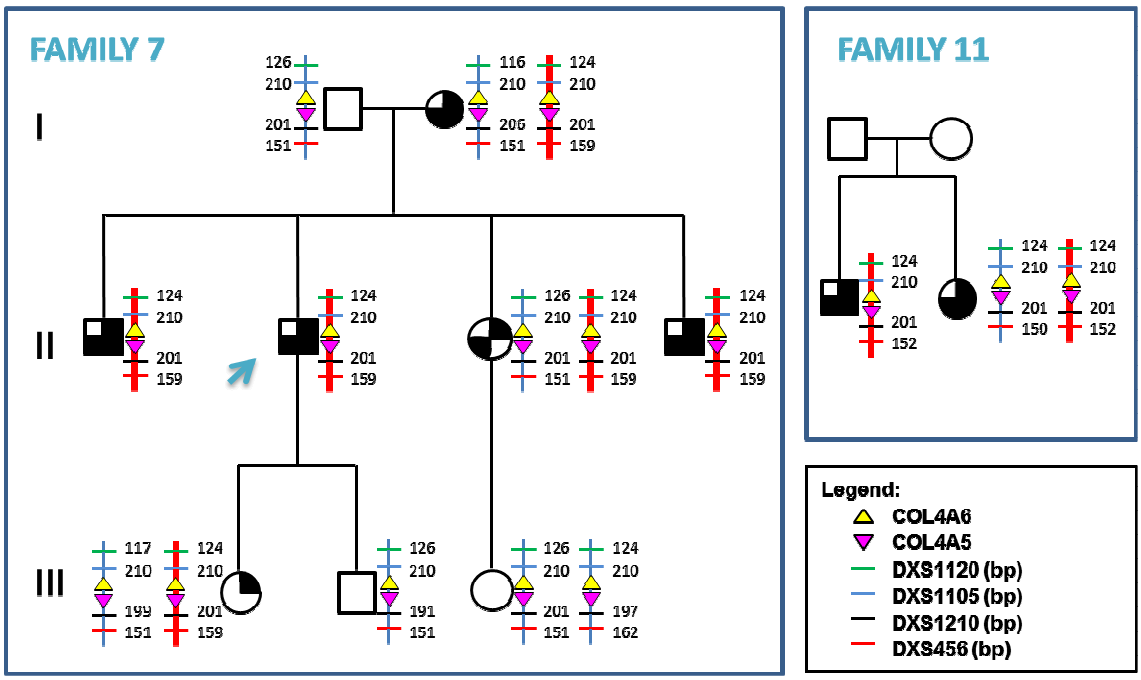
FH: family history.
AC: audiologic criteria; hearing loss, with and without confirmation by audiogram.
OC: ophthalmologic criteria; ocular signs, including anterior lenticonus, cataracts and retinopathy.
UC: ultrastructural criteria; characteristic glomerular basement membrane changes on electron microscopy examination of kidney biopsy, including thinning, thickening, lamellation and presence of electron-dense bodies.

Supplementary table S5. *COL4A5* mutation detection rate per number of diagnostic criteria met, among probands tested for *COL4A5*, with complete clinical evaluation (n=22).

		Diagnostic criteria				*COL4A5* molecular study				
		FH	AC	OC	UC	Probands with a pathogenic mutation identified	Probands without a pathogenic mutation identified	Total number of probands	Mutation detection rate	Mutation detection rate per number of diagnostic criteria
1 criterion	FH	0						0		0%
	AC		0					0		(0/2)
	OC			0				0		
	UC				2	0	2	2	0%	
2 criteria	FH		0					0		25%
				0				0		(2/8)
					5	1	4	5	20%	
	AC			0				0		
					3	1	2	3	33%	
	OC				0			0		
3 criteria	FH AC			0				0		14%
					4	0	4	4	0%	(1/7)
	FH OC				0			0		
	AC OC				3	1	2	3	33%	
4 criteria	FH AC OC				5	4	1	5	80%	80% (4/5)

FH: family history.
AC: audiologic criteria; hearing loss, with and without confirmation by audiogram.
OC: ophthalmologic criteria; ocular signs, including anterior lenticonus, cataracts and retinopathy.
UC: ultrastructural criteria; characteristic glomerular basement membrane changes on electron microscopy examination of kidney biopsy, including thinning, thickening, lamellation and presence of electron-dense bodies.

Supplementary figure S6



Supplementary figure S6. Pedigrees of families 7 and 11, with X-linked Alport syndrome segregating with the *COL4A5* mutation p.Gly1448Arg, in exon 47, showing the individual haplotypes for a set of four polymorphic microsatellite markers adjacent to the *COL4A5* and *COL4A6* genes. The numbers indicate the size of each allele, in base pairs.

Supplementary table S7. Correlation between genotype and phenotype in Portuguese male patients with truncating *COL4A5* mutations (n=14).

FAMILY	**15**	**15**	**1**	**1**	**6**	**33**	**41**	**19**	**60**	**60**	**60**	**9**	**31**	**45**
PATIENT	15/AS-2	**15/AS-1**	**1/AS-1**	1/AS-2	**6/AS-1**	33/AS-1	**41/AS-1**	**19/AS-1**	60/AS-4	**60/AS-1**	60/AS-5	**9/AS-1**	**31/AS-1**	**45/AS-1**
GENOTYPE														
Exon/intron	10	**10**	**30**	30	**47**				Intron 19	**Intron 19**	Intron 19	**Intron 20**	**Intron 46**	**Intron 49**
Nucleotide Change	c.598delC	**c.598delC**	c.2423delG	c.2423delG	**c.4444C>T**	del ex 1_13 +(del ex 1_2 COL4A5)	**del ex.2_29**	**del ex.2_31**	c.1165+1G>A	**c.1165+1G>A**	c.1165+1G>A	**c.1959+6C>G**	**C.4297+1G>A**	**c.4803+1 G>A**
Predicted effect on the protein	p.Pro197Glnfs*6	**p.Pro197Glnfs*6**	**p.Gly808Aspfs*11**	p.Gly808Aspfs*11	**p.Gln1482***				ND	**ND**	ND	**ND**	**ND**	**ND**
Mutation type	Frameshift	**Frameshift**	**Frameshift**	Frameshift	**Nonsense**	Large deletion	**Large deletion**	**Large deletion**	Splice site	**Splice site**	Splice site	**Splice site**	**Splice site**	**Splice site**
PHENOTYPE														
Gender, age at enrollment in years	M [35]	**M [29]**	**M [27]**	M [27]	**M [22]**	M [17]	**M [39]**	**M [25]**	M [29]	**M [27]**	M [21]	**M [49]**	**M [39]**	**M [39]**
Birth District	Faro	**Faro**	**Braga**	Braga	**Porto**	Setúbal	**Santarém**	**Lisboa**	Madeira	**Madeira**	Madeira	**Porto**	**Aveiro**	**Porto**
Affected relatives	+	**+**	**+**	+	**+**	+	**+**	**+**	+	**+**	+	**+**	**+**	**+**
Renal anomalies														
History of macroscopic hematuria[a]	NA	**- [29]**	NA	NA	**- [22]**	[19]	**+ [15]**	**+ [2]**	NA	**+ [2]**	NA	**+ [6]**	**- [53]**	NA
History of microscopic hematuria[a]	NA	**+ [10]**	**+ [18]**	+ [18]	**+ [18]**	+ [5]	**+ [15]**	**+ [2]**	+ [NA]	**NA**	+ [NA]	**+ [6]**	**+ [51]**	**+ [NA]**
History of proteinuria[c]	NA	**+ [14]**	**+ [18]**	+ [18]	**+ [18]**	+ [5]	**+ [25]**	**+ [7]**	+ [NA]	**+ [2]**	+ [NA]	**+ [25]**	**+ [51]**	**+ [NA]**
History of hypertension[a]	+ [NA]	**+ [21]**	NA	+ [18]	**+ [18]**	- [19]	**+ [20]**	**+ [NA]**	+ [NA]	**NA**	+ [NA]	**+ [25]**	**+ [<40]**	**+ [NA]**
History of CKD stage 2 or higher[a]	+ [NA]	**+ [21]**	**+ [18]**	+ [18]	**+ [18]**	NA	**+ [15]**	**+ [NA]**	+ [NA]	**+ [18]**	+ [NA]	**+ [26]**	**+ [51]**	**+ [NA]**
pCr (eGFR) at diagnosis of CKD stage 2 or higher	-	**3.4 (25)**	**7.7 (9)**	4.3 (19)	**1.6 (57)**	NA	**NA**	**NA**	NA	**NA**	NA	**1.7 (49)**	**3.9 (16)**	NA
History of renal replacement therapy[b]	HD [NA], Tx [21]	**PD [22], Tx [29]**	**HD [18], Tx [19]**	HD [19], Tx [19]	**DP [21], Tx [28]**	- [17]	**HD [30], Tx [33]**	**HD [20], Tx [21]**	- [79]	**HD [18]**	- [71]	**PD [28], Tx [29]**	**PD [53]**	**HD [32], Tx [33]**
pCr (eGFR) at onset of renal replacement therapy	NA	**NA**	**7.7 (9)**	7.3 (10)	**7 (10)**		**NA**	**9.0 (8)**		**NA**		**9 (7.2)**	**NA**	**8.6 [7]**
Hearing loss														
Self-noticed (subjective)[a]	+ [NA]	**+ [NA]**	**+ [6]**	+ [5]	**+ [10]**	+ [5]	**+ [8]**	**+ [6]**	+ [NA]	**+ [15]**	+ [NA]	**+ [<20]**	**+ [NA]**	**+ [13]**
Sensorineural hearing loss confirmed by audiogram[a]	NA	**NA**	**+ [18]**	+ [18]	**+ [20]**	+ [6]	**+ [20, 33]**	**+ [6]**	NA	**NA**	NA	**+ [25]**	**NA**	**+ [37]**
Ocular signs														
Anterior lenticonus[a]	NA	**NA**	**+ [19]**	+ [19]	**- [24]**	- [17]	**- [25]**	**NA**	NA	**+ [15]**	NA	**NA**	**- [51]**	**+ [34]**
Maculopathy[a]	NA	**NA**	**+ [19]**	+ [19]	**+ [24]**	- [17]	**- [25]**	**+ [26]**	+ [NA]	**+ [15]**	+ [NA]	**+ [33]**	**- [51]**	**+ [40]**
Cataracts[a]	NA	**NA**	**+ [19]**	- [30]	**- [24]**	+ [5]	**+ [25]**	**- [26]**	NA	**NA**	NA	**NA**	**- [51]**	**+ [34]**
GBM ultrastructural changes														
Thinning[c]	NA	**NA**	**+ [18]**	NA	**+ [19]**	NA [5]	**NA**	**+ [6]**	NA	**NA**	NA	**NA [28]**	**NA**	**NA [18]**
Thickening[c]	NA	**NA**	**+ [18]**	NA	**+ [19]**	NA [5]	**NA**	**+ [6]**	NA	**NA**	NA	**NA [28]**	**NA**	**NA [18]**
Lamellation[c]	NA	**NA**	**+ [18]**	NA	**+ [19]**	NA [5]	**NA**	**+ [6]**	NA	**NA**	NA	**NA [28]**	**NA**	**NA [18]**
Electron dense bodies[c]	NA	**NA**	**+ [18]**	NA	**+ [19]**	NA [5]	**NA**	**+ [6]**	NA	**NA**	NA	**NA [28]**	**NA**	**NA [18]**
EBM reactivity for the α5(IV) chain[d]	absent [28]	**absent [22]**	NA	NA	**NA**	NA	**NA**	**NA**	NA	**NA**	NA	**NA**	**NA**	**NA**
Diffuse leiomyomatosis[a]		**-**	**-**		**-**		**-**	**+ [25]**		**-**		**-**	**-**	**-**

[a] In brackets, the age at diagnosis or at last screening, in years. [b] In brackets, the age at start of hemodialysis (HD), peritoneal dialysis (PD) or of kidney transplant (Tx). [c] In brackets, the age at kidney biopsy. [d] In brackets, the age at skin biopsy. < Before the age indicated within brackets. CKD: chronic kidney disease. pCr: plasma creatinine level, in micromol/L. Conversion of micromoles/L (SI units) into mg/dL may be performed, dividing SI units by the conversion factor 88.4. eGFR: estimated glomerular filtration rate, in mL/min/1.73m^2. GBM: glomerular basement membrane. EBM: epidermal basement membrane. ND: effect on splicing not determined by mRNA analysis. Genetic probands are indicated in bold.

Supplementary table S8. Correlation between genotype and phenotype in Portuguese male patients with non-truncating *COL4A5* mutations (n=16).

FAMILY	53	23	48	48	48	48	57	57	50	7	7	7	11	28	34	34
PATIENT	53/AS-1	23/AS-1	48/AS-3	48/AS-9	48/AS-7	48/AS-1	57/AS-3	57/AS-1	50/AS-1	7/AS-3	7/AS-1	7/AS-2	11/AS-1	28/AS-1	34/AS-2	34/AS-6
GENOTYPE																
Exon/intron	8	19	18	18	18	18	24	24	31	47	47	47	47	Intron 30	Intron 30	Intron 30
Nucleotide Change	c.458G>A	c.715G>A	c.1009G>A	c.1009G>A	c.1009G>A	c.1009G>A	c.1719G>A	c.1719G>A	c.2633G>T	c.4342G>C	c.4342G>C	c.4342G>C	c.4342G>C	c.2510_2554del	c.2510-1_2523del	c.2510-1_2523del
Predicted effect on the protein	p.Gly153Asp	p.Gly239Arg	p.Gly337Ser	p.Gly337Ser	p.Gly337Ser	p.Gly337Ser	p.Gly573Asp	p.Gly573Asp	p.Gly878Val	p.Gly1448Arg	p.Gly1448Arg	p.Gly1448Arg	p.Gly1448Arg	ND	ND	ND
Mutation type	missense	missense	missense	missense	missense	missense	missense	missense	missense	missense	missense	missense	missense	Splice site	Splice site	Splice site
PHENOTYPE																
Gender, age at enrollment in years	M [31]	M [32]	M [78]	M [72]	M [44]	M [34]	M [38]	M [5]	M [24]	M [51]	M [48]	M [44]	M [49]	M [38]	M [67]	M [8]
Birth District	Guarda	Castelo Branco	Coimbra	Coimbra	Coimbra	Coimbra	Lisboa	Lisboa	Braga	Porto	Porto	Porto	Porto	Bragança	Vila Real	Vila Real
Affected relatives	+	+	-	-	-	+	-	+	No	+	-	-	+	+	-	-
Renal anomalies																
History of macroscopic hematuria[a]	+ [5]	+ [5-10]	- [78]	- [72]	+ [19]	+ [3]	+ [4]	- [5]	+ [5]	NA	+ [NA]	NA	NA	+ [6]	- [NA]	- [8]
History of microscopic hematuria[a]	+ [5]	+ [5]	+ [35]	- [31]	+ [34]	+ [5]	+ [4]	+ [2]	+ [5]	+ [25]	+ [31]	+ [23]	NA	NA	NA	+ [8]
History of proteinuria[a]	NA	+ [20]	+ [64]	- [31]	- [34]	+ [26]	+ [4]	+ [4]	+ [5]	+ [28]	+ [31]	+ [23]	NA	NA	- [NA]	- [8]
History of hypertension[a]	+ [12]	+ [18]	+ [50]	- [31]	- [44]	- [34]	- [38]	- [5]	+ [20]	-	+ [31]	NA	NA	+ [6]	- [NA]	- [8]
History of CKD stage 2 or higher[a]	+ [16]	+ [20]	+ [63]	- [NA]	- [44]	- [34]	+ [NA]	- [5]	+ [20]	+ [28]	+ [31]	+ [23]	+ [28]	+ [NA]	- [NA]	- [8]
pCr (eGFR) at diagnosis of CKD stage 2 or higher	NA	2.0(43)	2.8 (23)	NA	-	-	NA	-	2.1 (40)	1.0 (89)	1.3 (64)	2.1 (39)	NA	NA	NA	-
History of renal replacement therapy[b]	HD [17]	HD [20], Tx [21]	HD [64], Tx [65]	HD [47], Tx [48]	[44]	- [34]	HD [15]	- [5]	PD [29]	HD [45], Tx [52]	PD [44], Tx [48]	HD [25], Tx [28]	HD [28], Tx [30]	HD [17], Tx [24]	HD [NA], Tx [65]	- [8]
pCr (eGFR) at onset of renal replacement therapy	NA	5.1 (15)	9.8 (5)	10.1 (6)	--	--	NA	--	4.8 (18)	8.78 (7)	4.5 (14)	NA	NA	NA	NA	
Hearing loss																
Self-noticed (subjective)[a]	- [31]	+ [10-20]	+ [60]	- [36]	+ [37]	- [34]	+ [15]	- [5]	+ [NA]	+ [28]	+ [23]	+ [NA]	+ [NA]	+ [NA]	- [50]	- [8]
Sensorineural hearing loss confirmed by audiogram[a]	NA	+ [31]	+ [64]	- [49, 55]	+ [37]	No audiogram	+ [35]	- [5]	NA	No audiogram	+ [30]	+ [33]	No audiogram	NA	NA	+[5]
Ocular signs																
Anterior lenticonus[a]	- [31]	NA	- [67]	NA	NA	NA	- [38]	NA	+ [24]	- [52]	- [48]	- [46]	NA	NA	NA	NA
Maculopathy[a]	- [31]	NA	- [67]	NA	NA	NA	- [38]	NA	- [24]	- [52]	- [48]	+ [46]	- [51]	NA	NA	NA
Cataracts[a]	NA	NA	+ [67]	NA	NA	NA	NA	NA	- [24]	- [52]	- [48]	+ [46]	NA	NA	NA	NA
GBM ultrastructural changes																
Thinning[c]	NA [16]	NA	NA	NA	NA	+ [33]	NA [5]	NA	+ [21]	+ [28]	NA	NA	NA	NA [7]	NA	NA
Thickening[c]	NA [15]	NA	NA	NA	NA	+ [33]	NA [5]	NA	+ [21]	+ [28]	NA	NA	NA	NA [7]	NA	NA
Lamellation[c]	NA [15]	NA	NA	NA	NA	+ [33]	NA [5]	NA	+ [21]	+ [28]	NA	NA	NA	NA [7]	NA	NA
Electrondense bodies[c]	NA [15]	NA	NA	NA	NA	+ [33]	NA [5]	NA	+ [21]	+ [28]	NA	NA	NA	NA [7]	NA	NA
EBM reactivity for the α5(IV) chain[d]	NA	NA	NA	NA	NA	NA	absent [27]	NA	NA	NA	NA	NA	NA	NA	NA	NA
Diffuse leiomyomatosis[a]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

NA: Data not available. [a] In brackets, the age at diagnosis or at last screening, in years. [b] In brackets, the age at start of hemodialysis (HD), peritoneal dialysis (PD) or of kidney transplant (Tx). [c] In brackets, the age at kidney biopsy. [d] In brackets, the age at skin biopsy. < Before the age indicated within brackets. CKD: chronic kidney disease. pCr: plasma creatinine level, in micromol/L. Conversion of micromoles/L (SI units) into mg/dL may be performed, dividing SI units by the conversion factor 88.4. eGFR: estimated glomerular filtration rate, in mL/min/1.73m^2. GBM: glomerular basement membrane. EBM: epidermal basement membrane. ND: effect on splicing not determined by mRNA analysis. Genetic probands are indicated in bold.

Supplementary table S9. Correlation between genotype and phenotype in Portuguese female patients with truncating *COL4A5* mutations (n=13).

FAMILY	**1**	**1**	**26**	**6**	**6**	**6**	**33**	**41**	**19**	**19**	**18**	**60**	**3**
PATIENT	**1/AS-4**	**1/AS-5**	**26/AS-1**	**6/AS-2**	**6/AS-3**	**6/AS-4**	**33/AS-2**	**41/AS-2**	**19/AS-2**	**19/AS-3**	**18/AS-1**	**60/AS-2**	**3/AS-2**
GENOTYPE													
Exon/intron	30	30	33	47	47	47						Intron 19	Intron 20
Nucleotide Change	c.2412delG	c.2412delG	**c.2915G>T**	c.4444C>T	c.4444C>T	c.4444C>T	**del ex.1_15 + [del ex.1_2 *COL4A6*)**	del ex.2_25	del ex.2_51	del ex.2_51	**del ex.45_45**	c.1165+1G>A	c.1339+5C>G
Predicted effect on the protein	p.Gly805Aspfs*11	p.Gly805Aspfs*11	**p.Glu939***	p.Gln1482*	p.Gln1482*	p.Gln1482*						ND	ND
Mutation type	Frameshift	Frameshift	**Nonsense**	Nonsense	Nonsense	Nonsense	**Large deletion**	Large deletion	Large deletion	Large deletion	**Large deletion**	Splice site	Splice site
PHENOTYPE													
Gender, age at enrollment in years	F [46]	F [45]	**F [33]**	F [60]	F [35]	F [29]	**F [46]**	F [58]	F [43]	F [19]	**F [27]**	F [48]	F [11]
Birth District	Braga	Vila Real	**Viseu**	Porto	Porto	Porto	**Setúbal**	Santarém	Lisboa	Lisboa	**Braga**	Madeira	Porto
Affected relatives	+	+	**+**	+	+	+	**+**	+	+	+	**No**	+	+
Renal anomalies													
History of macroscopic hematuria[a]	NA	- [46]	**- [33]**	NA	- [30]	- [29]	**- [46]**	NA	- [44]	+[0.5]	**- [27]**	NA	+ [9]
History of microscopic hematuria[a]	+ [37]	+ [20]	**+ [33]**	+ [47]	+ [32]	+ [10]	**+ [9]**	+ [50]	+ [43]	+[0.5]	**+ [16]**	+ [NA]	+ [9]
History of proteinuria[a]	+ [37]	+ [20]	**+ [16]**	+ [47]	+ [32]	+ [10]	**+ [9]**	+ [NA]	- [44]	+[0.5]	**+ [15]**	+ [NA]	+ [25]
History of hypertension[a]	+ [35]	+ [24]	**- [33]**	+ [47]	- [32]	- [29]	**+ [33]**	+ [50]	- [44]	- [19]	**- [24]**	+ [NA]	+ [26]
History of CKD stage 2 or higher[a]	+ [37]	+ [40]	**+ [30]**	+ [62]	+ [32]	+ [29]	**+ [33]**	+ [50]	- [44]	+[9]	**+ [24]**	+ [NA]	+ [26]
pCr (eGFR) at diagnosis of CKD stage 2 or higher	1.0 (62)	NA	**1.1 (58)**	1.24 (44)	0.79 (94)	0.8 (85)	**1.2 (50)**	1.1 (51)	--	1.0 (72)	**0.84 (83)**	NA	1.1 (60)
History of renal replacement therapy[b]	- [46]	- [46]	**- [33]**	- [62]	- [30]	- [29]	**HD [46]**	- [60]	- [44]	- [19]	**- [24]**	- [48]	HD [31], Tx [33]
pCr (eGFR) at onset of renal replacement therapy	--	--	**--**	--	--	--	**8.8 (5)**	--	--	--	**--**	--	NA
Hearing loss													
Self-noticed (subjective)[b]	+ [38]	+ [NA]	**- [34]**	+ [50]	+ [NA]	- [30]	**+ [NA]**	- [60]	+ [40]	+[16]	**+ [23]**	- [48]	+ [31]
Sensorineural hearing loss confirmed by audiogram[c]	+ [38]	+ [41]	**- [34]**	+ [NA]	+ [36]	- [30]	**- [34], + [46]**	NA	+ [40, 42, 44]	+[16]	**- [19], + [23,24,25]**	NA	+ [37]
Ocular signs													
Anterior lenticonus[c]	- [48]	NA	NA	- [61]	- [36]	- [30]	**- [45]**	NA	NA	- [20]	**- [26]**	NA	- [44]
Maculopathy[c]	- [48]	NA	NA	- [61]	- [36]	+ [30]	**- [45]**	NA	+ [45]	+ [20]	**- [28]**	- [48]	+ [44]
Cataracts[a]	- [49]	NA	NA	- [61]	- [36]	- [30]	**- [45]**	NA	- [45]	+[20]	**- [28]**	NA	- [44]
GBM ultrastructural changes													
Thinning[c]	NA	NA	**NA [16]**	NA	NA	NA	**NA [8]**	NA	NA	NA	**+ [17]**	NA	+ [24]
Thickening[c]	NA	NA	**NA [16]**	NA	NA	NA	**NA [8]**	NA	NA	NA	**+ [17]**	NA	+ [24]
Lamellation[c]	NA	NA	**NA [16]**	NA	NA	NA	**NA [8]**	NA	NA	NA	**+ [17]**	NA	+ [24]
Electrondense bodies[c]	NA	NA	**NA [16]**	NA	NA	NA	**NA [8]**	NA	NA	NA	**+ [17]**	NA	+ [24]
EBM reactivity for the α5(IV) chain[d]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Diffuse leiomyomatosis[a]	-	-	-	-	-	-	-	-	+ [<18]	+[19]	-	-	-

NA: Data not available. [a] In brackets, the age at diagnosis or at last screening, in years. [b] In brackets, the age at start of hemodialysis (HD), peritoneal dialysis (PD) or of kidney transplant (Tx). [c] In brackets, the age at kidney biopsy. [d] In brackets, the age at skin biopsy. < Before the age indicated within brackets. CKD: chronic kidney disease. pCr: plasma creatinine level, in micromol/L. Conversion of micromoles/L (SI units) into mg/dL may be performed, dividing SI units by the conversion factor 88.4. eGFR: estimated glomerular filtration rate, in mL/min/1.73m^2. GBM: glomerular basement membrane. EBM: epidermal basement membrane. ND: effect on splicing not determined by mRNA analysis. Genetic probands are indicated in bold.

Supplementary table S10. Correlation between genotype and phenotype in Portuguese female patients with truncating *COL4A5* mutations (n=13).

FAMILY	53	53	53	48	48	48	48	57	57	57	57	57	49	7	7	7	7	11	11	34	34	34
PATIENT	53/AS-2	53/AS-4	53/AS-5	48/AS-8	48/AS-2	48/AS-12	48/AS-10	57/AS-4	57/AS-7	57/AS-2	57/AS-5	57/AS-6	**49/AS-1**	7/AS-4	7/AS-7	7/AS-9	7/AS-5	11/AS-2	11/AS-3	**34/AS-1**	34/AS-4	34/AS-5
GENOTYPE																						
Exon/Intron	8	8	8	18	18	18	18	24	24	24	24	24	**25**	47	47	47	47	47	47	**Intron 30**	Intron 30	Intron 30
Nucleotide Change	c.458G>A	c.458G>A	c.458G>A	c.1008G>A	c.1008G>A	c.1008G>A	c.1008G>A	c.1719G>A	c.1719G>A	c.1719G>A	c.1719G>A	c.1719G>A	**c.1844G>A**	c.4342G>C	c.4342G>C	c.4342G>C	c.4342G>C	c.4342G>C	c.4342G>C	**c.2510-1_2523del**	c.2510-1_2523del	c.2510-1_2523del
Predicted effect on the protein Mutation type	p.Gly153Asp missense	p.Gly153Asp missense	p.Gly153Asp missense	p.Gly337Ser missense	p.Gly337Ser missense	p.Gly337Ser missense	p.Gly337Ser missense	p.Gly573Asp missense	p.Gly573Asp missense	p.Gly573Asp missense	p.Gly573Asp missense	p.Gly573Asp missense	**p.Gly615Glu missense**	p.Gly1448Arg missense	p.Gly1448Arg missense	p.Gly1448Arg missense	p.Gly1448Arg missense	p.Gly1448Arg missense	p.Gly1448Arg missense	**ND Splice site**	ND Splice site	ND Splice site
PHENOTYPE																						
Gender, age at enrollment in years	F [56]	F [58]	F [49]	F [66]	F [55]	F [34]	F [30]	F [60]	F [48]	F [32]	F [31]	F [4]	**F [42]**	F [76]	F [47]	F [41]	F [21]	F [47]	F [16]	**F [40]**	F [39]	F [38]
Birth District	Guarda	Guarda	Guarda	Coimbra	Coimbra	Coimbra	Coimbra	Lisboa	Lisboa	Lisboa	Lisboa	Lisboa	**Porto**	Porto	Porto	Porto	Porto	Porto	Porto	**Vila Real**	Vila Real	Vila Real
Affected relatives	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Renal anomalies																						
History of macroscopic hematuria[a]	NA	- [58]	- [49]	- [66]	+ [12]	- [34]	- [30]	- [60]	- [48]	- [32]	NA	- [4]	NA	- [76]	NA	- [41]	NA	NA	- [16]	NA	NA	+ [NA]
History of microscopic hematuria[a]	+ [21]	+ [23]	+ [21]	+ [38]	+ [12]	+ [24]	+ [5]	+ [26]	+ [49]	+ [2]	+ [NA]	+ [4]	+ [NA]	+ [NA]	+ [NA]	+ [12]	+ [2]	+ [27]	+ [10]	+ **[28]**	+ [NA]	+ [NA]
History of proteinuria[a]	NA	+ [23]	+ [23]	+ [38]	+ [12]	+ [24]	NA	- [60]	+ [49]	- [32]	+ [31]	- [4]	+ **[92]**	- [75]	- [46]	+ [18]	+ [17]	+ [27]	- [16]	+ **[28]**	+ [NA]	+ [35]
History of hypertension[a]	+ [54]	+ [30]	+ [23]	+ [40]	+ [32]	+ [24]	- [30]	+ [55]	- [48]	- [32]	- [31]	- [4]	+ **[88]**	+ [70]	NA	- [41]	- [21]	- [47]	- [16]	- **[40]**	+ [NA]	+ [30]
History of CKD stage 2 or higher[a]	- [56]	+ [45]	+ [21]	+ [38]	+ [20]	+ [24]	NA	- [60]	+ [49]	- [32]	- [31]	- [4]	- **[41]**	+ [72]	- [47]	- [39]	- [21]	+ [27]	- [16]	- **[41]**	- [39]	+ [30]
pCr (eGFR) at diagnosis of CKD stage 2 or higher	- -	0.9 (78)	1.1 (58)	1.2 (50)	1.07 (65)	2.2 (27)	NA	- -	0.85 (71)	NA	- -	- -	- -	1.03 (59)	NA	- -	- -	0.9 (86)	- -	- -	- -	1.0 (65)
History of renal replacement therapy[b]	- [56]	- [53]	HD [41], Tx [42]	HD [48], Tx [49]	HD [39], Tx [40]	HD [33]	- [30]	- [60]	- [49]	- [32]	- [31]	- [4]	- **[41]**	- [76]	- [47]	- [41]	- [21]	- [48]	- [16]	- **[41]**	- [39]	- [38]
pCr (eGFR) at onset of renal replacement therapy	- -	- -	6.6 (7)	7.9 (6)	7.4 (6)	7.6 [6]	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -
Hearing loss																						
Self-noticed (subjective)[a]	+ [NA]	- [53]	+ [NA]	+ [39]	- [55]	NA	- [30]	NA	- [48]	- [32]	- [31]	- [4]	- **[41]**	+ [NA]	+ [46]	- [41]	- [21]	+ [46]	- [16]	- **[36]**	- [39]	- [38]
Sensorineural hearing loss confirmed by audiogram[a]	NA	NA	- [23], + [47]	+ [57]	- [53]	No audiogram	No audiogram	NA	NA	NA	NA	NA	- **[39]**	+ [76]	+ [46]	NA	- [22]	- [28], + [46]	- [12,16]	- **[36]**	NA	NA
Ocular signs																						
Anterior lenticonus[a]	NA	NA	- [23]	NA	- [53]	NA	NA	- [60]	NA	NA	NA	NA	- **[42]**	- [76]	- [49]	- [40]	- [21]	- [48]	- [16]	NA	NA	NA
Maculopathy[a]	NA	NA	- [23]	NA	- [53]	NA	NA	- [60]	NA	NA	NA	NA	- **[42]**	NA	+ [49]	- [40]	- [21]	- [48]	- [16]	NA	NA	NA
Cataracts[a]	+ [NA]	- [53]	- [23]	NA	- [53]	NA	NA	NA	NA	NA	NA	NA	- **[42]**	+ [74]	- [49]	- [40]	- [21]	- [48]	- [16]	NA	NA	NA
GBM ultrastructural changes																						
Thinning[c]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA [18]	NA [10]	+ [28]	NA	NA	NA	NA
Thickening[c]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA [18]	NA [10]	+ [28]	NA	NA	NA	NA
Lamellation[c]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA [18]	NA [10]	+ [28]	NA	NA	NA	NA
Electrondense bodies[c]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA [18]	NA [10]	+ [28]	NA	NA	NA	NA
EBM reactivity for the α6(IV) chain[d]	NA	NA	NA	NA	NA	NA	NA	Intermittent [49]	Intermittent [37]	Normal [20]	Intermittent [20]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Diffuse leiomyomatosis[a]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

[a] In brackets, the age at diagnosis or at last screening, in years. [b] In brackets, the age at start of hemodialysis (HD), peritoneal dialysis (PD) or of kidney transplant (Tx). [c] In brackets, the age at kidney biopsy. [d] In brackets, the age at skin biopsy. < Before the age indicated within brackets. CKD: chronic kidney disease. pCr: plasma creatinine level, in micromol/L. Conversion of micromoles/L (SI units) into mg/dL may be performed, dividing SI units by the conversion factor 88.4. eGFR: estimated glomerular filtration rate, in mL/min/1.73m^2. GBM: glomerular basement membrane. EBM: epidermal basement membrane. ND: effect on splicing not determined by mRNA analysis. Genetic probands are indicated in bold.

Supplementary table S11. Phenotypic comparisons between hemizygous males with truncating and non-truncating pathogenic *COL4A5* mutations (n=30).

	Truncating (n=14)		Non-truncating (n=16)		p-value
		N		N	
PHENOTYPE					
Age at enrollment (median (interquartile range))	28 (11.8)	14	41 (19)	16	0.046
History of renal abnormalities					
History of macroscopic hematuria (%) Age at diagnosis (median (interquartile range))	50.0 3.5 (10.5)	4/8 4	69.2 5 (3)	9/13 7	0.646 0.527
History of microscopic hematuria (%) Age at diagnosis (median (interquartile range))	100.0 15 (13)	12/12 9	100.0 8 (27.5)	13/13 13	a 0.969
History of proteinuria (%) Age at diagnosis (median (interquartile range))	100.0 16.5 (13)	13/13 10	92.3 23 (26)	12/13 11	1.000 0.172
History of hypertension (%) Age at diagnosis (median (interquartile range))	91.7 20 (5)	11/12 5	57.1 20 (19)	8/14 7	0.081 0.417
History of CKD stage 2 or higher (%) Age at diagnosis (median (interquartile range)) eGFR at diagnosis (median (interquartile range))	100.0 18 (7) 23 (39)	13/13 8 7	75.0 26 (10) 41.5 (35)	12/16 8 6	0.107 0.204 0.224
History of renal replacement therapy (%) Age at onset (median (interquartile range)) eGFR at onset (median (interquartile range))	78.6 21.5 (12) 9 (2.3)	11/14 10 6	75.0 25 (28) 10.5 (10)	12/16 11 6	1.000 0.512 0.936
GBM ultrastructural abnormalities					
Age at kidney biopsy (median (interquartile range))	18 (14*)	3	28 (12*)	3	0.025
Thinning (%)	100.0	3/3	100.0	3/3	a
Thickening (%)	100.0	3/3	100.0	3/3	a
Lamellation (%)	100.0	3/3	100.0	3/3	a
Electrondense bodies (%)	100.0	3/3	100.0	3/3	a
Hearing loss					
Self-noticed or subjective (%) Age at self-noticed or subjective (median (interquartile range))	100.0 7 (6.3)	14/14 8	75.0 36 (27)	12/16 7	0.103 <0.001
Audiogram (%) Age at diagnosis of hearing loss by audiogram (median (interquartile range))	100.0 19 (15)	8/8 8	88.9 33 (15.8)	8/9 8	1.000 0.027
Ocular abnormalities					
Anterior Lenticonus (%) Age at diagnosis (median (interquartile range))	50.0 19 (14)	4/8 4	14.3 24 (0*)	1/7 1	0.282 0.806
Maculopathy (%) Age at diagnosis (median (interquartile range))	75.0 24 (14)	9/12 7	12.5 46 (0*)	1/8 1	0.020 0.112
Cataracts (%) Age at diagnosis (median (interquartile range))	50.0 22 (23.3)	4/8 4	40.0 56.5 (21*)	2/5 2	1.000 0.049
Other abnormalities					
Leiomyomatosis (%) Age at diagnosis (median (interquartile range))	7.1 24 (0*)	1/14 1	0.0 -	0/16 -	0.467 -

SD: standard deviation. CKD: chronic kidney disease. eGFR: estimated glomerular filtration rate. GBM: glomerular basement membrane. Age is expressed in years. eGFR is expressed in ml/min/1.73m^2. *The range is indicated instead of the interquartile range. [a]Statistics not computed because the proportions in the comparison groups are the same.

Supplementary table S12. Phenotypic comparisons between heterozygous females with truncating and non-truncating pathogenic *COL4A5* mutations (n=35).

	Truncating (n=13)		Non-truncating (n=22)		p-value
		N		N	
PHENOTYPE					
Age at enrollment (median (interquartile range))	43 (16)	13	41.5 (22)	22	0.720
History of renal abnormalities					
History of macroscopic hematuria (%) Age at diagnosis (median (interquartile range))	22.2 1.75 (2.5*)	2/9 2	14.3 12 (-)	2/14 1	1.000 0.225
History of microscopic hematuria (%) Age at diagnosis (median (interquartile range))	100.0 26 (37)	13/13 12	100.0 21 (20.5)	22/22 16	a 0.271
History of proteinuria (%) Age at diagnosis (median (interquartile range))	92.3 19.5 (29.3)	12/13 10	70.0 27 (13)	14/20 13	0.202 0.717
History of hypertension (%) Age at diagnosis (median (interquartile range))	53.8 35 (22.3)	7/13 6	52.4 32.5 (26)	11/21 10	0.934 0.578
History of CKD stage 2 or higher (%) Age at diagnosis (median (interquartile range)) eGFR at diagnosis (median (interquartile range))	92.3 32 (14) 60 (33)	12/13 11 11	42.9 31 (21.5) 65 (23)	9/21 9 9	0.005 0.684 0.939
History of renal replacement therapy (%) Age at onset (median (interquartile range)) eGFR at onset (median (interquartile range))	15.4 38.5 (15*) 5 (0*)	2/13 2 1	18.2 40 (11.8) 6 (0.8)	4/22 4 4	1.000 0.707 0.114
GBM ultrastructural abnormalities					
Age at kidney biopsy (median (interquartile range))	20.5 (7*)	2	28 (0*)	1	0.225
Thinning (%)	100.0	2/2	100.0	1/1	a
Thickening (%)	100.0	2/2	100.0	1/1	a
Lamellation (%)	100.0	2/2	100.0	1/1	a
Electrondense bodies (%)	100.0	2/2	100.0	1/1	a
Hearing loss					
Self-noticed or subjective (%) Age at self-noticed or subjective (median (interquartile range))	69.2 34.5 (21.3)	9/13 6	30.0 46 (7*)	6/20 3	0.027 0.407
Audiogram (%) Age at diagnosis of hearing loss by audiogram (median (interquartile range))	81.8 37.5 (14.5)	9/11 8	50.0 47 (20.5)	5/10 5	0.183 0.003
Ocular abnormalities					
Anterior Lenticonus (%) Age at diagnosis (median (interquartile range))	0.0 -	0/8 -	0.0 -	0/10 -	a b
Maculopathy (%) Age at diagnosis (median (interquartile range))	40.0 37 (22.3*)	4/10 4	11.1 49 (0*)	1/9 1	0.303 0.156
Cataracts (%) Age at diagnosis (median (interquartile range))	11.1 20 (0*)	1/9 1	18.2 66 (20*)	2/11 2	1.000 0.157
Other abnormalities					
Leiomyomatosis (%) Age at diagnosis (median (interquartile range))	15.4 18 (0*)	2/13 1	0.0 -	0/22 -	0.131 -

SD: standard deviation. CKD: chronic kidney disease. eGFR: estimated glomerular filtration rate. GBM: glomerular basement membrane. Age is expressed in years. eGFR is expressed in ml/min/1.73m^2. *The range is indicated instead of the interquartile range. [a]Statistics not computed because the proportions in the comparison groups are the same.

2.1.2. Deletion of the 5'exons of *COL4A6* is not needed for the development of diffuse leiomyomatosis in Alport syndrome patients (Manuscript 2)

Published in J Med Genet. 2013 Nov;50(11):745-53.

Deletion of the 5'exons of *COL4A6* is not needed for thedevelopment of diffuse leiomyomatosis in Alport syndrome patients

Maria João Nabais Sá[1,2], Nathalie Fieremans[3], Arjan P.M. de Brouwer[4], Rita Sousa[5], Fernando Teixeira e Costa[6], Maria José Brito[7], Fernanda Carvalho[8], Márcia Rodrigues[9], Francisco Teixeira de Sousa[10], Joana Felgueiras[10], Fernando Neves[11], Adelino Carvalho[11], Umbelina Ramos[12], José Ramón Vizcaíno[12], Susana Alves[1], Filipa Carvalho[1], Guy Froyen[3]*, João Paulo Oliveira[1,2]*

[1] Department of Genetics, Faculty of Medicine, University of Porto, Porto, Portugal
[2] Unit of Research and Development in Nephrology (FCT-725), Faculty of Medicine, University of Porto, Porto, Portugal
[3] Human Genome Laboratory, Department of Human Genetics, VIB Center for the Biology of Disease, KU Leuven, Leuven, Belgium
[4] Department of Human Genetics*, Institute for Genetic and Metabolic Disease,* Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands
[5] Department of Gastroenterology, [6] Department of Nephrology and [7] Department of Pathology, Hospital Garcia de Orta, Almada, Portugal
[8] Unit of Renal Morphology, Department of Nephrology, Hospital Curry Cabral
[9] Department of Genetics, Hospital Dona Estefânia, Centro Hospitalar de Lisboa Central, Lisboa, Portugal
[10] Nephrocare Montijo, Fresenius Medical Care Portugal, Montijo, Portugal
[11] Nephrocare Santarém, Fresenius Medical Care Portugal, Santarém, Portugal
[12] Department of Pathology, Hospital de Santo António – Centro Hospitalar do Porto, Porto, Portugal
* Both authors contributed equally to this project and should be considered co-last authors.

Corresponding author:
Maria João Nabais Sá, Department of Genetics, Faculty of Medicine, University of Porto, Alameda Prof. Hernâni Monteiro, 4200 - 319 Porto,

Portugal. E-mail: mjs.jano@gmail.com. Telephone: +351965546341. Fax: +351225513648.

ABSTRACT

Background

Alport syndrome (AS), a hereditary type IV collagen nephropathy, is a major cause of end-stage renal disease in young people. About 85% of the cases are X-linked (ATS), due to mutations in the *COL4A5* gene. Rarely, families have a contiguous gene deletion comprising at least exon 1 of *COL4A5* and the first exons of *COL4A6*, associated with the development of diffuse leiomyomatosis (ATS-DL). We report three novel deletions identified in families with AS, one of which challenges the current concepts on genotype-phenotype correlations of ATS/ATS-DL.

Methods

In the setting of a multicentric study aiming to describe the genetic epidemiology and molecular pathology of AS in Portugal, three novel *COL4A5* deletions were identified in two families with ATS and in one family with ATS-DL. These mutations were initially detected by PCR and MLPA, and further mapped by high-resolution X chromosome-specific oligo-array and PCR.

Results

In the ATS-DL family, a *COL4A5* deletion spanning exons 2 through 51, but not extending to *COL4A6*, segregated with the disease phenotype. A *COL4A5* deletion encompassing exons 2 through 29 was identified in one of the ATS families. In the second ATS family, a deletion of exon 13 of *COL4A5* through exon 3 of *COL4A6* was detected.

Conclusions

These observations suggest that deletion of the 5’ exons of *COL4A6* and of the common promoter of the *COL4A5* and *COL4A6* genes is not essential for the development of leiomyomatosis in patients with ATS, and that *COL4A5_COL4A6* deletions extending into *COL4A6* exon 3 may not result in ATS-DL.

Key words

Alport syndrome, leiomyomatosis, *COL4A5*, *COL4A6*.

BACKGROUND

X-linked Alport syndrome (MIM#301050; ATS) is a childhood onset progressive glomerulopathy with high frequency sensorineural hearing loss (SNHL) and typical ocular signs.(Kruegel, Rubel et al. 2013) Electron microscopy (EM) examination of kidney biopsies of ATS patients, which remain a crucial diagnostic test for the disease,(Cosgrove 2012) shows characteristic changes of the glomerular basement membrane (GBM), including irregular thinning, thickening and splitting. These abnormalities are thought to be an expression of the higher susceptibility of the GBM to proteolytic degradation,(Khoshnoodi, Pedchenko et al. 2008) in all forms of Alport syndrome (AS). ATS is fully penetrant and consistently severe in hemizygous males, most of whom have reached end-stage renal disease (ESRD) by the age of 30 years, but has lower penetrance and is clinically less severe in heterozygous females.

In rare families, ATS is associated with diffuse leiomyomatosis (MIM#308940; ATS-DL). Diffuse leiomyomatosis (DL) is a benign neoplastic condition characterized by aberrant proliferation of well differentiated smooth muscle cells, involving the gastrointestinal, tracheobronchial and female genital tracts.(Garcia Torres and Guarner 1983) DL of the oesophagus is a constant finding in families with ATS-DL, being the initial clinical manifestation in most patients, frequently mimicking achalasia.(Cochat, Guibaud et al. 1988; Bourque, Spigland et al. 1989) In contrast to sporadic oesophageal leiomyomata, which form circumscribed and often solitary nodules, these patients have diffuse thickening of the esophageal musculature, particularly of the lower third. Histologically, there is extensive replacement of the normal fibre pattern by irregular, plexiform fibres with whorl formation.(Heidet, Boye et al. 1998) The gender-related penetrance and severity of the renal, audiological and ocular manifestations of ATS-DL are similar to those described in ATS. In ATS-DL, however, the esophageal involvement is identically severe in both genders, even in females with no evidence of renal disease, including microscopic hematuria.(Dahan, Heidet et al. 1995)

ATS is caused by mutations in the *COL4A5* gene (MIM*303630),(Barker, Hostikka et al. 1990) the gene encoding the alpha-5 chain of collagen type IV (α5(IV)), which is a major structural component of the mammalian GBM.(Khoshnoodi, Pedchenko et al. 2008) So far, more than 700 different pathogenic mutations confined to *COL4A5* have been described in affected individuals (Human Gene Mutation Database®; http://www.hgmd.cf.ac.uk/ac/gene.php?gene=COL4A5, last accessed on March 9, 2014). In males, large deletions, nonsense and frameshift mutations that lead to truncated proteins and/or mutations located closer to the 5' end of the gene were shown to correlate with a younger age of onset of chronic kidney disease (CKD) and a higher risk of developing SNHL and specific ocular changes, as compared with other types of *COL4A5* mutations. In females, no clear-cut genotype-phenotype correlations have ever been established.(Jais, Knebelmann et al. 2003) *COL4A5* is paired head-to-head with *COL4A6* (MIM*303631), at chromosome location Xq22.3, sharing a bidirectional promoter.(Khoshnoodi, Pedchenko et al. 2008) *COL4A6* encodes the alpha-6 chain of collagen type IV (α6(IV)), which assembles with two α5(IV) chains in long triple helical molecules. These α5(IV)/α5(IV)/α6(IV) heterotrimers are present in the basement membrane (BM) of the skin, smooth muscle, and of the Bowman's capsule and distal tubules in the kidney, but not in the GBM.(Ninomiya, Kagawa et al. 1995; Peissel, Geng et al. 1995) Absence of α6(IV) in the GBM explains why mutations confined to *COL4A6* would not cause ATS.

A contiguous gene deletion involving the 5' exons of *COL4A5* and *COL4A6* was identified in all patients with ATS-DL reported so far.(Antignac, Zhou et al. 1992; Zhou, Mochizuki et al. 1993; Renieri, Bassi et al. 1994; Dahan, Heidet et al. 1995; Heidet, Cohen-Solal et al. 1997; Van Loo, Vanholder et al. 1997; Heidet, Boye et al. 1998; Ueki, Naito et al. 1998; Segal, Peissel et al. 1999; Guillem, Delcambre et al. 2001; Mothes, Heidet et al. 2002; Wang, Ding et al. 2002; Anker, Arnemann et al. 2003; Thielen, Barker et al. 2003; Sugimoto, Yanagida et al. 2005; Oohashi, Naito et al. 2011; Uliana, Marcocci et al. 2011) Deletion of a 4.2 kb critical region containing

exon 1 of *COL4A5*, exons 1', 1 and 2 of *COL4A6* and the common promoter region that regulates the expression of the two adjacent genes, is thought to be required for the development of DL.(Thielen, Barker et al. 2003) With the exception of a single family,(Uliana, Marcocci et al. 2011) deletions extending beyond exon 3 of *COL4A6* were associated with ATS only.(Heidet, Dahan et al. 1995; Heidet, Cohen-Solal et al. 1997) To the best of our knowledge, mutations involving only *COL4A5* or *COL4A6* have not been found in patients with ATS-DL.

Herein, we report the clinical and molecular characterization of three Portuguese families with ATS or ATS-DL bearing deletions of *COL4A5*. Remarkably, the family with ATS-DL had a deletion confined to *COL4A5*, an observation that questions the current genotype-phenotype correlation paradigm in this disorder.

METHODS

Probands and their relatives were enrolled in a multicentric study aiming to describe the phenotype and identify the genotype of families with AS in Portugal. Written informed consent was obtained from all patients. Current as well as retrospective clinical and laboratory data were collected at enrolment, concerning typical renal, audiological and ophthalmologic manifestations of AS. Chronic kidney disease (CKD) was defined according to the National Kidney Foundation clinical practice guidelines (http://www.kidney.org/professionals/KDOQI/guidelines_ckd/p4_class_g1.htm) and the glomerular filtration rate was estimated by a standard equation (http://www.kidney.org/professionals/kdoqi/gfr_calculator.cfm). Clinically relevant gastrointestinal tract symptoms were investigated as appropriate. Upper gastrointestinal workup of the proband of family 1 was detailed elsewhere.(Sousa, Figueiredo et al. 2013)

For the purpose of inclusion in this report, we selected "large" *COL4A5* deletions, defined as those involving at least 50 kb and/or 10 exons. GenBank® reference sequences NG_011977.1 and NG_012059.2 (http://www.ncbi.nlm.nih.gov/nuccore/) were used as the reference nucleotide

sequences of *COL4A5* and *COL4A6*, respectively. In order to facilitate the comparison of our data with previous reports of deletions involving the *COL4A5* and *COL4A6* genes, *COL4A6* exons are numbered according to the original description of the two alternative promoters,(Sugimoto, Oohashi et al. 1994) where exon 1' corresponds to exon 1 of reference sequence NG_012059.2, exon 1 to exon 2, exon 2 to exon 3, and so forth. Similarly, due to the existence of two alternatively spliced exons within intron 41 of the *COL4A5* gene, which have been termed 41A and 41B,(Guo, Van Damme et al. 1993; Martin and Tryggvason 2001) the sequential exon numbers in reference sequence NG_011977.1 differ from the original numbering beyond exon 41. For this reason, the last *COL4A5* exon herein numbered 51 corresponds to exon 53 in the GenBank® sequence.

Genomic DNA was extracted from peripheral blood leukocytes using a commercial kit (Citomed; Lisbon, Portugal). PCR amplification of all exons and adjacent intronic regions of the *COL4A5* gene was performed in affected males, using previously described primers and conditions.(Martin, Heiskari et al. 1998) Exon 1 of the *COL4A5* gene, exons 1', 1, 2, 3 and 4 of the *COL4A6* gene, and their intervening common promoter region, were amplified using the PCR primers and conditions described in supplementary table S1. Amplified regions were sequenced with an ABI Prism® 310 Genetic Analyzer (Life Technologies; Foster City, CA), using the BigDye® Terminator v3.1 Cycle Sequencing kit (Life Technologies). Deletions of the *COL4A5* coding region and deletions of exons 1', 1 and 2 of *COL4A6* were further investigated using the SALSA P191/P192 Alport Multiplex Ligation-dependent Probe Amplification (MLPA) commercial kit (MRC-Holland; Amsterdam, The Netherlands). The MLPA assay was performed according to the manufacturer's protocol, available at the MRC-Holland internet site (http://www.mrc-holland.com). The extension of deletions in family 1 was determined by microarray-based comparative genomic hybridization (array-CGH), using a high-resolution custom-designed X chromosome-specific 244k oligo-array (Agilent; Santa Clara, CA), as described elsewhere.(Froyen, Van Esch et al. 2007) Confirmation and mapping of the deletions was achieved by

standard PCR assay. Iterative rounds of PCR were performed at both the proximal and distal breakpoints to narrow down the breakpoint regions, using the primer pairs indicated in Supplementary table S2. Breakpoint positions are reported according to the most recently updated human genome reference sequence at the UCSC Genome Bioinformatics Site (http://genome.ucsc.edu/) (hg19; Feb 2009). Carrier status analysis was performed by quantitative PCR (qPCR) using the SYBRgreen relative quantitation method on a LC480 apparatus (Roche; Basel, Switzerland), as described previously.(Vandewalle, Van Esch et al. 2009) Primer sequences used for the qPCR assays can be obtained upon request. In order to track its inheritance across the three generations of the pedigree available for study, the risk haplotype of the ATS-DL family was characterized with four microsatellite polymorphic markers flanking the *COL4A5* gene (DXS1120, DXS1105, DXS1210, DXS456).(Srivastava, McMillan et al. 1999; Tazon-Vega, Ars et al. 2007)

RESULTS

Clinical Investigation

The demographic and clinical features of patients are summarised in table 1.

Family 1

Family 1 includes two siblings and their mother affected with ATS-DL (figure 1-A). AS was diagnosed in the proband (III-1) when he was five years old, based on typical EM findings on a kidney biopsy (figure 2) obtained in the work-up of persistent urinary abnormalities. High frequency SNHL was first documented at age six years. The patient started haemodialysis for renal replacement therapy (RRT) at age 20, eventually receiving a kidney transplant six months later. Typical AS retinopathy was diagnosed at age 26.

At age 25, due to persistent upper gastrointestinal symptoms, including long-standing odynophagia, dysphagia to solids and liquids, regurgitation, and retrosternal and epigastric pain, the proband underwent a barium oesophagography, an oesophagogastroduodenoscopy and a stationary manometry, whose results were consistent with the diagnosis of achalasia.

Unsuccessful standard surgical treatment for achalasia with Heller myotomy and Dor fundoplication prompted further evaluation by computer tomography (CT), which showed prominent circumferential thickening of the distal two-thirds of the oesophagus and of the cardia (figure 3). Endoscopic ultrasound imaging confirmed the marked increase in the thickness of the esophageal wall, most prominently of the second endosonographic layer, suggesting the involvement of the muscularis mucosa, and endoscopic tunnel biopsies were diagnostic of leiomyomatosis (figures 4-A,B). As the patient suffered from bronchial asthma since early childhood, a positron emission tomography (PET)–CT was performed to screen for pulmonary involvement by DL. It showed abnormal uptake of the radiotracer by the distal oesophagus, right colon and rectum. A colonoscopy to exclude colon cancer was normal and a rectal biopsy showed abundant smooth muscle cells but the tissue sample was too superficial to allow the histological diagnosis of leiomyomatosis. In the patient's clinical setting, these results were interpreted as suggestive of colonic and rectal involvement by DL.

The proband's mother (II-2) has bilateral high frequency SNHL and microscopic hematuria. She reported to have had surgeries for "achalasia" at ages 18 and 22. At age 30, hysterectomy, right adnexectomy and left salpingectomy were performed due to diagnostic hypothesis of uterine fibromyoma and right adnexal mass. At age 36, she underwent total colectomy with proctectomy for the treatment of severe long-standing constipation, which was attributed to aganglionic megacolon. Gross examination showed an irregularly dilated colon, reaching a maximum diameter of 15cm, and a strikingly dilated rectum, reaching a maximum diameter of 8 cm, with a markedly thickened muscular layer. Histologically, however, the diagnostic hypothesis of aganglionic megacolon was not confirmed. After the diagnosis of ATS-DL in the son, microscopic examination of the same fragments was performed, which showed, not only the presence of ganglion cells in the submucosa and muscle layer, but also the proliferation of well-differentiated smooth muscle cells in the *muscularis mucosa* and *muscularis propria* of the colon and rectum, confirming the diagnosis of DL

(figure 5). A barium oesophagography obtained at the age of 44, for the investigation of recurrent complaints of dysphagia for solids and liquids, showed significantly delayed progression of the barium swallows along the distal two thirds of the oesophagus, especially of the liquid boluses, resulting in their accumulation in mid-oesophagus. Endoscopic ultrasound imaging additionally revealed thickening of the distal oesophagus, particularly of the second layer. In view of the patient's clinical and family history, these findings were considered to be highly suggestive of esophageal leiomyomatosis, and further invasive diagnostic procedures were avoided.

The proband's sister (III-3) presented with recurrent macroscopic hematuria and persistent urinary abnormalities in the first year of life, and eventually reached CKD stage 3 at age 19 years. She has a long-standing history of severe constipation, with a normal colonoscopy result before the diagnosis of ATS-DL was suspected. At age 20, due to sporadic symptoms of dysphagia, she underwent a barium esophagography which was consistent with esophageal leiomyomatosis. In addition, she has microcephaly, moderate intellectual disability and dysmorphic facial features, which have been attributed to a *de novo* 13q34 terminal deletion syndrome.

The proband's maternal grandmother (I-2) reported history of sporadic microscopic hematuria, hypertension, SNHL and bilateral cataracts. The proband's brother (III-2), maternal aunt (II-3) and maternal uncle (II-4) are asymptomatic and have no evidence of CKD.

Family 2

In family 2, a 34 years old male and his mother were diagnosed with ATS (figure 1-B). Urinary abnormalities and CKD were detected in the proband (III-1) at age 15 years, on the diagnostic investigation of progressive SNHL, which had been first noticed when he was eight years old. AS was clinically suspected but a kidney biopsy was not conclusive. Bilateral cataracts were detected at age 25. The patient reached ESRD at age 30. The proband's mother (II-2) has CKD stage 3. No other cases of ESRD are known in this family.

Family 3

Family 3 includes mother and son with ATS (figure 1-C). The proband (II-2) is a female who started RRT at age 46. Urinary abnormalities had been present since early childhood. Pure-tone audiometry screening, performed at enrolment for the present study, revealed moderate SNHL on the right ear. Her 18-year-old son (III-1) has persistent urinary abnormalities and bilateral moderate SNHL since early childhood, and a right cataract was diagnosed at age 5 years. Neither the proband nor her son reported symptoms that might be suggestive of DL.

Table 1. Demographic, clinical and genetic characterization of three Alport syndrome families with *COL4A5* deletions.

Family	1				2		3	
Patient	Proband (III-1)	Sister (III-3)	Mother (II-2)	Grandmother (I-2)	Proband (III-1)	Mother (II-2)	Proband (II-2)	Son (III-1)
Genotype								
Mutation	del ex.2_51 [*COL4A5*]				del ex.2_29 [*COL4A5*]		del ex.1_13 [*COL4A5*] / / del ex.1_3 [*COL4A6*]	
Mutation status	hemizygous	heterozygous	heterozygous	no deletion	hemizygous	heterozygous	heterozygous	hemizygous
Clinical history and phenotype								
Gender, age at enrollment in years	M, 26y	F, 20y	F, 45y	F, 69y	M, 34y	F, 60y	F, 47y	M, 18y
History of macroscopic hematuria[a]	Yes [2]	Yes [<1]	No	NA	Yes [15]	NA	No	No
History of microscopic hematuria[a]	Yes [2]	Yes [<1]	Yes [43]	Yes [66]	Yes [15]	Yes [60]	Yes [3]	Yes [5]
History of proteinuria[a]	Yes [7]	Yes [<1]	No	No	Yes [15]	Yes [NA]	Yes [3]	Yes [5]
Degree of proteinuria (mg/day) at diagnosis / most recent[a]	256 [7] / -	- / 250 [19][+]	-	-	NA	- / 340 [60]	NA	NA
History of hypertension[a]	Yes [NA]	No	No	Yes [NA]	Yes [20]	Yes [50]	Yes [35]	No
History of CKD stage 3 or higher[a]	Yes [NA]	Yes [19]	No	No	Yes [15]	Yes [60]	Yes [38]	NA
pCr (eGFR) at diagnosis of CKD stage 3	NA	124 (48)	-	-	NA	97 (51)	106 (50)	NA
History of renal replacement therapy[b]	HD [20], Tx [21]	No	No	No	HD [30], Tx [33]	No	HD [46]	No
Last available pCr (eGFR)[†,a]	97 (81) [26]	124 (48) [19]	62 (91) [44]	71 (77) [66]	NA	97 (51) [60]	-	44 (217) [11]
Kidney biopsy[c]	Yes [5]*	No	No	No	Yes [NA]	No	No	Yes [NA]
Sensorineural hearing loss[‡,a]	Yes [6]	Yes [16]	Yes [40]	Yes [67]	Yes [8]	No	Yes [46]	Yes [6]
Ocular signs[a]	Yes [26]	Yes [20]	Yes [45]	Yes [NA]	Yes [25]	NA	NA	Yes [5]
Diffuse leiomyomatosis[a]	Yes [25] **	Yes [19] ~	Yes [<18] ***	No	No	No	No	No

NA: Data not available. [a] In brackets, the age at diagnosis or at last screening, in years. [b] In brackets, the age at start of haemodialysis (HD) or of kidney transplant (Tx). [c] In brackets, the age at kidney biopsy. < Before the age indicated within brackets. [†]For patients not on HD. [‡] SNHL confirmed by audiogram in all affected patients. [+]Under treatment with an angiotensin-converting-enzyme inhibitor. * GBM changes shown in figure 2. pCr: plasma

creatinine level, in micromol/L (to convert into mg/dL divide by 88.4). eGFR: estimated glomerular filtration rate, in mL/min/1.73m^2. ** DL diagnosis histologically confirmed (see figure 4); *** DL diagnosis histologically confirmed (see figure 5); ~ DL diagnosis clinically suspected.

Molecular Investigation

In the proband (III-1) of family 1, a *COL4A5* hemizygous deletion spanning exons 2 to 51 was detected by PCR and confirmed by MLPA analysis. No point mutations were detected in exon 1 of *COL4A5*, in exons 1', 1, 2, 3 and 4 of *COL4A6*, nor in the shared promoter region. High-resolution X-chromosome array-CGH showed that there were no additional genes included in the deletion (Supplementary figure S1). For junction analysis, breakpoint mapping was first performed by iterative rounds of PCR. When breakpoint regions were <3 kb, a PCR with forward primer set 2 combined with reverse primer set 7 (Supplementary table S2) was done yielding a 4 kb PCR product specific for the patient (Supplementary figure S2-A). A nested PCR was then performed using the complementary sequences from reverse primer set 2 and forward primer set 7 to yield a band of approximately 3 kb, which was subsequently sequenced (Supplementary figure S2-B). The *COL4A5* deletion was mapped precisely to chrX:107,773,340-107,950,610, with the proximal breakpoint located in intron 1, 9.5 kb proximal to exon 2, at the end of a MA3-LINE1 repeat, and the distal breakpoint located 10 kb distal to the last exon 51, within a MB5-LINE1 repeat. As both breakpoints locate in different subfamilies of LINE1 repeats (MA3 and MB5) with no significant homology, this junction did not likely result from non-allelic homologous recombination (NAHR) (Supplementary figure S2-C). Mapping the breakpoint to the nucleotide level however, cannot exclude the occurrence of additional, but subtle rearrangements in *COL4A5* or *COL4A6*. Regular PCR and high-resolution X-oligo-array demonstrated that at least all other exons flanking this deletion are neither deleted nor duplicated in the proband of family 1. Moreover, qPCR at five different loci within *COL4A6* intron 2 and one in exon 3 all revealed the expected copy number of 1.00 showing that intron 2 does not contain any copy number variations larger than 25 kb. The MLPA and qPCR assays showed that the proband's mother (II-2) and sister (III-3) were heterozygous for the same deletion but that it was not present in the unaffected proband's brother (III-2), maternal aunt (II-3) and maternal uncle (II-4). Of note, this deletion was also not detected by MLPA in peripheral

blood genomic DNA of the proband's maternal grandmother (I-2), who has several clinical manifestations that might fit in the diagnosis of AS. Since haplotype analysis showed that she transmitted the same haplotype both to her affected daughter (II-2) and unaffected son (II-4) (Supplementary figure S3), it might be speculated that instead the *COL4A5* rearrangement having occurred *de novo* in her daughter, this woman is a mosaic carrier of the *COL4A5* deletion in her gonads and other ectodermal tissues affected in AS.

In family 2, a deletion of exons 2 to 29 of the *COL4A5* gene was detected by PCR and confirmed by MLPA in the proband (II-1). The proximal breakpoint was mapped in intron 1 in a 700 bp region 24.5 kb distal to *COL4A5* exon 1 (chrX:107,707,634-107,708,329), which contains a ME1-LINE1 repeat. The distal breakpoint was mapped within a region of 2.9 kb in intron 29 (chrX:107,856,319-107,858,111) also including a LINE1 of the ME1 subfamily. Although we were unable to obtain the junction sequence, the most likely mechanism for this rearrangement is NAHR between LINE1 repeats. His mother (II-2) carries the same *COL4A5* deletion, as shown by MLPA.

In family 3, a deletion of exons 1 to 13 of *COL4A5* and of exons 1', 1 and 2 of *COL4A6*, was identified by MLPA in the female proband (II-2). The deletion also includes exon 3 of *COL4A6*, as confirmed by PCR in her affected son (III-1). Since *COL4A6* exon 4 was retained, the proximal breakpoint of this deletion locates in intron 3 of *COL4A6*. As demonstrated by PCR in the son, the shared promoter region of both genes was also deleted.

DISCUSSION

We report the detailed molecular characterization of three deletions involving the *COL4A5* gene identified in Portuguese patients presenting with ATS or ATS-DL. The ATS families carried either a 149 kb deletion encompassing exons 2 to 29, or a contiguous gene deletion that included exons 1 to 13 of *COL4A5* and exons 1' to 3 of *COL4A6*. According to current paradigms of genotype-phenotype correlation in ATS, these two genetic defects are fully consistent with the patients' clinical phenotypes. Surprisingly, however, in the family manifesting ATS-DL, molecular analyses revealed a

177 kb deletion involving exons 2 to 51 of *COL4A5*, which would not be expected to cause DL. Direct sequencing of genomic DNA excluded the presence of single nucleotide or of any other small variants in exon 1 of *COL4A5*, in the intergenic common promoter region and in exons 1', 1, 2, 3 and 4 of *COL4A6*. Although we did not perform pulsed field gel electrophoresis (PFGE) for long range mapping and detection of more complex rearrangements, a technique formerly used by other investigators to characterize large *COL4A5* (Vetrie, Boye et al. 1992) or *COL4A5_COL4A6* deletions (Heidet, Dahan et al. 1995; Heidet, Cohen-Solal et al. 1997; Guillem, Delcambre et al. 2001; Mothes, Heidet et al. 2002), cloning the junction clearly demonstrated that the deletion in this family was a simple rearrangement, with no inversion or insertion at the deletion breakpoints. It was also demonstrated by high-resolution array-CGH that exons and genes flanking the *COL4A5* deletion were neither deleted nor duplicated. Although qPCR at several locations within *COL4A6* intron 2 showed the expected copy number at all the tested loci, it does not formally exclude the presence of a second, smaller rearrangement between the qPCR probed regions. Such a rearrangement might be of pathogenic importance as *COL4A6* intron 2 is thought to constitute the proximal boundary of the critical region for the development of DL in patients with ATS.(Thielen, Barker et al. 2003)

Until now, approximately 30 families with ATS-DL were found to carry deletions comprising the 5' regions of *COL4A5* and *COL4A6*, all of them including the 4.2 kb critical region (figure 6). (Antignac, Zhou et al. 1992; Zhou, Mochizuki et al. 1993; Renieri, Bassi et al. 1994; Dahan, Heidet et al. 1995; Heidet, Dahan et al. 1995; Heidet, Cohen-Solal et al. 1997; Heidet, Boye et al. 1998; Ueki, Naito et al. 1998; Segal, Peissel et al. 1999; Guillem, Delcambre et al. 2001; Mothes, Heidet et al. 2002; Wang, Ding et al. 2002; Anker, Arnemann et al. 2003; Thielen, Barker et al. 2003; Sugimoto, Yanagida et al. 2005; Oohashi, Naito et al. 2011; Uliana, Marcocci et al. 2011) In contrast, contiguous *COL4A5_COL4A6* deletions extending upstream beyond exon 3 of *COL4A6*,(Heidet, Dahan et al. 1995; Heidet, Cohen-Solal et al. 1997) or encompassing the entire *COL4A5* and *COL4A6* genes,(Thielen,

Barker et al. 2003) were identified in patients with ATS who did not develop DL. Furthermore, intragenic *COL4A5* deletions of at least one exon have been described in more than 30 ATS families but none of these patients developed DL, including two cases who had deletions that also involved exons 2 to 51 of *COL4A5*.(Plant, Green et al. 1999) These observations led to the postulate that the contiguous deletion of the 5' exons of *COL4A5* and *COL4A6*, with the proximal breakpoint located within the intron 2 of *COL4A6*, was the critical genetic defect underlying ATS-DL.(Zhou, Mochizuki et al. 1993; Antignac, Knebelmann et al. 1994; Heidet, Dahan et al. 1995; Segal, Peissel et al. 1999; Zheng, Harvey et al. 1999; Thielen, Barker et al. 2003).

Contiguous *COL4A5_COL4A6* deletions remove the intergenic bidirectional promoters that drive the transcription of the two genes,(Sugimoto, Oohashi et al. 1994; Sund, Maeshima et al. 2005) thereby preventing their normal transcription and expression. Both the α5(IV) and α6(IV) collagen chains are absent in the BM of esophageal leiomyomata removed from patients with ATS-DL.(Heidet, Cai et al. 1997) Lack of expression of α5(IV) and α6(IV) in the tumour BM has also been demonstrated in sporadic esophageal leiomyomata, in association with a somatic deletion of the 5' ends of *COL4A5* and *COL4A6*, mimicking the germline mutation in ATS-DL.(Heidet, Boye et al. 1998) These findings suggest that the mechanisms leading to smooth muscle cell proliferation are similar in the hereditary and sporadic leiomyomata, strengthening the case for the pathogenic role of the specific contiguous gene deletion in ATS-DL.(Heidet, Boye et al. 1998) Unfortunately, the expression status of α6(IV) collagen chains in the normal esophageal smooth muscle basement membranes of ATS patients who do not have deletions extending into *COL4A6*, is not known.(Miner 1999) However, absence of α6(IV) chains from other α6(IV)-containing BM, including of the dermal-epidermal junction and/or of the kidney Bowman's capsule and distal tubules, has been observed in patients with ATS-DL (Segal, Peissel et al. 1999) or with ATS.(Ninomiya, Kagawa et al. 1995; Peissel, Geng et al. 1995; Hino, Takemura et al. 1996; Naito, Kawai et al. 1996; Sasaki, Zhou et al. 1998) This suggests that the

α6(IV) chain cannot assemble into BM without the α5(IV) chain,(Miner 1999) an assumption supported by experimental data in the canine model of X-linked AS.(Zheng, Harvey et al. 1999) These dogs carry a nonsense point mutation in *COL4A5* that is sufficient to prevent incorporation of the collagen α6(IV) chain into smooth muscle BM, independent of a reduction in α6(IV) mRNA levels. However, the lack of the α6(IV) chain does not result in any obvious complications in the affected animals, particularly the development of leiomyomata. This implies that the absence of the α6(IV) chain, alone or in combination with absence of the α5(IV) chain, is not enough to cause DL, and that some other mechanisms, other than expression of α6(IV) and its incorporation into BM, are pathogenically relevant in this disease.(Zheng, Harvey et al. 1999)

Several hypotheses have been proposed to explain the apparent requirement for the proximal breakpoint of deletions associated with ATS-DL to be located within *COL4A6* intron 2: (*i*) gain of function of *COL4A6* conferred by deletions at that position, in a manner abrogated by more extensive deletions,(Heidet, Cohen-Solal et al. 1997) due to activation of a cryptic promoter, resulting in the transcription of sequences that are normally not expressed,(Heidet, Dahan et al. 1995) or formation of pathogenic fusion transcripts;(Thielen, Barker et al. 2003) (*ii*) loss of an unspecified gene or of its regulatory elements within intron 2 of *COL4A6* that might be involved in the regulation of smooth muscle cell proliferation;(Heidet, Dahan et al. 1995; Ueki, Naito et al. 1998) and finally (*iii*) alteration of the chromatin structure, influencing the expression of neighboring genes by an epigenetic mechanism.(Thielen, Barker et al. 2003) The detection of a *COL4A6* transcript that included exon 4, but not exon 3, in an esophageal tumour sample of a patient with ATS-DL,(Heidet, Dahan et al. 1995) lends support to the hypothesis that an abnormal *COL4A6* mRNA transcript or truncated α6(IV) collagen chain might be involved in the pathogenesis of DL. The epigenetic model assumes that the critical ATS-DL region may contain a silencer element(s) for establishment of repressed chromatin, or act as a transcriptional insulator of genes involved in the pathogenesis of DL.(Thielen,

Barker et al. 2003) However, none of these hypotheses is entirely satisfactory (Uliana, Marcocci et al. 2011) and the molecular mechanisms of the smooth muscle overgrowth in ATS-DL patients remains to be elucidated.

Since the α5(IV) collagen chain has binding sites for numerous components of the extracellular matrix (ECM) and interacts with multiple integrin and non-integrin cell receptors,(Khoshnoodi, Pedchenko et al. 2008) it is also possible that the absence or the presence of abnormal α5(IV) chains in the BM induces changes in the structure or function of other ECM component(s) and/or of cell surface receptors. As a result, degradation of the physical support for ECM components provided by the BM type IV collagen network, and the deregulation of the tissue-specific ECM-cell interactions that coordinate intracellular signalling essential for cell growth and proliferation, might favour the development of the leiomyomata.(Zhou, Mochizuki et al. 1993; Heidet, Cai et al. 1997; Heidet, Boye et al. 1998; Thorner, Heidet et al. 1998) Further studies are needed to test these assumptions.

The recent report of a deletion extending proximally into the intron 3 of *COL4A6* in a family with ATS-DL,(Uliana, Marcocci et al. 2011) although leading to reconsider the location of the centromeric limit of the critical region for DL, would still be compatible with a putative pathogenic role for the truncated *COL4A6* mRNA lacking the first three exons. Indeed, it might be speculated that a truncated *COL4A6* transcript causes the development of leiomyomatosis by a dominant activating effect,(Heidet, Boye et al. 1998) while the complete absence of a *COL4A6* transcript does not. This hypothesis implies the presence of a cryptic promoter within the ~90 kb large intron 3 of *COL4A6*. Therefore, mapping the breakpoint sites in intron 3 in those few reported cases with contradictory phenotypic outcomes (Heidet, Dahan et al. 1995; Heidet, Cohen-Solal et al. 1997; Wang, Ding et al. 2002; Uliana, Marcocci et al. 2011) might provide additional proof for the existence of such a cryptic promoter.

Our finding of a deletion confined to *COL4A5* segregating with ATS-DL further challenges the current paradigm of genotype-phenotype correlation in this disorder, showing that contiguous deletion of the *COL4A5_COL4A6*

common promoter region and the 5' exons of *COL4A6* is not an absolute requisite for the development of DL in ATS patients. The fact that DL developed in family 1, carrying a *COL4A5* deletion with proximal breakpoint very close to exon 2, while it did not develop in family 2, carrying a *COL4A5* deletion with proximal breakpoint much closer to exon 1, might indicate the presence of a regulatory domain within the large intron 1 that could affect the expression of a transcript involved in DL. Mapping the breakpoints of other similar *COL4A5* deletions, like those reported by Plant *and colleagues* (Plant, Green et al. 1999) will be most important to test for this hypothesis. On the other hand, failure to identify more ATS-DL patients with deletions confined to *COL4A5* may be due to the rarity of this disease and the difficulty in recognizing its clinical manifestations, in the context of more common differential diagnoses.

CONCLUSIONS

This is the first report on a family with ATS-DL carrying a deletion confined to *COL4A5*, suggesting that the deletion of the 5' exons of *COL4A6* and of the common *COL4A5_COL4A6* promoter region is not essential for the development of DL in patients with ATS. Deletion breakpoint mapping might prove essential to pinpoint the underlying molecular mechanism resulting in ATS-DL.

Given the implications of the recognition of genotype-phenotype correlations in ATS-DL for diagnosis and genetic counselling, the following practical points are worth emphasising: (*i*) clinicians should be aware of the full spectrum of ATS-DL manifestations, as its diagnosis requires a high suspicion index; (*ii*) leiomyomatosis is often severely symptomatic, both in males and females with ATS-DL, before patients develop advanced CKD; *(iii)* besides the esophageal involvement mimicking achalasia, ATS-DL may present with megacolon secondary to colon and anorectal involvement (Guillem, Delcambre et al. 2001; Claus, Geboes et al. 2008); (*iv*) the diagnosis of ATS-DL should be considered in any female with gastrointestinal and gynaecological manifestations suggestive of DL, even without any clinical

or laboratory evidence of nephropathy, due to the risk of ATS-DL in male offspring; and finally *(v)* the inclusion of primers for the exons 3 and 4 of *COL4A6* in the MLPA commercial kits for ATS is justified for a comprehensive molecular diagnosis.

Acknowledgements

The authors thank all patients and their families for their invaluable contribution to this study. This work is part of the PhD project of Maria João Nabais Sá (supervisor: João Paulo Oliveira), to be presented at the ICBAS – "*Instituto de Ciências Biomédicas Abel Salazar*" (Abel Salazar Biomedical Sciences Institute), University of Porto, Portugal. Nathalie Fieremans is a PhD fellow from the IWT (Agentschap voor Innovatie door Wetenschap en Technologie), Belgium. These studies were partially supported by FEDER funds through the COMPETE-"Programa Operacional Factores de Competitividade", and by Portuguese national funds through the FCT – "*Fundação para a Ciência e a Tecnologia*" (Foundation for Science and Technology) within the project FCOMP-01-0124-FEDER-HMSP-ICJ/SAU-ICT/0003/2009; by a research grant from the Portuguese Society of Nephrology; by the Unit of Research & Development in Nephrology (FCT-725) at the Department of Nephrology of the Faculty of Medicine, University of Porto, Portugal; and by the São João Hospital Centre. The authors thank NephroCare-Portugal for the thorough ascertainent of patients with Alport syndrome in their dialysis centers. The authors thank Laboratório Roriz, Lda (Lisboa, Portugal) for providing archival paraffin blocks containing surgical samples of the colon and rectum of the proband's mother of the ATS-DL family. The authors thank the Department of Pathology of Hospital de Santo António/Centro Hospitalar do Porto (Porto, Portugal), namely Anabela Ferreira for providing the time and expertise required to prepare the biological material for histological review and Dr. Francisca Costa for providing the time and expertise for reviewing and obtaining photographs of the colectomy and proctectomy fragments for this manuscript.

Parts of these data were presented as posters at the European Human Genetics Conference 2012, Nürnberg, Germany, June 23-26, 2012 (P12.007) and at the 62nd Annual Meeting of the American Society of Human Genetics, November 6-10, 2012 San Francisco, California (3122F).

Figures

Figure 1

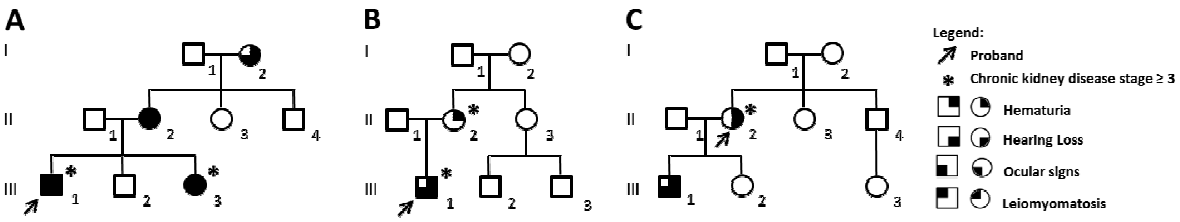


Figure 1. Pedigrees of family 1 (A), family 2 (B), and family 3 (C).

Figure 2

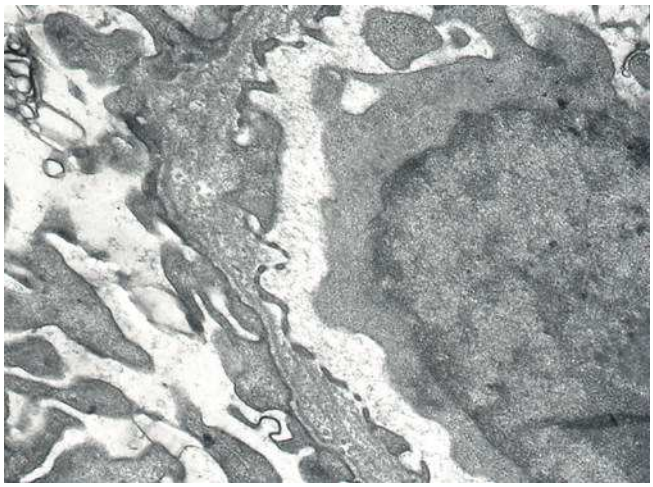


Figure 2. Electron microscopy of the kidney biopsy of the proband of family 1 performed at 5 years, showing the typical glomerular basement membrane lesions of Alport syndrome: alternating thinning and thickening, splitting, and electron-dense bodies surrounded by a lucent halo.

Figure 3

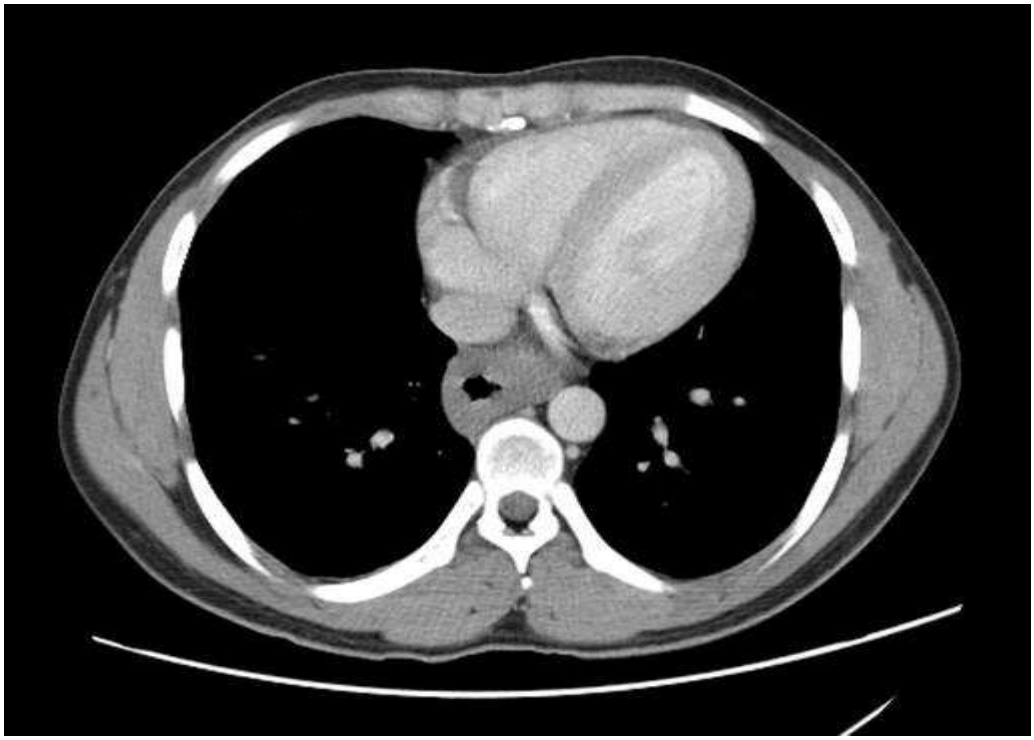


Figure 3. Axial section of the lower half of the thorax from the proband of family 1 (III-1), obtained during arterial and venous phases of a computer tomography scan, showing “isodense to muscle”, homogeneous and circumscribed wall thickening of the distal segment of the oesophagus with luminal gas bubble relatively centred and symmetrical.

Figure 4

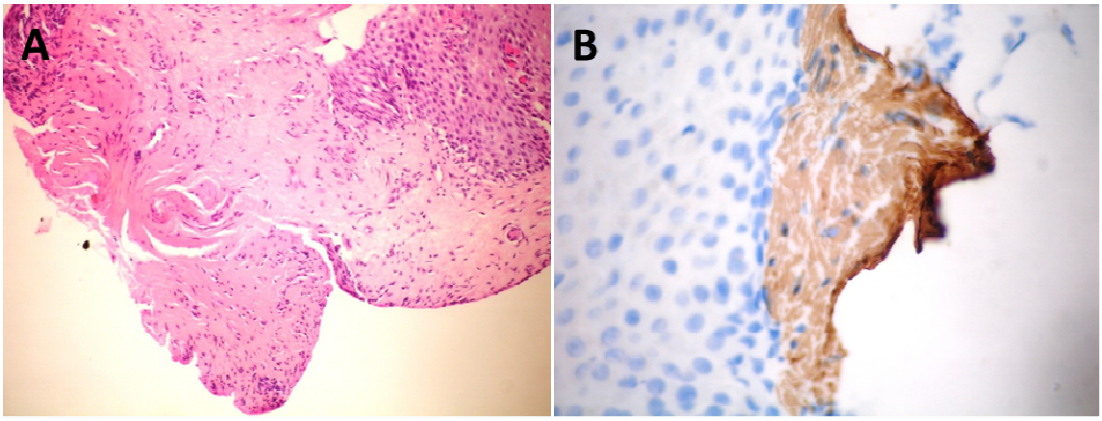


Figure 4. A – Light microscopy of the esophageal biopsy of the proband of family 1 performed at 25 years, showing muscle cell proliferation arranged in intersecting bundles (H&E (hematoxylin and eosin); x200). B – Immunohistochemistry of the esophageal biopsy of the proband of family 1 showing desmin-positive, DOG1 (“Discovered on GIST-1”)-negative spindle cells (desmin; x300).

Figure 5

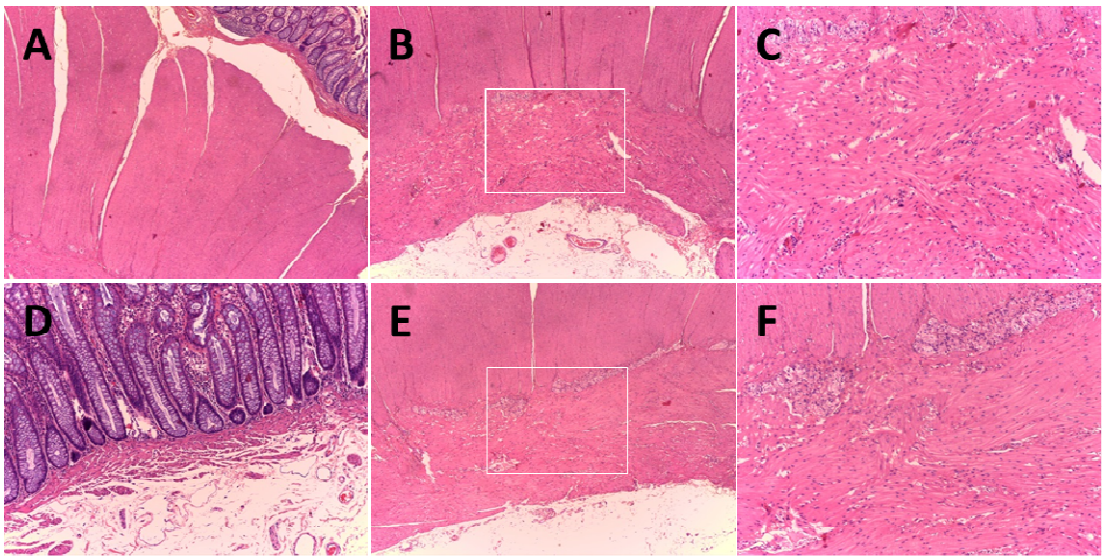


Figure 5. Light microscopy of sections from the colon (A,B,C) and rectum (D,E,F) of the proband's mother of family 1, obtained after total colectomy with protectomy, performed at 36 years. A – Proliferation of smooth muscle cells in the *muscularis mucosa* and markedly thickened circular muscle layer of the colon (H&E (hematoxylin and eosin); x40). B –Poorly delineated smooth muscle proliferation of the muscular longitudinal layer and ganglion cells from the myenteric plexus of the colon [detail from the white box is magnified in C] (H&E; x40). C – Well-differentiated smooth muscle cells, which lost their usual parallel orientation (H&E; x100). D – Thickened *muscularis mucosa* under vertically oriented mucosal crypts of the rectum (H&E; x100). E – Smooth muscle cell proliferation of the *muscularis propria* and presence of ganglion cells from the myenteric plexus of the rectum, excluding Hirschsprung disease [detail from the white box is magnified in F] (H&E; x40). F – Interlacing bundles of smooth muscle cells infiltrating the myenteric plexus (H&E; x100).

Figure 6

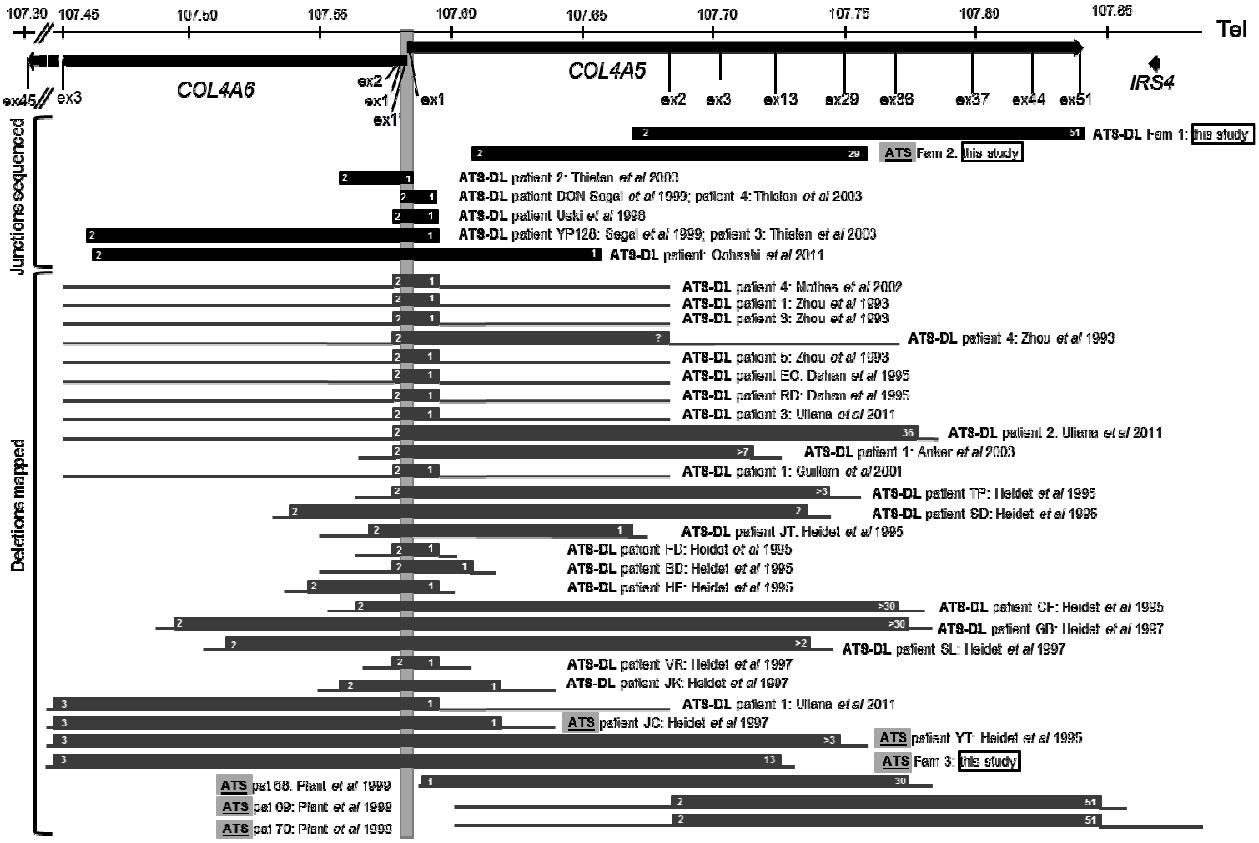


Figure 6. Schematic overview of the most currently reported deletions within *COL4A5* in ATS or ATS-DL patients. The 400 kb region (107.45 Mb to 107.85 Mb; UCSC Hg19) at Xq22.3 containing the complete *COL4A5* (exon 1 to 51) and part of the *COL4A6* (exon 1 to 3) genomic organization is shown above. The seven deletions of which the junctions were mapped and sequenced are shown first followed by 29 deletions for which only the breakpoints were mapped. Horizontal grey bars indicate regions that were demonstrated as deleted while the grey lines represent the regions in which the breakpoints must be located. The first or last exon that is deleted is given by its number, which is placed at the beginning or end of each grey bar, respectively. For each deletion, it is mentioned whether it concerns ATS patients (underlined, highlighted in grey) or ATS-DL patients (no highlight), followed by the reference in which the patient(s) was described. The deletions identified in our three families are highlighted in a box labelled 'this study'.

Supplementary Material

Supplementary table S1. Primer sequences, annealing temperatures and product sizes for PCR amplification of the common promoter region of the *COL4A5* and *COL4A6* genes, the exon 1 of *COL4A5* and exons 1', 1 and 2 of *COL4A*6.

	Forward primer	Reverse primer	Annealing temperature	PCR product size	References
***COL4A5* exon 1** †	5'-AAGCCTCACTGTCCCTCTC-3'	5'-AAAGGAAGATAAAGGGACCC-3'	60°C	296 bp	Adapted from (Martin, Heiskari et al. 1998)
***COL4A5-COL4A6* intergenic region** †	5'-AAATTCCCGGCTGGCTCTA-3'	5'-CAGGGAGCACCGACCAAAA-3'	58°C	621 bp	(Sugimoto, Oohashi et al. 1994; Mothes, Heidet et al. 2002)
***COL4A6* exon 1'** †	5'-CCAGACTAGTTGACTGAGC-3'	5'-CCCTAAGTATCTCCACAGC-3'	58°C	1150 bp	Primers designed for this study
***COL4A6* exon 1 and 2** †	5'-GGCAGCTGAATCGATATCTC-3'	5'-CTCGTGGTGAAACTCTCTGC-3'	60°C	1218 bp	Primers designed for this study
***COL4A6* exon 3** ††	5'-TGAGAAAGGAGCGAGAGT-3'	5'-GGGAGGGAAGAATCAAGT-3'	58°C	360 bp	Primers designed for this study
***COL4A6* intron 3** ††	5'-GGAAAAGTGTTGAAAGGG-3'	5'-GATAATGAAAGAGAAGGTGG-3'	58°C	437 bp	Primers designed for this study
***COL4A6* exon 4** ††	5'- CACTCCTTCTGTCCACCT -3'	5'-CCTGGCACCAAACTACTC-3'	58°C	274 bp	Primers designed for this study

† Cycling conditions were: initial denaturation for 5 minutes at 94°C, followed by 35 cycles of denaturation at 94°C for 60 seconds, annealing for 60 seconds, and extension at 72°C for 90 seconds, ending with a final extension at 72°C for 10 minutes.

†† Cycling conditions were: initial denaturation for 5 minutes at 94°C, followed by 35 cycles of denaturation at 94°C for 30 seconds, annealing for 30 seconds, and extension at 72°C for 40 seconds, ending with a final extension at 72°C for 10 minutes.

Supplementary table S2. PCR primers to map the breakpoint regions in the proband of family 1.

N°	Amplicon location [a]	Forward primer (5' to 3')	Reverse primer (5' to 3')	Size (bp)	Present [b]
Mapping the proximal breakpoint in COL4A5 intron 1					
1	107,761,475 - 107,761,850	GTCCCATTTCACTATTTCC	TCCCATATGTTTTTTCCCC	376	yes
2	107,770,364 - 107,770,677	CCCATGTTGTGTCGATTTA	AAGTGAGATAGGGTGAAAG	314	yes
3	107,773,302 - 107,773,649	AAACATCCTACCCAACAAC	TTGCCTGTTCACTCTGAT	348	no
4	107,775,848 - 107,776,325	GGGAAAAGAAGCAGTCAAA	TTAGGTCAGGGTTGCATT	478	no
Mapping the distal breakpoint downstream of COL4A5					
5	107,946,359 - 107,946,617	TAGAGAGGCAGGAAAGGT	TGAGTAGGAGAAGGAAGTGT	259	no
6	107,948,701 - 107,949,182	TGGGGTGTTGATTTGTTG	TTGTGCACTTGGATTAGG	482	no
7	107,951,237 - 107,951,625	CTGTCCTGGGCTTTTATTT	CCCATCGCCACAAAAAAA	389	yes
8	107,953,812 - 107,954,214	GAAGAATAGATGCACTGGAA	GGGAGCGGAAAATGGAGA	403	yes

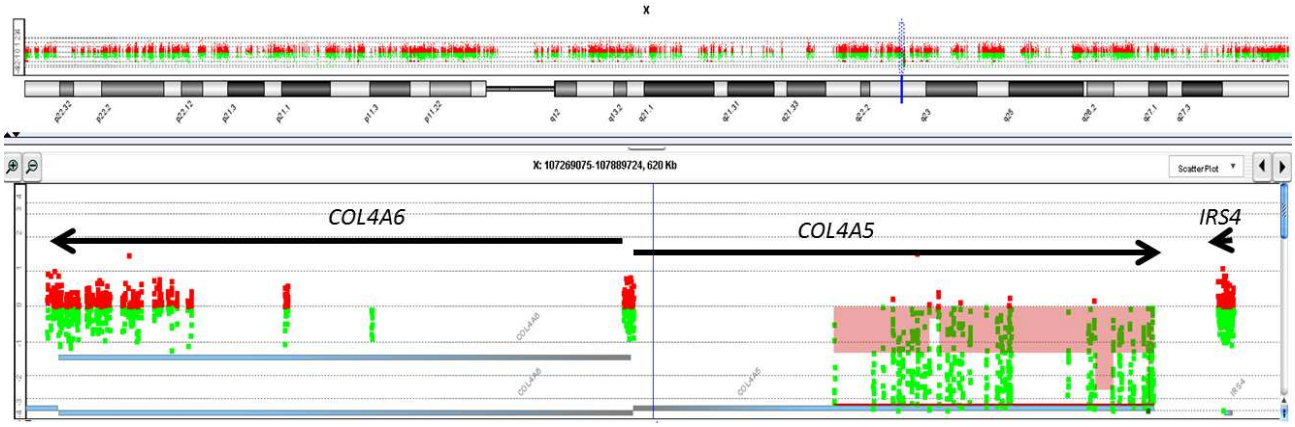
[a] Positions based on UCSC Hg 19.

[b] The column 'Present' indicates whether a PCR fragment was obtained (yes) or not (no).

The proximal breakpoint is located in a 2.7 kb region between positions 107,770,677 and 107,773,302.

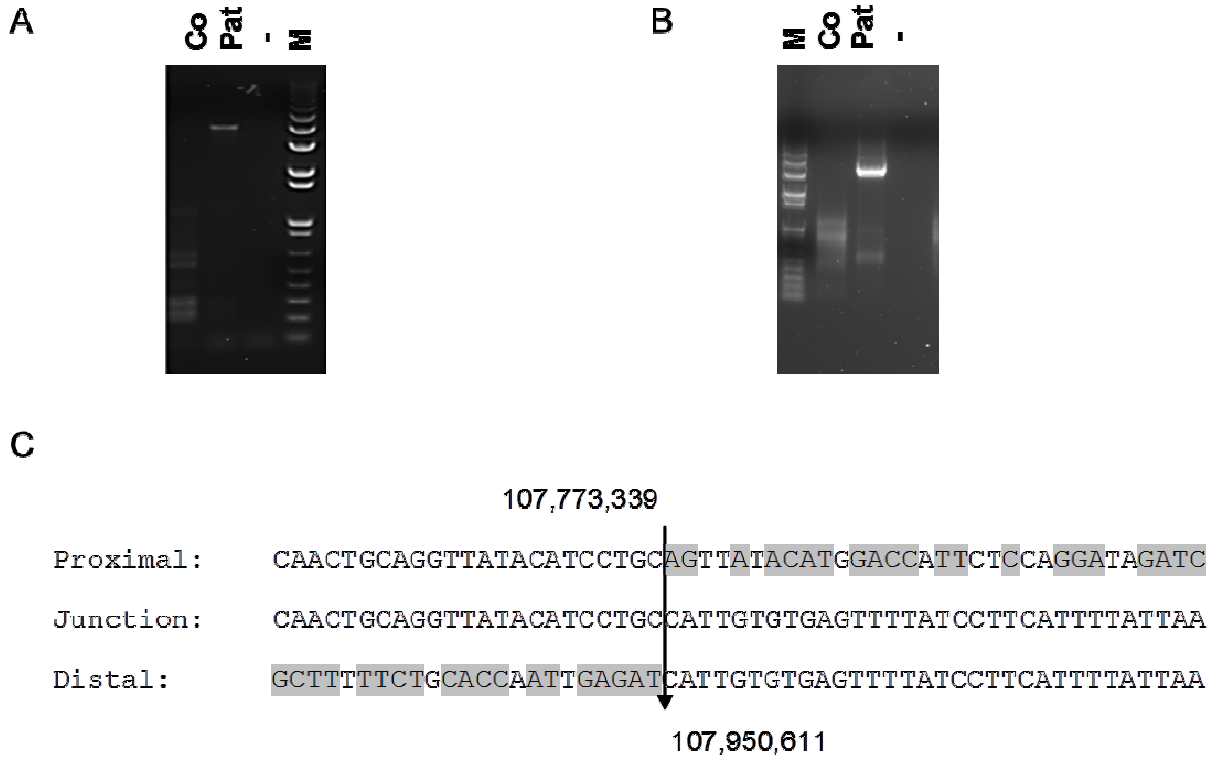
The distal breakpoint is located in a 2.0 kb region between positions 107,949,182 and 107,951,237.

Supplementary figure S1



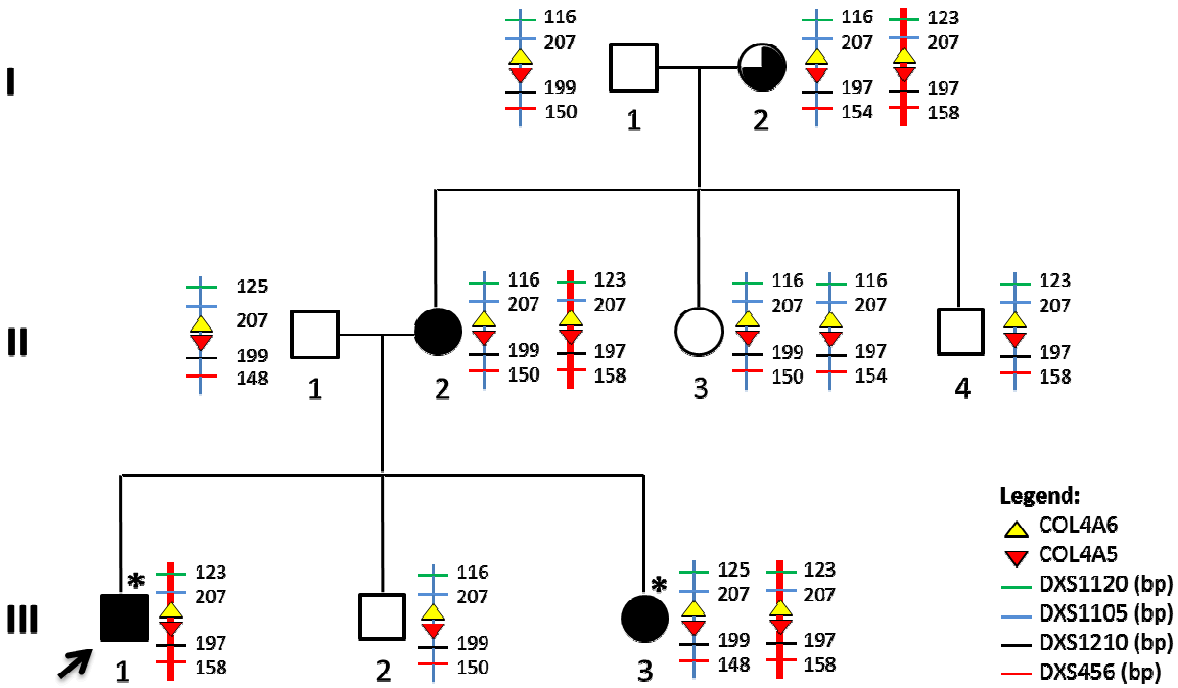
Supplementary figure S1. X chromosome-specific exome array-CGH data of the *COL4A6-COL4A5* region at Xq22 for the proband III-1 of family 1, diagnosed with ATS-DL. The deletion (shaded area) includes exon 2 to exon 51 of *COL4A5*. Locations of probes are: last normal: 107,570,227; first deleted: 107,669,482; last deleted: 107,827,516; and first normal: 107,859,356. Therefore, the proximal breakpoint region is located in intron 1 and is 9.5 kb proximal to exon 2 while the distal breakpoint region is situated between the last exon of *COL4A5* and the *IRS4* gene, being located 10 kb distal to the last exon 51. The genomic positions indicated here are based on UCSC Hg18.

Supplementary figure S2



Supplementary figure S2. Junction analysis of *COL4A5* deletion in family 1. **A.** Agarose gel analysis showing the results of the PCR with forward primer 2 combined with reverse primer 7 (Supplementary table S2). For the patient (Pat) a 4 kb PCR product was obtained, which was absent in the control (Co). M is the 100 bp DNA ladder. **B.** Agarose gel image with the 3 kb junction DNA fragment. Nested PCR on PCR products obtained in A was done with the complementary sequences of reverse primer 2 and forward primer 7. **C.** Junction sequence from the nested PCR band. Both proximal (upper) and distal (lower) regions are located within LINE1 sequences. However, the proximal breakpoint is within a MA3 subtype while the distal one is located in a MB5 subtype thereby having no homology, excluding NAHR. No microhomology is found at the junction. Mismatches at the proximal and distal regions are highlighted in grey. Positions of the break sites are indicated and are based on UCSC Hg19.

Supplementary figure S3



Supplementary figure S3. Pedigree of family 1 with haplotypes for polymorphic markers adjacent to *COL4A5-COL4A6* genes (DXS1120, DXS1105, DXS1210, DXS456). Numbers indicate lengths of PCR products in base pairs (bp) for different alleles of the chosen microsatellites.

2.2. Patients with pathogenic COL4A3 *and* COL4A4 *mutations*

2.2.1. Collagen type IV-related nephropathies in Portugal: spectrum of pathogenic *COL4A3* and *COL4A4* mutations and clinical characterisation of 25 families (Manuscript 3)

Manuscripts 1 and 3 were submitted simultaneously to the Journal of Medical Genetics, on the 11th April 2014 (Manuscript 3 ID: jmedgenet-2014-102472).

Collagen type IV-related nephropathies in Portugal: spectrum of pathogenic *COL4A3* and *COL4A4* mutations and clinical characterisation of 25 families

Maria João Nabais Sá[1,2], Helen Storey[3], Frances Flinter[4], Mato Nagel[5], Susana Sampaio[2,6], Rui Castro[7], José Augusto Araújo[8], Maria Augusta Gaspar[9], Carlos Soares[10], Ana Oliveira[11], António Castro Henriques[12], António Gomes da Costa[13], Cristina Pinto Abreu[14], Pedro Ponce[15], Rui Alves[16], Liliana Pinho[11], Sérgio Estrela Silva[17], Carla Pinto de Moura [18,19], Luís Mendonça[20], Fernanda Carvalho[21], Manuel Pestana[6,22], Susana Alves[1], Filipa Carvalho[1], João Paulo Oliveira[1,2,20]

[1] Department of Genetics, Faculty of Medicine, University of Porto, Porto, Portugal

[2] Unit of Research and Development of Nephrology (FCT-725), Faculty of Medicine, University of Porto, Porto, Portugal

[3] DNA Laboratory, GSTS Pathology, Guy's and St. Thomas' Hospital National Health Service Foundation Trust, London, UK

[4] Genetics Centre, Guy's and St. Thomas' Hospital National Health Service Foundation Trust, London, UK

[5] Center for Nephrology and Metabolic Diseases, Weisswasser, Germany

[6] Department of Nephrology, Hospital de São João, Porto, Portugal

[7] Department of Nephrology, Centro Hospitalar de Trás-os-Montes e Alto Douro, Vila Real, Portugal

[8] Department of Nephrology, Hospital dos Marmeleiros, Funchal, Portugal

[9] Dialysis Clinic, NephroCare Restelo, Fresenius Medical Care, Lisboa, Portugal

[10] Department of Nephrology, Hospital de Braga, Braga, Portugal

[11] Dialysis Clinic Paredes, Diaverum, Paredes, Portugal

[12] Dialysis Clinic, NephroCare Braga, Fresenius Medical Care, Braga, Portugal

[13] Department of Nephrology, Hospital de Santa Maria, Lisboa, Portugal

[14] Dialysis Clinic Lumiar, Diaverum, Lisboa, Portugal

[15] Dialysis Clinic, NephroCare Lumiar, Fresenius Medical Care, Lisboa, Portugal

[16] Dialysis Clinic, NephroCare Viseu, Fresenius Medical Care, Viseu, Portugal
[17] Department of Ophthalmology, Hospital de São João, Porto, Portugal
[18] Department of Otolaryngology, Hospital de São João, Porto, Portugal
[19] Medical Genetics Outpatient Clinic, Hospital de São João, Porto, Portugal
[20] Department of Ophthalmology, Hospital de Braga, Braga, Portugal
[21] Unit of Renal Morphology, Department of Nephrology, Hospital Curry Cabral, Lisboa, Portugal
[22] Nephrology and Infectious Diseases Research and Development Group – INEB, University of Porto, Porto, Portugal

Corresponding author: Maria João Nabais Sá, Department of Genetics, Faculty of Medicine, University of Porto, Alameda Prof. Hernâni Monteiro, 4200 - 319 Porto, Portugal. E-mail: mjs.jano@gmail.com. Telephone: +351965546341. Fax: +351225513648.

Key words: Alport syndrome, Thin Basement Membrane Nephropathy, *COL4A3*, *COL4A4*, *COL4A5*.

ABSTRACT

Background

Pathogenic mutations in the *COL4A3/COL4A4* genes are responsible for the autosomally inherited collagen IV-related glomerular basement membrane (GBM) nephropathies, including the recessive (ARAS) and dominant (ADAS) forms of Alport syndrome (AS) and thin basement membrane nephropathy (TBMN). The purpose of this study was to identify pathogenic *COL4A3/COL4A4* mutations in unrelated Portuguese patients (n=40) with clinical suspicion of ARAS/ADAS/TBMN, and assess relevant genotype-phenotype correlations.

Methods

Mutation scanning of *COL4A3*/*COL4A4* was performed by Sanger sequencing of all exons and splice site regions. In 35 cases where X-linked AS (XLAS) could not be excluded by pedigree analysis, prior comprehensive mutational analysis of *COL4A5* had not shown any pathogenic mutations. Clinical phenotypes were compared between the apparently homozygous/compound heterozygous and the apparently heterozygous patients.

Results

Seventeen novel and four previously reported pathogenic *COL4A3/COL4A4* mutations were identified in 62.5% (25/40) of the probands. Regardless of the mutated gene, all patients with ARAS manifested chronic renal failure (CRF) and hearing loss, whereas only a minority of the apparently heterozygous patients had CRF or extrarenal symptoms. CRF was diagnosed at a significantly younger age in patients with ARAS. Overall, a pathogenic *COL4A3*/*COL4A4*/*COL4A5* mutation was identified in >50% of patients with fewer than three diagnostic criteria of AS.

Discussion

The occurrence of pathogenic *COL4A3*/*COL4A4* mutations in Portuguese patients with AS was higher, while the prevalence of XLAS was lower than expected. In such epidemiological context, next-generation sequencing of all three genes simultaneously may be the most cost-effective first-tier approach to the diagnosis of collagen IV-related GBM nephropathies.

INTRODUCTION

Alport syndrome (AS) is a hereditary glomerulopathy associated with sensorineural hearing loss (SNHL) and ocular anomalies, caused by mutations in the genes encoding the α3 (*COL4A3*), α4 (*COL4A4*) and α5 (*COL4A5*) chains of collagen type IV.(Kashtan 1999; Kashtan 2001; Tryggvason and Patrakka 2009) Although rare in the general population,(Levy and Feingold 2000) AS is one of the most common genetic disorders leading to advanced chronic renal failure (CRF) and end-stage renal disease (ESRD) in teenagers and young adults. Overall, AS was the primary renal disease in 0.56% of all patients commencing renal replacement therapy (RRT) in Europe, but accounted for 1.61% of the children requiring RRT before 15 years of age.(Rigden, Mehls et al. 1996) Pathogenic mutations in the *COL4A5* gene, mapping at chromosome location Xq22.3, cause X-linked AS (XLAS; MIM#301050), which is identified in approximately 80-85% of the families.(Kashtan 2001; Tryggvason and Patrakka 2009) Pathogenic mutations in either the *COL4A3* or *COL4A4* genes, mapping at chromosome location 2q36.3, cause the autosomally inherited forms of AS: recessive AS (ARAS; MIM#203780), which accounts for about 15% of the reported families, and dominant AS (ADAS; MIM#104200), which so far has been described in a small number of kindreds.(Kashtan 2001; Tryggvason and Patrakka 2009)

Familial persistent or recurrent hematuria, without significant proteinuria or extrarenal complications, is associated with heterozygosity for pathogenic mutations in the *COL4A3* or *COL4A4* genes in up to 60% of the families.(Lemmink, Nillesen et al. 1996; Badenas, Praga et al. 2002) Usually detected during childhood, this dominantly inherited condition is characterized, on electron microscopy (EM) examination of kidney tissue, by diffuse thinning of the glomerular basement membrane (GBM) to about half its normal thickness.(Rogers, Kurtzman et al. 1973; Yoshikawa, Matsuyama et al. 1988; Tiebosch, Frederik et al. 1989) Although commonly referred to as “benign familial hematuria” (BFH; MIM#141200), such term is a misnomer because a significant proportion of affected individuals develop proteinuric chronic kidney disease (CKD) and eventually need RRT later in life.(van Paassen, van Breda

Vriesman et al. 2004; Voskarides, Damianou et al. 2007; Pierides, Voskarides et al. 2009; Temme, Peters et al. 2012) For this reason, “thin basement membrane nephropathy” (TBMN) has been proposed as a better descriptor for this disorder, reflecting the underlying ultrastructural abnormality.(Savige, Rana et al. 2003) The lack of clinical or genetic markers to predict the natural history of renal disease in heterozygote carriers of pathogenic mutations in *COL4A3* or *COL4A4* prompted the recommendations for comprehensive family screening and periodic surveillance of affected individuals for signs of disease progression.(Frasca, Onetti-Muda et al. 2005; Marcocci, Uliana et al. 2009; Temme, Peters et al. 2012) Furthermore, such patients should be counseled about the possibility of having children affected with ARAS, particularly if the partner is consanguineous.(Frasca, Balestra et al. 2008) Likewise, patients with ARAS should recognize that their own progeny, as well as other first-degree relatives, may develop TBMN and, more rarely, even progressive CKD.

The spectrum of clinical manifestations associated with pathogenic mutations in the *COL4A3* and *COL4A4* genes, ranging from isolated hematuria to progressive CKD with ESRD, together with the mutations in the *COL4A5* gene causing XLAS, demonstrate the importance of these type IV collagen chains for the integrity of kidney structure and function, and led to the nosological concept of collagen IV-related GBM nephropathies.(Kashtan 2001; Longo, Porcedda et al. 2002) Genetic heterogeneity and clinical variability complicate the differential diagnosis of the various collagen IV-related GBM nephropathies, particularly in small pedigrees, as well as in young patients and sporadic cases. TBMN and ADAS may be regarded as the two extremes of severity of the familial patterns of expression of heterozygous *COL4A3* or *COL4A4* pathogenic mutations. The distinction between ADAS and TBMN is not easy, especially in children and young adults. Accurate diagnosis and prognostic information may also be difficult with XLAS in families where only females are affected.(Longo, Porcedda et al. 2002; Marcocci, Uliana et al. 2009) ARAS may be easily confused with XLAS in affected males, when the family history is uninformative. Immunohistochemical analysis of basement membrane expression of type IV collagen in a skin or kidney biopsy may be of

help in the differential diagnosis between those conditions.(Kashtan 2001; Tryggvason and Patrakka 2009) Comprehensive clinical investigation of probands and of their at-risk relatives, including screening for renal and extrarenal signs of AS in older family members, along with thorough pedigree analysis, are baseline key steps in the diagnostic process of collagen IV-related GBM nephropathies and genetic counseling of affected families.(Marcocci, Uliana et al. 2009)

Mutational analyses of the *COL4A3* and *COL4A4* genes can be used to establish the genetic diagnosis of autosomal collagen IV-related GBM nephropathies, obviating the need for a kidney biopsy.(Hoefele, Lange-Sperandio et al. 2010; Deltas, Pierides et al. 2012) Prior identification of the disease-causing mutation(s) in the family is usually a prerequisite for preimplantation genetic diagnosis (PIGD), for prenatal diagnosis (PND) of at-risk pregnancies, and for screening at-risk asymptomatic relatives in order to identify those who need specific medical follow-up. In very large pedigrees where the inheritance pattern is clearly established it may be possible to use linkage analysis to track the inheritance of the disease if specific mutation(s) cannot be identified. Furthermore, molecular genetic testing is a valuable adjunct in the differential diagnosis between collagen IV-related GBM nephropathies and other monogenic causes of glomerular hematuria, as well as with non-genetic diseases.(Norby and Cosio 2005; Deltas, Pierides et al. 2012; Deltas, Pierides et al. 2013)

The purpose of this study was to estimate the relative prevalence of pathogenic *COL4A3* and *COL4A4* mutations among Portuguese patients with diagnoses of collagen IV-related GBM nephropathies, in order to implement a cost-effective laboratory approach to their genetic diagnosis in our population.

MATERIAL AND METHODS

Patients and families

Between 1/January/2009 and 30/June/2012, in the setting of a nationwide study to identify the disease-causing mutations in Portuguese patients with the clinical diagnosis of AS, 40 apparently unrelated probands either with no detectable

pathogenic mutations in the *COL4A5* gene (n=35), as screened by Sanger sequencing and multiplex ligation-dependent probe amplification (MLPA; MRC-Holland), or with family history suggestive of autosomal transmission of kidney disease (n=5), were selected for molecular analysis of the *COL4A3* and *COL4A4* genes (Supplementary figure S1). In one of these five probands, direct sequencing of *COL4A3* and *COL4A4* was performed since the proband was a young female with AS who had family history of hematuria, but not of CKD. Four of these five probands of families with the diagnosis of AS (n=3) or TBMN (n=1), who attended the Inherited Kidney Diseases Clinic of São João Hospital Centre (CHSJ), Porto, Portugal, were known to carry pathogenic *COL4A3* or *COL4A4* mutations, identified through an outsourcing molecular genetics laboratory service (GENetic DIAgnostic Network, GENDIA; Antwerp, Belgium).

Patients were candidate for genetic analyses if they had history of urinary anomalies (e.g., intermittent or persistent microscopic hematuria with or without proteinuria) of unknown etiology and, at least, one of the other major diagnostic criteria of AS (Flinter, Cameron et al. 1988): (*i*) family history of hematuria, CRF and/or hearing loss; (*ii*) ultrastructural GBM changes, including thinning, thickening, splitting and electron-dense bodies inclusion; (*iii*) bilateral, high frequency SNHL; (*iv*) specific ocular lesion, including anterior lenticonus and/or central and peripheral dot-and-fleck retinopathy. CRF was defined as plasma creatinine (pCr) ≥1.2 mg/dl in males and ≥0.9 mg/dl in females, and advanced CRF was defined as pCr >1.5 mg/dl in males and >1.2 mg/dl in females. These cut-off values roughly correspond to the upper limits of the estimated glomerular filtration rate respectively of CKD stages 2 and 3, in young adult subjects, according to the guidelines for CKD evaluation, classification and stratification of the National Kidney Foundation (New York, NY, USA) [http://www.kidney.org/professionals/kdoqi/guidelines_ckd/p4_class_g1.htm].

In order to describe the phenotypic spectrum and the natural history of the disease in our cohort, a standardized questionnaire collecting data about the diagnostic criteria and other relevant features of AS identified on each patient was completed by his/her nephrologist or geneticist [manuscript submitted to the JMG simultaneously]. Whenever necessary and possible, retrospective

details of individual or familial medical histories were obtained by review of archive clinical records. In addition, available audiograms and EM photomicrographs of kidney biopsies were systematically rechecked.

As used herein, the term “proband” refers to the patient who first underwent mutational analyses for collagen IV-related GBM nephropathies in each family. Regarding the clinical phenotypic description, the diagnosis of AS was made strictly in patients or families who manifested three or four diagnostic criteria as proposed by Flinter and colleagues (Flinter, Cameron et al. 1988), i.e., in patients or families in whom at least one extra-renal criterion had manifested. Diagnosis of familial hematuria was made in families in whom extra-renal criteria of AS did not manifest, whether or not manifesting proteinuria or CRF. The term TBMN was strictly applied to those patients or families with microscopic hematuria in whom ultrastructural examination of a kidney biopsy showed thin GBM, whether or not manifesting proteinuria or CRF. In such families, the diagnosis of TBMN was assumed for all the relatives of a biopsied patient who also manifested microscopic hematuria with or without additional evidence of renal disease. Regarding the molecular description, the diagnosis of ARAS was made in patients presenting with typical manifestations of AS who were (apparently) compound heterozygotes or homozygotes for pathogenic *COL4A3* and/or *COL4A4* mutations, even if not all diagnostic criteria of AS had manifested.(Savige, Gregory et al. 2013) When one pathogenic *COL4A3* or *COL4A4* mutation was detected, the phenotype, at the familial level, was decisive to interpret the molecular result. The term ‘carrier of ARAS’ was used to denote a patient who carried a single pathogenic *COL4A3* or *COL4A4* mutation, whose phenotype ranged from absence of microscopic hematuria to progressive CRF and ESRD. The term ‘collagen type IV-related GBM nephropathy’ was used to refer to a patient or family in whom at least a pathogenic *COL4A5*, *COL4A4* or *COL4A3* mutation was identified.

The study was approved by the Health Ethics Commission of CHSJ, and patient enrollment for the genetic analyses required written informed consent.

Molecular study of the *COL4A3* and *COL4A4* genes

Genomic DNA samples of 30 probands without a pathogenic *COL4A5* mutation, identified in the first-tier molecular genetics diagnostic study, were screened for pathogenic mutations in *COL4A3* and *COL4A4* using tagged primer Sanger sequencing, either at the DNA Laboratory, Guy's and St. Thomas' National Health Service Foundation Trust, London, United Kingdom (n=18), or at the Center for Nephrology and Metabolic Diseases, Weisswasser, Germany (n=12). Polymerase chain reaction (PCR) products covering the entire coding sequence and the splice regions of the two genes were generated, purified and automatically sequenced by capillary electrophoresis, using standard laboratory methods and state-of-art equipment and techniques (details provided upon request).

The presence of the *COL4A3* and *COL4A4* sequence variants identified in the probands was confirmed in a second set of sequencing analyses at the Department of Genetics, Faculty of Medicine, University of Porto, Porto, Portugal. The confirmatory assays were limited to the relevant exons and corresponding exon-intron boundaries in each family, and were carried out as described in supplementary tables S2 and S3.(Boye, Mollet et al. 1998; Heidet, Arrondel et al. 2001) The same laboratory approach was used for screening at-risk relatives within the affected families.

Sequence variants were described according to the *COL4A3* reference sequence NM_000091.4 (http://www.ncbi.nlm.nih.gov/nuccore/NM_000091.4) and the *COL4A4* reference sequence NM_000092.4 (http://www.ncbi.nlm.nih.gov/nuccore/NM_000092.4), with the first nucleotide position corresponding to the first base of the translation start codon, using the nomenclature recommended by the Human Genome Variation Society.(den Dunnen and Antonarakis 2000)

The pathogenicity of the identified variants was presumed from the type of mutation – with nonsense mutations, splicing mutations within the consensus donor or acceptor splice regions and frameshifting mutations regarded as disease-causing –, and supported by the observed genotype-phenotype

correlations within the family. The inheritance pattern in each family was inferred from pedigree analysis. The possible impact of novel *COL4A3* and *COL4A4* missense point mutations upon the structure and function of the corresponding protein and/or mRNA splicing was predicted by *in silico* analyses with several bioinformatic tools (see details in the footer of Table I).

Genotype-phenotype correlations and statistical analyses

For genotype-phenotype correlation analyses, the study cohort consisted of both the probands and their affected and adult unaffected relatives prospectively identified on genetic screening. Subjects were divided according to their genetic status into a subcohort of homozygotes and compound heterozygotes for pathogenic *COL4A3* or *COL4A4* mutations, clinically corresponding to patients with ARAS, and a subcohort of individuals apparently carrying a single pathogenic mutation in either *COL4A3* or *COL4A4*, clinically corresponding to a wide phenotypic spectrum, including individuals without microscopic hematuria and individuals with AS who may have a second yet unidentified pathogenic *COL4A3* or *COL4A4* mutation.

The chi-square or the Fisher's exact tests were used as appropriate for comparisons of proportions. Survival analysis was used for comparisons of time to events. Non-parametric statistical tests were used to compare selected continuous clinical outcomes. The chi-square test for trend was used to assess the association between the numbers of diagnostic criteria of AS identified in each patient and the probability of finding a pathogenic mutation(s) in any of the COL4A3/COL4A4/COL4A5 genes. The 95% confidence intervals (95%CI) for prevalences were estimated by the modified Wald method.(Agresti and Coull 1998) The data were analyzed with the statistics software packages PASW® Statistics 18 (SPSS Inc., Chicago, IL, USA) or GraphPad Prism® 5.0 (GraphPad Software, Inc.; La Jolla, CA, USA).

RESULTS

COL4A3 and *COL4A4* mutations

Three previously reported (Heidet, Arrondel et al. 2001; Badenas, Praga et al. 2002; Wang, Rana et al. 2004) and 10 novel pathogenic mutations scattered along the *COL4A3* gene, as well as one recurrently described variant (Lemmink, Mochizuki et al. 1994; van der Loop, Heidet et al. 2000; Heidet, Arrondel et al. 2001; Badenas, Praga et al. 2002; Longo, Porcedda et al. 2002; Tazon Vega, Badenas et al. 2003; Wang, Rana et al. 2004) of uncertain significance (VUS) (p.Leu1474Pro) and one novel VUS (p.Phe1504Leu), were identified in 18 apparently unrelated families (Table I). Of the novel pathogenic mutations, five are small duplications or deletions that result in shifting of the reading frame and truncated protein translation; two are nonsense mutations; two involve critical nucleotides in canonical donor or acceptor splice junctions; and one is a glycine-changing missense mutation. All three previously known pathogenic mutations were typical glycine substitutions in the collagenous domain of the α3-chain. In the Portuguese cohort, one of these mutations (p.Gly407Arg) was identified in 8 apparently unrelated pedigrees, but all the other were unique to a single family. The p.Leu1474Pro VUS was identified in apparent compound heterozygosity with different pathogenic mutations, in two female probands. As both of these patients presented with a severe renal disease phenotype, with early progression to ESRD, they were included with the other homozygotes and compound heterozygotes for the genotype-phenotype correlation analyses. On the other hand, the p.Phe1504Leu VUS was not included in these correlation analyses, since this variant was not unequivocally causative of the proband's phenotype. This variant was detected in a female, in whom pathogenic mutations in *COL4A5* and *COL4A4* were not identified, who presented at 27-years-old with persistent microscopic hematuria, proteinuria and hypertension, diagnosed while she was pregnant, and whose mother, who was said to have had AS, had already deceased. At age 31, ultrastructural examination of kidney biopsy revealed GBM of variable thickness, with areas of thinning and wrinkling of the membrane, as well as

electron-dense bodies, but she did not have typical audiologic or ocular findings of AS.

One previously reported (Lemmink, Nillesen et al. 1996; Tazon Vega, Badenas et al. 2003) and 7 novel pathogenic mutations scattered along the *COL4A4* gene, as well as one recurrently described VUS (p.Gly999Glu),(Buzza, Dagher et al. 2003; Wang, Rana et al. 2004; Slajpah, Gorinsek et al. 2007) were identified in 9 families (Table II). Of the novel pathogenic mutations, two small duplications and one small deletion result in frameshifting and truncated protein translation; one is a small in-frame deletion; two involve critical nucleotides in canonical donor splice junctions; and one is a glycine-changing missense mutation. The previously reported pathogenic mutation was a typical glycine substitution in the collagenous domain of the α4-chain. All *COL4A4* mutations were unique to each family. The *COL4A4* p.Gly999Glu VUS was detected in heterozygosity in a 37-year-old female who had been on RRT since the age of 18 years and the sequencing of the *COL4A3* and *COL4A5* genes did not show any additional variants. However, as p.Gly999Glu in *COL4A4* seems to be of questionable pathogenicity, reaching a polymorphic allelic frequency of 0.11 in the Slovenian population,(Slajpah, Gorinsek et al. 2007) this proband has not been included in the genotype-phenotype correlation analyses.

Genetic screening of at-risk family members allowed identification of: 23 relatives of probands with pathogenic *COL4A3* mutations, three of whom were compound heterozygotes, while 20 were heterozygotes for one pathogenic mutation previously found in the proband; and 11 relatives of probands with pathogenic *COL4A4* mutations, all of whom were heterozygotes for the mutation found in the proband. Three of the heterozygote relatives (p.Pro135Glnfs*18 and p.Tyr481* in *COL4A3*; p.Ser1555* in *COL4A4*), including two relatives of patients with ARAS, were totally asymptomatic and in 6 there was insufficient clinical data to support a clinical diagnosis. These subjects were also included in the genotype-phenotype correlation analyses, according to their genetic status. In 11 of the families with *COL4A3* mutations and five of the families with *COL4A4* mutations the disease phenotype segregated with the identified mutation(s) at least in one relative of the proband.

None of the novel mutations identified in this study was present in more than 8,000 alleles from more than 4,000 unrelated European-American individuals listed on the Exome Variant Server.(Exome Variant Server)

Table I. Pathogenic mutations and allelic variants of unknown significance identified in the *COL4A3* gene in Portuguese patients with Alport syndrome and Thin Basement Membrane Nephropathy.

Mutation type	Exon or Intron	Nucleotide change	Predicted effect on the protein	Bioinformatic prediction of pathogenicity‡	Mutation status	Reference#	Family
Missense mutations							
	Exon 19	c.1114 G>C	p.(Gly372Arg)	Disease causing [a, b, c]	Compound heterozygosity		**56**
	Exon 21	c.1219G>C	p.(Gly407Arg)	Disease causing [a, b, c]	Compound heterozygosity	(Heidet, Arrondel et al. 2001)	**5,9,16,46,54, 55,A3_4 - 4/1, A3_4 - 6/1**
					Heterozygosity		
	Exon 28	c.2083G>A	p.(Gly695Arg)	Disease causing [a, b, c]	Heterozygosity	(Wang, Rana et al. 2004)	**A3_4 - 5/1**
	Exon 35	c.2954G>T	p.(Gly985Val)	Disease causing [a, b, c]	Compound heterozygosity	(Badenas, Praga et al. 2002)	**56**
Frameshift mutations							
	Exon 2	c.92_95dupGTGT	p.(Lys34Leufs*2)	Disease causing [a]	Homozygosity		**35**
					Heterozygosity		
	Exon 7	c.402delT	p.(Pro135Glnfs*18)	Disease causing [a]	Compound heterozygosity		**9**
					Heterozygosity		
	Exon 26	c.1845dupA	p.(Pro616Thrfs*30)	Disease causing [a]	Compound heterozygosity		**44**
	Exon 28	c.2111delC	p.(Pro704Leufs*43)	Disease causing [a]	Compound heterozygosity		**44**
	Exon 35	c.2914_2915delAG	p.(Gly973Argfs*53)	Disease causing [a]	Compound heterozygosity		**61**
Nonsense mutations							
	Exon 23	c.1443T>A	p.(Tyr481*)	Disease causing [a]	Homozygosity		**61, 62**
					Compound heterozygosity		
	Exon 30	c.2371C>T	p.(Arg791*)	Disease causing [a]	Compound heterozygosity		**38**

Splice site mutations							
	Intron 33	c.2657-1G>T	ND	Disease causing [d, e, f]	Compound heterozygosity		**63**
	Intron 44	c.3955+1G>C	ND	Disease causing [d, e, f]	Homozygosity Compound heterozygosity Heterozygosity		**38, 12**
Variant of unknown significant							
	Exon 48	c.4421T>C	p.(Leu1474Pro)	Disease causing [a, b, c]	Compound heterozygosity	(Lemmink, Mochizuki et al. 1994; van der Loop, Heidet et al. 2000; Heidet, Arrondel et al. 2001; Badenas, Praga et al. 2002; Longo, Porcedda et al. 2002; Tazon Vega, Badenas et al. 2003; Wang, Rana et al. 2004)	**55, 63**
	Exon 49	c.4510T>C	p.(Phe1504Leu)	Disease causing [a, b, c]	Heterozygosity		**47**

[#] Mutation previously reported in HGMD (http://www.hgmd.cf.ac.uk/, last accessed on March 1, 2014) and/or LOVD (http://www.lovd.nl/3.0/home, last accessed on March 1, 2014).
[‡] Superscript letters refer to different bioinformatic predictions of pathogenicity:
[a]: Predicted by Mutationt@ster (http://www.mutationtaster.org/, last accessed on March 1, 2014).
[b]: Predicted by Polyphen 2 (http://genetics.bwh.harvard.edu/pph2/, last accessed on March 1, 2014).
[c]: Predicted by SNPs&GO (http://snps-and-go.biocomp.unibo.it/snps-and-go/, last accessed on March 1, 2014).
[d]: Predicted by NNSplice (http://www.fruitfly.org/seq_tools/splice.html, last accessed on March 1, 2014).
[e]: Predicted by NetGene2 (http://www.cbs.dtu.dk/services/NetGene2/, last accessed on March 1, 2014).
[f]: Predicted by Human Splicing Finder (http://www.umd.be/HSF/, last accessed on March 1, 2014).
ND: effect on splicing not determined by mRNA analysis.

Table II. Pathogenic mutations and allelic variants of unknown significance identified in the *COL4A4* gene in Portuguese patients with Alport syndrome and Thin Basement Membrane Nephropathy.

Mutation type	Exon or Intron	Nucleotide change	Predicted effect on the protein	Bioinformatic prediction of pathogenicity‡	Mutation status	Reference#	Family
Missense mutations							
	Exon 30	c.2690G>A	p.(Gly897Glu)	Disease causing [a, b, c]	Heterozygosity	(Lemmink, Nillesen et al. 1996; Tazon Vega, Badenas et al. 2003)	**17**
	Exon 38	c.3506G>T	p.(Gly1169Val)	Disease causing [a, b, c]	Heterozygosity		**37**
Frameshift mutations							
	Exon 9	c.568_569dupGG	p.(Asp191Glyfs*29)	Disease causing [a]	Heterozygosity		**22**
	Exon 28	c.2216dupC	p.(Val741Cysfs*47)	Disease causing [a]	Homozygosity Heterozygosity		**4**
	Exon 47	c.4664_4665delCT	p.(Ser1555*)	Disease causing [a]	Heterozygosity		**25**
Small (in-frame) deletion							
	Exon 20	c.1323_1340del (18bp)†	p.(Gly442_Pro447del)	Disease causing	Heterozygosity		**8**
Splice site mutations							
	Intron 2	c.71+1G>A	ND	Disease causing [d, e, f]	Heterozygosity		**A3_4 - 2/1**
	Intron 31	c.2860+1G>A	ND	Disease causing [d, e, f]	Homozygosity		**40**

Variant of unknown significant							
	Exon 33	c.2996G>A	p.(Gly999Glu)	Disease causing [a, b, c]	Heterozygosity	(Buzza, Dagher et al. 2003; Wang, Rana et al. 2004; Slajpah, Gorinsek et al. 2007)	**30**

† c.1323_1340delTGGCTTGCCTGGAGCACC

[#] Mutation previously reported in HGMD (http://www.hgmd.cf.ac.uk/, last accessed on March 1, 2014) and/or LOVD (http://www.lovd.nl/3.0/home, last accessed on March 1, 2014).

[‡] Superscript letters refer to different bioinformatic predictions of pathogenicity:

[a]: Predicted by Mutationt@ster (http://www.mutationtaster.org/, last accessed on March 1, 2014).

[b]: Predicted by Polyphen 2 (http://genetics.bwh.harvard.edu/pph2/, last accessed on March 1, 2014).

[c]: Predicted by SNPs&GO (http://snps-and-go.biocomp.unibo.it/snps-and-go/, last accessed on March 1, 2014).

[d]: Predicted by NNSplice (http://www.fruitfly.org/seq_tools/splice.html, last accessed on March 1, 2014).

[e]: Predicted by NetGene2 (http://www.cbs.dtu.dk/services/NetGene2/, last accessed on March 1, 2014).

[f]: Predicted by Human Splicing Finder (http://www.umd.be/HSF/, last accessed on March 1, 2014).

ND: effect on splicing not determined by mRNA analysis.

Genotype-phenotype correlations

The diagnosis of ARAS was genetically confirmed in 12 probands and in three of their relatives (Supplementary table S4). Five of the probands were apparently homozygous for *COL4A3* (n=3) or *COL4A4* (n=2) mutations, the other seven being compound heterozygotes for *COL4A3* mutations. In two of these 12 families (12 and 35), it was confirmed that both probands' parents were heterozygotes for the pathogenic *COL4A3* mutation found in homozygosity in the proband. Parental consanguinity was observed in one family with a proband who was homozygous for a pathogenic *COL4A4* mutation (family 40) and in one family in which the proband was a compound heterozygote for two pathogenic *COL4A3* mutations (family 61). The higher prevalence of *COL4A3* mutations in probands diagnosed with ARAS was statistically significant (p=0.02).

Apparent heterozygosity for pathogenic *COL4A3* (n=7) or *COL4A4* (n=6) mutations was identified in 13 probands with clinical and/or pathology diagnoses of familial hematuria, TBMN or AS (Supplementary table S5) and, in the 8 families in whom at least one of the probands' parents was enrolled, the same pathogenic mutation was found to have been inherited from one of the parents (families 8, 16, 17, 22, 25, 37, 54, A3/4_2). In some probands with heterozygous *COL4A3* or *COL4A4* mutations who had personal or family history of typical extra-renal manifestations of AS, we hypothesize that a second pathogenic mutation remained undetected, as it is the case of the proband of family 8 who reached ESRD at age 23 and had typical AS maculopathy, and in whom direct sequencing of *COL4A4* detected a single pathogenic mutation (p.Gly442_Pro447del). Moreover, in this proband, pathogenic *COL4A5* and *COL4A3* variants were not found, by direct sequencing of both genes. Additionally, we hypothesize that audiological and ocular manifestations, which may mimic extra-renal features of AS, may have other distinct causes, which should be pursued. The *COL4A4* mutation p.Val741Cysfs*47 caused ARAS, in homozygosity, and was also found in the 58-year-old heterozygote sister of the proband who manifested microscopic hematuria and CKD, concurrently with hearing loss.

Regarding the severity of the renal disease, six out of seven pathogenic *COL4A3* mutations and all the six pathogenic *COL4A4* mutations identified in apparently heterozygous probands were associated with microscopic hematuria and proteinuric CKD and, in five probands, even with stage 2 CKD (p.(Gly407Arg) in the *COL4A3* gene; and p.Gly897Glu, p.Asp191Glyfs*29, c.71+1G>A in the *COL4A4* gene). The recurrent p.Gly407Arg *COL4A3* mutation caused ARAS when present together with a second pathogenic mutation and, in heterozygotes, was associated with a wide array of age-dependent clinical phenotypes, ranging from absence of microscopic hematuria, to isolated microscopic hematuria and to ESRD.

The proportion of probands (as compared with affected relatives of probands), was significantly higher in the subcohort of homozygous/compound heterozygous patients than in the subcohort of apparently heterozygous patients (Table III), but the median ages at enrollment were similar in the two subcohorts. Gender distribution also did not significantly differ between the two groups. The diagnosis of microscopic hematuria was made at a significantly younger median age in the homozygous/compound heterozygous patients (12 vs. 26 years; p=0.006). As compared with the (apparently) heterozygous patients, the prevalence of proteinuria, hypertension and CRF were significantly higher in the former, while the median ages at diagnosis were significantly lower. Proteinuria and hypertension were diagnosed at median ages of 16 and 25 years respectively (p=0.002) and of 18 and 42 years (p<0.001) in each of the sub-cohorts. CRF was diagnosed at a median age of 20 years in the homozygous/compound heterozygous patients and at a median age of 46 years (p<0.001) in the (apparently) heterozygous patients. All homozygous/compound heterozygous patients, but only 14% of the (apparently) heterozygous patients, had already started RRT (p<0.001); however, the difference in the median ages at start of RRT between the two groups did not reach statistical significance. Significantly more homozygous/compound heterozygous patients reported subjective hearing loss. Among patients who underwent formal ophthalmological examination, the prevalence of anterior lenticonus and of dot-and-fleck retinopathy was also significantly higher in this subcohort. Diagnostic

kidney biopsies were performed at significantly younger ages in homozygous/compound heterozygous patients, but the prevalence of distinctive GBM ultrastructural features of AS did not differ between the two groups.

The overall *COL4A3/COL4A4/COL4A5* mutation detection rates in the 24 probands that had been comprehensively evaluated for all the diagnostic criteria of AS, are presented in Table IV. A highly significant linear trend ($p<0.01$) was observed between the number of diagnostic criteria of AS identified in each patient and the probability of detecting at least a pathogenic mutation in any of the three genes. Of note, the probability of identifying a pathogenic *COL4A3, COL4A4 or COL4A5* mutation in patients with microscopic hematuria and two additional diagnostic criteria of AS was 57%.

In general, the severity of the clinical expression of AS in the homozygous/compound heterozygous patients was comparable with that of males with XLAS, both groups showing high rates of ESRD and of SNHL manifesting before the fourth decade of life. When the disease expression was compared between males and females carrying a *COL4A3* or *COL4A4* mutation in apparent heterozygosity and females with XLAS (i.e., heterozygotes for a pathogenic mutation in *COL4A5*), no major differences were identified except in the prevalence of microscopic hematuria, which was a constant finding in XLAS females but was not manifested by a significant number of the heterozygotes for a pathogenic mutation in *COL4A3* or *COL4A4* that were prospectively identified on family screening.

Table III. Descriptive characteristics of Portuguese Alport syndrome patients with pathogenic *COL4A3* and *COL4A4* mutations by mutation status (homozygotes and compound heterozygotes versus heterozygotes).

	Homozygotes and Compound Heterozygotes (n=15)		Heterozygotes (n=44)		p-value
		N		N	
Phenotype					
Index-cases (%)	80	12/15	30	13/44	0.001
Male gender (%)	53	8/15	32	14/44	0.137
Age at enrollment (median (interquartile range))	42 (17)	15	46 (26.8)	44	0.296
Renal anomalies					
History of macroscopic hematuria (%) Age at diagnosis (median (interquartile range))	55 8 (15*)	6/11 3	12 17 (12*)	3/25 3	0.012 0.343
History of microscopic hematuria (%) Age at diagnosis (median (interquartile range))	100 12 (14.3)	9/9 6	77 26 (20.5)	27/35 26	0.175 0.006
History of proteinuria (%) Age at diagnosis (median (interquartile range))	100 16 (13.5)	12/12 9	70 25 (19.5)	23/33 22	0.042 0.002
History of hypertension (%) Age at diagnosis (median (interquartile range))	92.3 18 (14.8)	12/13 8	52 42 (18.5)	17/33 13	0.016 0.000
History of CKD stage 2 or higher (%) Age at diagnosis (median (interquartile range)) eGFR at diagnosis (median (interquartile range))	100 20 (10) 59 (15*)	14/14 9 3	51 46 (19.5) 67 (20)	18/35 17 15	0.001 0.000 0.236
History of renal replacement therapy (%) Age at onset (median (interquartile range)) eGFR at onset (median (interquartile range))	100 23 (12) 9.5 (3*)	15/15 15 2	14 36 (22) 14 (-)	5/36 5 1	0.000 0.042 0.221
GBM ultrastuctural anomalies Age at diagnosis (median (interquartile range))	 17.5 (7.8)	 6	 34 (19)	 9	 0.004
Thinning (%)	100	5/5	100	8/8	a
Thickening (%)	100	4/4	100	3/3	a
Lamellation (%)	100	5/5	60	3/5	0.444
Electrondense bodies (%)	100	5/5	67	2/3	0.375
Hearing loss					
Self-noticed or subjective (%) Age at self-noticed or subjective (median (interquartile range))	93.3 9 (19)	14/15 9	35 36 (32*)	12/34 3	0.000 0.086
Hearing loss confirmed by audiogram (%) Age at diagnosis of hearing loss by audiogram (median (interquartile range))	90 32 (22)	9/10 7	53 44 (26.5)	10/19 9	0.098 0.009
Ocular anomalies					
Anterior Lenticonus (%) Age at diagnosis (median (interquartile range))	33 30 (-)	3/9 1	0 -	0/25 -	0.014 -
Maculopathy (%) Age at diagnosis (median (interquartile range))	63 34 (25*)	5/8 3	9.1 40.5 (35*)	3/33 2	0.003 0.610
Cataracts (%) Age at diagnosis (median (interquartile range))	83 42.5 (17*)	5/6 2	12 58 (15*)	3/25 3	0.002 0.207

SD: standard deviation. Age is expressed in years. eGFR is expressed in ml/min/1.73m^2. *The range is indicated instead of the interquartile range. [a] Statistics cannot be computed for this variable because the proportion in the two groups is a constant.

Table IV. *COL4A3*, *COL4A4* and *COL4A5* detection mutation rate per number of diagnostic criteria met, among probands tested for *COL4A3*, *COL4A4* and *COL4A5*, with complete clinical evaluation (n=23).

Diagnostic criteria met	***COL4A5-positive probands (n=7)***	***COL4A3 or COL4A4-positive probands (n=10)***	***COL4An-negative probands (n=6)***	***Total of probands studied (n=23)***	***COL4A5 mutation detection rate per number of diagnostic criteria met (%)***	***COL4A3/4 mutation detection rate per number of diagnostic criteria met (%)***	***COL4An mutation detection rate per number of diagnostic criteria met (%)***
1	0	0	2	2	0	0	0
2	2	2	3	7	29	29	57
3	1	7	1	9	11	78	89
4	4	1	0	5	80	20	100

DISCUSSION

We have identified 17 novel and 4 previously reported pathogenic *COL4A3* or *COL4A4* mutations in 25 (62.5%) patients from a cohort of 40 apparently unrelated probands from Portuguese families with a diagnosis of non-XLAS or of TBMN. Exclusion of X-linked inheritance was based on pedigree evidence or the negative result of previous mutational analysis of *COL4A5*. Although gross deletions and gross insertions/duplications of *COL4A3* or *COL4A4* have rarely been reported in patients with collagen IV-related GBM nephropathies [Human Gene Mutation Database – HGMD®; http://www.hgmd.cf.ac.uk/ac/gene.php?gene=COL4A4 and http://www.hgmd.cf.ac.uk/ac/gene.php?gene=COL4A3, accessed on March 1, 2014], the lack of testing for deletions/duplications not readily detectable by sequence analysis was a major limitation of our study, and deep intronic mutations – like those affecting the normal splicing of *COL4A5* mRNA, described in families with XLAS (King, Flinter et al. 2002; King, Flinter et al. 2006) –, might have been overlooked as well. Consequently, we may have slightly underestimated the contribution of pathogenic *COL4A3* or *COL4A4* mutations to the molecular genetics epidemiology of collagen IV-related GBM nephropathies in Portugal, and even possibly missed their genetic diagnosis in some families.

The overall frequencies of different types of pathogenic *COL4A3* and *COL4A4* mutations identified in our cohort did not significantly differ from those reported to the HGMD®. Small deletions/duplications were predominant, constituting 42.9% (9/21) of the total. In all but one case, these mutations caused a shift in the mRNA reading frame with premature termination of translation. Missense mutations, all of them involving glycine substitutions, accounted for 28.6% (6/21). Like in other genetic disorders of collagen,(Shoulders and Raines 2009) glycine substitutions in the collagenous domain of the α3, α4 or α5 chains of type IV collagen are a common type of pathogenic mutation in AS.(Kashtan 2001) Such mutations are thought to interfere with the normal folding of the mutant chain into triple helices with other type IV collagen α-chains, leading to increased susceptibility to proteolytic degradation. The *COL4A3* glycine substitution p.Gly407Arg was identified in 8 apparently unrelated probands, making it a

particularly frequent disease-causing mutation in Portuguese families with collagen IV-related GBM nephropathies. Two nonsense point mutations were identified in *COL4A3*, accounting for 9.5% (2/21) of all pathogenic mutations identified in this study. The four splice site mutations, which accounted for 19% (4/21) of the total, were all novel and affected the highly conserved guanine residues at positions +1 or –1, respectively in the consensus donor and acceptor splice regions.

Twelve probands were apparently homozygous or compound heterozygotes for pathogenic *COL4A3* or *COL4A4* mutations. Although we were not able to confirm the heterozygous condition of their parents in all cases, the phenotypes of these probands, and their family histories, were compatible with a clinical diagnosis of ARAS. In the ARAS patients, the disease was five times more frequently caused by *COL4A3* mutations than by mutations in *COL4A4*.

Taking into account the data reported herein and the results of our previous scanning for *COL4A5* mutations, [manuscript submitted to the JMG simultaneously] the prevalences of XLAS and ARAS in Portuguese families with clinical and/or EM diagnosis of AS are respectively 35.5% (22/62; 95%CI: 25–48%) and 19.4% (12/62; 95%CI: 11–31%). On the other hand, in 15 out of 62 (24%) of those families where all three genes were analyzed, no pathogenic mutations could be identified. Overall, the results of our *COL4A3/COL4A4/COL4A5* genetic analyses allowed us to confirm the diagnosis of a collagen IV-related GBM nephropathy in about two-thirds of 65 families, of which 46.8% (22/47) had XLAS and 53.2% (25/47) had ARAS or familial hematuria/TBMN with or without progressive renal function deterioration. The lower than expected prevalence of XLAS in our cohort is in accordance with the results of a recent Italian study,(Fallerini, Dosa et al. 2013) which showed that the prevalence of patients with familial hematuria or TBMN with CRF, carrying one pathogenic *COL4A3* or *COL4A4* mutation may be underestimated, at least in specific populations. The >50% mutation detection rate in the Italian patients with only 1 or 2 diagnostic criteria of AS, is also consistent with our own findings. The proportion of Portuguese families with no pathogenic mutation identified in any of the three genes was significantly lower than in the Italian study.

The hemizygous condition for pathogenic *COL4A5* mutations, as well as the homozygous, compound or double heterozygous conditions for pathogenic *COL4A3* and/or *COL4A4* mutations, are associated with a significantly higher risk of developing ESRD and extrarenal complications of AS, manifesting at younger ages, than the heterozygous condition for any of the three genes.(Jais, Knebelmann et al. 2000; Jais, Knebelmann et al. 2003; Marcocci, Uliana et al. 2009; Temme, Peters et al. 2012; Storey, Savige et al. 2013) Likewise, we observed a significantly higher frequency of ESRD and SNHL in our patients with apparently homozygous or compound heterozygous pathogenic mutations in *COL4A3* or *COL4A4*, when compared with the apparently heterozygous patients; median ages at the diagnosis of advanced CRF were in the early second decade of life in the former and in the beginning of the fourth decade of life in the latter patients; and the apparently heterozygous condition was less often associated with extrarenal manifestations. These observations are consistent with current knowledge about the collagen IV-related GBM nephropathies, suggesting that we did not misclassify a significant number of families with ARAS as having a heterozygous condition. Furthermore, the lower than expected prevalence of XLAS in Portuguese patients diagnosed with AS, which was a major finding of this study, cannot be explained by the methodological limitations discussed above.

Given the age-related penetrance of their clinical manifestations, the phenotypic prevalence of familial hematuria/TBMN with progressive CRF in heterozygous patients for pathogenic *COL4A3* or *COL4A4* mutations increases with age, even into late adulthood.(Voskarides, Damianou et al. 2007; Pierides, Voskarides et al. 2009; Temme, Peters et al. 2012) The risk of CKD progression to ESRD in late adulthood for patients carrying a single *COL4A3* or *COL4A4* pathogenic mutation (Voskarides, Damianou et al. 2007; Marcocci, Uliana et al. 2009; Temme, Peters et al. 2012) must be acknowledged for proper planning of clinical follow-up and in genetic counseling. Furthermore, it is particularly important to it be addressed with prospective donors when live-related kidney transplantation is being considered for patients with collagen IV-related GBM nephropathies.(Gross, Weber et al. 2009)

It is recommended that all at-risk relatives of patients diagnosed with any of the autosomally inherited collagen IV-related GBM nephropathies should be comprehensively screened for the typical disease manifestations, either renal or extrarenal, and that individuals with urinary abnormalities (even asymptomatic), or with hypertension, should be periodically re-evaluated for CKD progression.(Frasca, Onetti-Muda et al. 2005; Temme, Peters et al. 2012)

Genotyping is the most clinically useful diagnostic method of collagen IV-related GBM nephropathies, providing an effective tool for age- and phenotype-independent family screening, and for accurate genetic counselling. It also allows for earlier therapeutic intervention with renin–angiotensin–aldosterone system (RAAS) blockade in carriers of single *COL4A3* or *COL4A4* pathogenic from families with autosomal dominant collagen IV-related GBM nephropathies, and is a key requisite for primary prevention of those diseases by PIGD or PND. It is advisable that all carriers of a single *COL4A3* or *COL4A4* mutation have lifelong renal follow-up, even if asymptomatic at baseline.

Given the high number of exons per gene in *COL4A3*, *COL4A4*, *COL4A5*, (respectively 52, 48 and 51) and the absence of mutational hot-spots due to the private nature of most pathogenic mutations, a primary molecular genetics diagnostic approach to collagen IV-related GBM nephropathies by Sanger sequencing becomes quite laborious, time consuming and expensive, particularly in those families without a clear inheritance pattern. Next-generation sequencing (NGS) is a promising new analytical method for detection of pathogenic mutations in genetically heterogeneous disorders like AS.(Artuso, Fallerini et al. 2012; Hertz, Thomassen et al. 2012) The relative prevalence of XLAS as compared with ARAS in our cohort was about fivefold lower than expected, and the mutation detection rate in *COL4A5/COL4A4/COL4A3* was >50%, even in patients with only one or two diagnostic criteria of AS. In these conditions, NGS may be particularly cost-effective as first-tier approach to the genetic diagnosis of patients with clinical suspicion of AS.

ACKNOWLEDGEMENTS

This work is part of the PhD project of Maria João Nabais Sá (supervisor: João Paulo Oliveira), to be presented at the ICBAS – "Instituto de Ciências Biomédicas Abel Salazar" (Abel Salazar Biomedical Sciences Institute), University of Porto, Portugal. The studies reported herein were mainly supported by the European Regional Development Fund (ERDF) through the Competitive Factors Operational Programme (COMPETE), and by the Portuguese Foundation for Science and Technology (FCT), in the framework of the Cooperation Agreement between Portugal and Harvard Medical School-2009, through the project FCOMP-01-0124-FEDER-HMSP-ICJ/SAU-ICT/0003/2009. Additional funding was obtained from a research grant of the Portuguese Society of Nephrology. The Unit of Research & Development in Nephrology (FCT-725) at the Department of Nephrology of the Faculty of Medicine, University of Porto, and the São João Hospital Centre, Porto, Portugal, provided secretarial support.

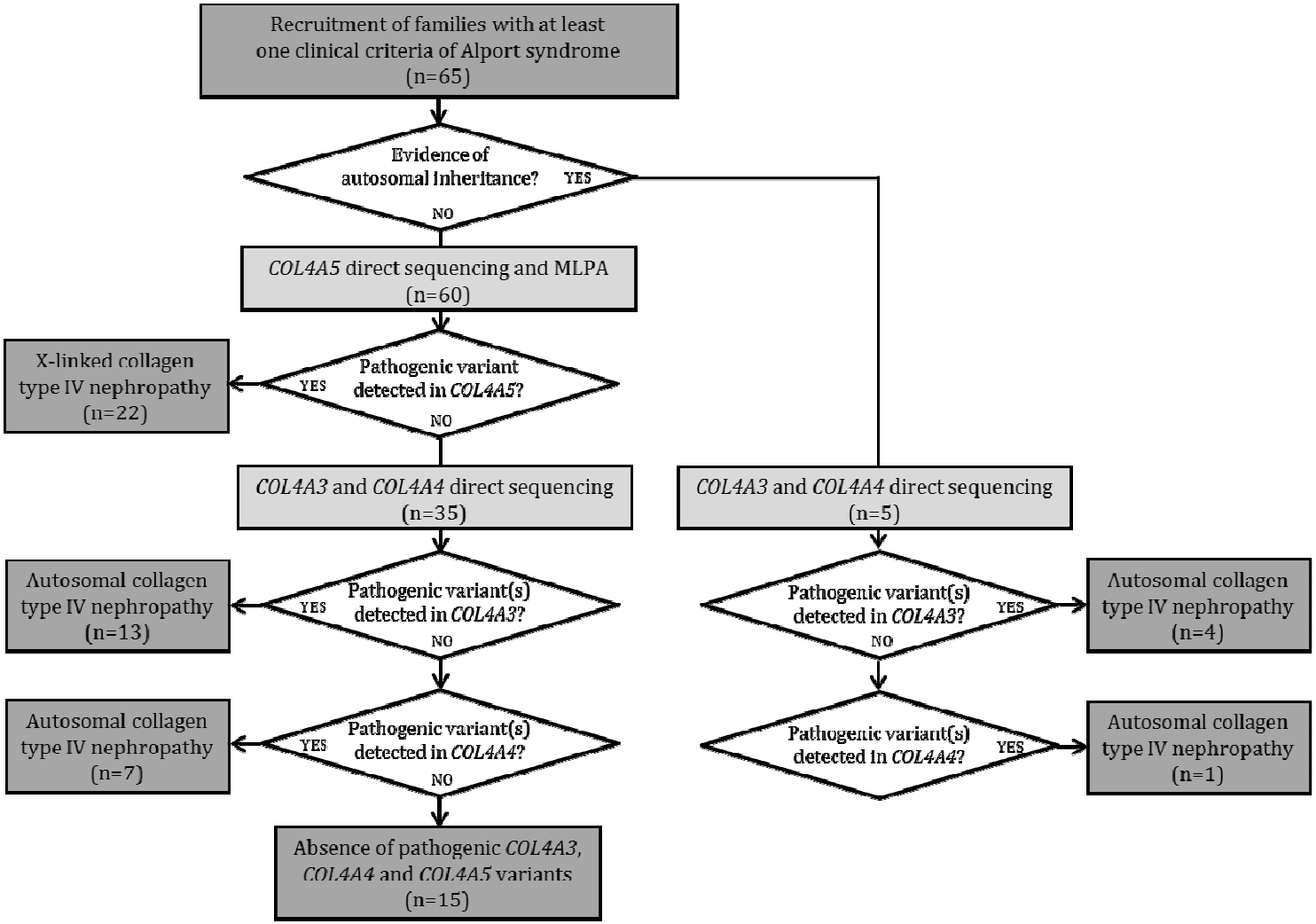
Complete genotyping of the *COL4A3* and *COL4A4* genes was performed in 30 families as research collaboration with *GSTS Pathology* (London, UK) and *Moldiag – Center for Nephrology and Metabolic Diseases* (Weisswasser, Germany). In 5 additional families, the complete genotyping of the *COL4A3* and *COL4A4* genes was performed as outsource diagnostic service by *Gendia* (Antwerp, Belgium).

The authors thank the patients and their relatives for the invaluable collaboration in this study; NephroCare – Portugal, for having allowed the nationwide ascertainment of patients with Alport syndrome in their dialysis center; Dr.s Adelino Carvalho (NephroCare, Santarém), Ana Bernardo (NephroCare, Covilhã), Bernardo Faria (NephroCare, Guarda), Carlos Botelho (NephroCare, Viseu), Joana Felgueiras (NephroCare, Montijo), Liliana Pinho (Dialysis clinic, Paredes), Luís Carvalho (NephroCare, Restelo) and Miguel Gonçalves Rocha (Hospital de Braga) for contributing clinical data of patients enrolled in this study; Dr.s Ana Oliveira, Augusta Praça, Isabel Tavares, João Frazão, Ricardo Neto (Department of Nephrology, São João Hospital Centre, Porto), and Dr. Ana Gomes (Department of Nephrology, Hospital Centre, Vila Nova de Gaia) for referring patients for this study.

The authors also thank Dr. Francis Earl Cook, from the Department of Epidemiology, Harvard School of Public Health, Boston, MA, for his advice and supervision of the epidemiological and statistical analyses.

Parts of these data were presented as posters at the European Human Genetics Conference 2012, Nürnberg, Germany, June 23-26, 2012 (P12.007) and at the 62nd Annual Meeting of the American Society of Human Genetics, November 6-10, 2012 San Francisco, California (3122F).

Supplementary figure S1. Study flowchart.



Supplementary table S2: Primer sequences, annealing temperatures and product size for PCR amplification of the *COL4A3* gene.[†]

Exon	Forward primer	Reverse primer	Annealing temperature	PCR product size
2	CTCCTAACAGATAGTGTCC	CGGATCTTTCTGCTACCGAT	56ºC	267 bp
7	AGCCTCATGACCCAGTAGC	GAATGTCTACTATGGCTACC	58ºC	308 bp
19	TGA TGT TTG GTG AGC TGT C	TCC TGC CCT CCA GCA AAT G	58ºC	417 bp
21	TCTCCATTGTGCAATTTTTA	CTAAGCTGTGAGGAGGGTTT	55ºC	367 bp
23	TCT GAG GAC TCA ATG TAG C	CGT TGT GAC ATC CAC CAA T	56ºC	242 bp
26	ACGGAAGAAACCTGCAGTG	CCTGGTATCTGTGAGTTGAC	56ºC	483 bp
28	GTGCAAAAGGGATAGGACG	CTTCTAAATATCCACAACAA	55ºC	259 bp
30	GTTGATGACATGGTAGTGG	GGAAGAAAGGTAGGGCAGG	58ºC	391 bp
33	GAC CCA TCT CCT AGA CTA A	GGC TAA CAG TGC TGA GAA C	56ºC	499 bp
35	TCG TTC TGT CGC CCA CGC T	CTG CTG GAA CAC TAT CAG A	58ºC	392 bp
44	ACTTCAGCTTATTCTCACCC	GGAGCCTTCCTACTTCACTC	61ºC	333 bp
48	GCC ACT CTT CTC TAG GAT T	GTA TCA GGA GAC ACT CAA G	56ºC	458 bp

[†] Cycling conditions (adapted from (Heidet, Arrondel et al. 2001)) were: initial denaturation for 5 minutes at 94°C, followed by 35 cycles of denaturation at 94°C for 60 seconds, annealing for 60 seconds, and extension at 72°C for 90 seconds, ending with a final extension at 72°C for 10 minutes. The PCR products were automatically sequenced in forward and reverse directions in an ABI Prism® 310 Genetic Analyzer (Applied Biosystems) using the BigDye® Terminator v3.1 Cycle Sequencing kit, and the corresponding electropherograms were visually inspected for the anticipated mutation in each case.

Supplementary table S3: Primer sequences, annealing temperatures and product size for PCR amplification of the *COL4A4* gene.[†]

Exon	Forward primer	Reverse primer	Annealing temperature	PCR product size
2	TCTGGAAGAGAAGACTGGCA	AAGCAGGCAATCACACTGAG	60ºC	152 bp
9	ACCTCCGCATATCCCTTCC	TTCATGTTCTGTGGTCGCC	58ºC	375 bp
20	CTCCAGCTCCGTCTCTTTC	TGCTTTCTTAGTGGCACTG	60ºC	365 bp
28	ATTGGTTCTATACTTGCACA	TCTATGCACCAAAAGGACAG	55ºC	309 bp
30	TGCTGTGTGTGAAGCCAGTG	AGGACAAGAGCAAGGGAGG	61ºC	400 bp
31	GACTCTGTCCACCTCCCAC	CCAAGCTTCAGACAAGTCC	61ºC	328 bp
33	GCT TAA CTA TTA CCT AGC TC	TCT CCT TGA GCT ACA CCT T	56ºC	383 bp
38	GCGTTTGTGGCTAGAGTGAG	ATACCAGGGAGGGTACCAC	61ºC	218 bp
47	GCAAAGGCCAGATGGAGG	GCGGGAGAAGGTGTTAGG	56ºC	433 bp

[†] Cycling conditions (adapted from (Boye, Mollet et al. 1998)) were: initial denaturation for 5 minutes at 94°C, followed by 35 cycles of denaturation at 94°C for 60 seconds, annealing for 60 seconds, and extension at 72°C for 90 seconds, ending with a final extension at 72°C for 10 minutes. The PCR products were automatically sequenced in forward and reverse directions in an ABI Prism® 310 Genetic Analyzer (Applied Biosystems) using the BigDye® Terminator v3.1 Cycle Sequencing kit, and the corresponding electropherograms were visually inspected for the anticipated mutation in each case.

Supplementary table S4. Genotype and phenotype of Portuguese probands with pathogenic homozygous or compound heterozygous *COL4A3* or *COL4A4* mutations and of their relatives with heterozygous, homozygous or compound heterozygous condition.

FAMILY	35/AS-1	9/AS-1	56/AS-1	55/AS-1	62/AS-1	61/AS-1	44/AS-1	38/AS-1	63/AS-1	12/AS-1	4/AS-1	40/AS-1
GENOTYPE												
Gene	COL4A3	COL4A3	COL4A3	COL4A3	COL4A3	COL4A3	COL4A3	COL4A3	COL4A3	COL4A3	COL4A4	COL4A4
Mutation status	Hom	Comp Het	Comp het	Comp het	Hom	Comp het	Comp Het	Comp Het	Comp Het	Hom	Hom	Hom
Allele 1												
Exon	2	7	19	21	23	23	26	30	Intron 33	Intron 44	28	Intron 31
Nucleotide change	c.92_95dupGTGT	c.402delT	c.1114G>C	c.1219G>C	c.1443T>A	c.1443T>A	c.1845dupA	c.2371C>T	c.2657-1G>T	c.3955+1 G>C	c.2219_2220insC	c.2860+1G>A
Predicted effect on the protein	p.(Lys34Leufs*2)	p.(Pro135Glnfs*18)	p.(Gly372Arg)	p.(Gly407Arg)	p.(Tyr481*)	p.(Tyr481*)	p.(Pro616Thrfs*30)	p.(Arg791*)	ND	ND	p.(Val741Cysfs*47)	ND
Mutation type	frameshift	frameshift	missense	missense	nonsense	nonsense	frameshift	nonsense	splicing	splicing	frameshift	splicing
Allele 2												
Exon	2	21	35	48	23	35	28	Intron 44	48	Intron 44	28	Intron 31
Nucleotide change	c.92_95dupGTGT	c.1219G>C	c.2954G>T	c.4421T>C	c.1443T>A	c.2914_2915delAG	c.2111delC	c.3955+1G>C	c.4421T>C	c.3955+1 G>C	c.2219_2220dupC	c.2860+1G>A
Predicted effect on the protein	p.(Lys34Leufs*2)	p.(Gly407Arg)	p.(Gly985Val)	p.(Leu1474Pro)	p.(Tyr481*)	p.(Gly973Argfs*53)	p.(Pro704Leufs*43)	ND	p.(Leu1474Pro)	ND	p.(Val741Cysfs*47)	ND
Mutation type	frameshift	missense	missense	missense	nonsense	frameshift	frameshift	splicing	missense	splicing	frameshift	splicing
PHENOTYPE AND ORIGIN												
Gender, [age at enrollment in years]	F, [19]	M, [50]	F, [52]	F, [46]	M, [50]	M, [38]	M, [33]	M, [45]	F, [33]	F, [26]	M, [51]	M, [42]
Birth district	Vila Real	Porto	Lisboa	Braga	Madeira	Madeira	Lisboa	Castelo Branco	Lisboa	Porto	Porto	Viseu
Proband's clinical and histopathological phenotype												
History of kidney disease												
Macroscopic hematuria[a]	NA	- [NA]	+ [20]	- [NA]	NA	NA	+ [NA]	+ [5]	+ [8]	+ [NA]	NA	- [NA]
Microscopic hematuria[a]	+ [5]	+ [19]	NA	NA	+ [30]	+ [NA]	+ [<5]	NA	+ [8]	+ [15]	NA	+ [9]
Proteinuria[a]	+ [5]	+ [17]	+ [20]	NA	+ [30]	+ [NA]	+ [16]	+ [NA]	+ [8]	+ [15]	NA	+ [9]
Hypertension[a]	+ [NA]	+ [17]	+ [NA]	+ [NA]	+ [30]	NA	+ [16]	+ [NA]	+ [13]	- [NA]	+ [29]	+ [9]
CKD stage 2 or higher[a]	+ [NA]	+ [27]	+ [34]	+ [NA]	+ [NA]	+ [NA]	+ [16]	+ [NA]	+ [17]	+ [19]	+ [28]	+ [20]
pCr (eGFR) at diagnosis of CKD stage 2 or higher	NA	1.6 (52)	NA	NA	NA	NA	NA	NA	1.2 (59)	1.07 (57)	NA	NA
Renal replacement therapy[b]	DP [14], KTx [15]	HD [31], KTx [31]	HD [42], KTx [NA]	HD [17]	HD [43]	HD [24]	HD [23]	HD [23], KTx [30]	PD [22]	PD [24]	HD [29], KTx [32]	HD [20]
pCr (eGFR) at onset of renal replacement therapy	NA	6.1 (11) [30]	NA	NA	NA	NA	NA	NA	NA	6.1 (8)	NA	NA
Hearing loss												
Self-noticed (subjective)[a]	+ [NA]	+ [9]	+ [24]	- [45]	+ [30]	+ [NA]	+ [8]	+ [8]	+ [NA]	+ [NA]	+ [8]	+ [19]
SNHL, audiologically confirmed[a]	+ [18]	+ [36]	NA	- [45]	NA	NA	+ [NA]	+ [9]	+[32]	+ [25]	+ [NA]	NA
Ocular signs												
Anterior lenticonus[a]	NA	- [34]	- [NA]	- [45]	+ [30]	NA	+ [NA]	NA	NA	- [26]	- [51]	NA
Maculopathy[a]	NA	+ [34]	- [NA]	- [45]	NA	+ [NA]	NA	NA	NA	+ [26]	+ [51]	NA
Cataracts[a]	+ [NA]	+ [34]	NA	NA	NA	NA	NA	+ [NA]	NA	- [26]	+ [51]	NA
Ultrastructural GBM abnormalities on kidney biopsy[c]	[11]	[19]	[20]		[31]				[14]	[16]		
Thinning	+	+	+	NA	NA	NA	NA	NA	NA	+	NA	NA
Thickening	+	+	NA	NA	NA	NA	NA	NA	NA	+	NA	NA
Lamellation	+	+	+	NA	NA	NA	NA	NA	NA	+	NA	NA
Electrondense bodies	+	+	+	NA	NA	NA	NA	NA	NA	+	NA	NA
Family history[d]	+ (2)	+ (5)[f]	+ (1)	+ (1)[f]	- (7)	+ (1)[f]	+ (1)[f]	NA (0)	+ (0)	- (7)	+ (1)	+ (0)[f]
Parental consanguinity	-	-	-	-	-	+	-	-	-	-	-	+
Heterozygous relatives												
Microhematuria[a]	- [Mo/56, Fa/59]	+ [D_3/27, Si/34], - [D_2/24, Br_2/45]	+ [D/17]		- [Si/NA]	+ [Br/NA]				NA	+ [Si/36]	
Proteinuria[a]	- [Mi/56,] NA [Fa/59]	- [D_2/24, Br_2/45, Si/49]	+ [D/17]		- [Si/NA]	NA				NA	NA	
CKD stage 2 or higher[a]	+ [Mo/56, Fa/59]	+ [Si/49], - [D_2/24, Br_2/45]	- [D/26]		- [Si/NA]	NA				NA	+ [Si/46]	
Hearing loss[a]	- [Mo/56, Fa/59]	- [D_2/24, Br_2/45]	- [D/26]		- [Si/NA]	NA				NA	+ [Si/58]	
Ultrastructural GBM abnormalities[a]	NA	NA	NA		NA	NA				NA	NA	
Homozygous or compound heterozygous relatives												
Microhematuria[a]		+ [Br_2/NA]		NA			NA					
Proteinuria[a]		NA		+ [Si/24]			+ [Si/NA]					
CKD stage 2 or higher[a]		HD [Br_2/18]		HD [Si/31]			HD [Si/19]					
Hearing loss[a]		+ [Br_2/7]		+ [Si/40]			+ [Si/9]					
Ultrastructural GBM abnormalities[a]		NA		+ [Si/24] §			NA					

CKD: chronic kidney disease. pCr: plasma creatinine level, in micromol/L; to convert micromoles/L (SI units) into mg/dL dividing by 88.4. eGFR: estimated glomerular filtration rate, in mL/min/1.73m^2. SNHL: sensorineural hearing loss. GBM: glomerular basement membrane. [a] In brackets, the age at diagnosis or at last screening, in years. [b] In brackets, the age at start of hemodialysis (HD), peritoneal dialysis (PD) or of kidney transplant (KTx). [c] In brackets, the age at kidney biopsy. [d] +: Family history of hematuria and/or end stage renal disease (ESRD); in parenthesis, the number of enrolled relatives who were confirmed to be heterozygotes; † Family history of ESRD. [e] In brackets, kinship / age at diagnosis or at last screening in the relative, in years. <: before the age indicated within brackets. Fa: father; Mo: mother; Br: brother; Si: sister; D: daughter; (if there are more than one relative of the same kinship, they are identified in subscript by the order of age). Hom: homozygous. Comp het: compound heterozygous. +: Clinical feature present. –: Clinical feature not present. § GBM changes observed in this patient: thinning, thickening, lamellation and electrondense bodies. ND: effect on splicing not determined by mRNA analysis. NA: data not available.

Supplementary table S5. Genotype and phenotype of Portuguese probands and their relatives with pathogenic heterozygous *COL4A3* or *COL4A4* mutations.

FAMILY	5/AS-1	16/AS-1	46/AS-1	54/AS-1	A3/4 - 4/1	A3/4 - 5/1	A3/4 - 6/1	8/AS-1	17/AS-2	22/AS-1	25/AS-1	37/AS-1	A3/4 - 2/1
GENOTYPE													
Gene	COL4A3	COL4A3	COL4A3	COL4A3	COL4A3	COL4A3	COL4A3	COL4A4	COL4A4	COL4A4	COL4A4	COL4A4	COL4A4
Mutation status	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het
Exon	21	21	21	21	21	28	21	20	30	9	47	38	Intron 2
Nucleotide change	c.1219G>C	c.1219G>C	c.1219G>C	c.1219G>C	c.1219G>C	c.2083G>A	c.1219G>C	c.1323_134Ddel18	c.2690G>A	c.568_569dupGG	c.4664_4665delCT	c.3506G>T	c.71+1G>A
Predicted effect on the protein	p.(Gly407Arg)	p.(Gly407Arg)	p.(Gly407Arg)	p.(Gly407Arg)	p.(Gly407Arg)	p.(Gly695Arg)	p.(Gly407Arg)	p.(Gly442_Pro447del)	p.(Gly897Glu)	p.(Asp191Glyfs*29)	p.(Ser1555*)	p.(Gly1169Val)	ND
Mutation type	missense	missense	missense	missense	missense	missense	missense	frameshift	missense	frameshift	frameshift	missense	splicing
PHENOTYPE AND ORIGIN													
Gender, [age at enrollment in years]	M, [50]	F, [27]	M, [42]	F, [28]	F, [71]	F, [44]	F, [54]	M, [26]	F, [29]	F, [51]	F, [25]	F, [45]	F, [42]
Birth district	Porto	Porto	Porto	Aveiro	Braga	Porto	Braga	Braga	Braga	Porto	Braga	Vila Real	Braga
Proband's clinical and histopathological phenotype													
History of kidney disease													
Macroscopic hematuria[a]	-	-	+ [25]	+ [17]	-	+ [39]	NA	-	+ [13]	-	-	-	- [42]
Microscopic hematuria[a]	+ [29]	+ [2]	+ [36]	+ [11]	+ [50]	+ [39]	+ [17]	+ [22]	+ [6]	+ [25]	+ [13]	+ [33]	+ [25]
Proteinuria[a]	+ [38]	- [29]	+ [36]	+ [12]	+ [50]	+ [39]	+ [20]	+ [22]	+ [13]	+ [25]	+ [20]	+ [33]	+ [25]
Hypertension[a]	+ [40]	- [18]	+ [NA]	-	+ [50]	- [39]	-	+ [21]	NA	+ [49]	- [25]	- [45]	- [35]
CKD stage 2 or higher[a]	+ [38]	- [29]	NA	- [27]	+ [60]	- [45]	+ [40]	+ [20]	+ [25]	+ [50]	- [25]	- [45]	+ [38]
pCr(eGFR) at diagnosis of CKD stage 2 or higher	1.4 (57)	---	NA	---	1.06 (53)	---	0.78 (82)	1.58 (56)	0.9 (76)	0.81 (75)	---	---	1.0 (62) [38]
Renal replacement therapy[b]	PD [47]	- [29]	- [42]	- [27]	- [71]	- [45]	- [54]	HD [23], PD [23]	- [29]	- [51]	- [25]	- [45]	- [42]
pCr(eGFR) at onset of renal replacement therapy	4.4 (14)	---	---	---	---	---	---	NA	---	---	---	---	---
Last available pCr(eGFR)[c,e]	NA	0.6 (119) [28]	NA	0.58 (125) [27]	1.68 (39) [70]	0.6 (111) [39]	0.8 (75) [54]	NA	0.9 (76) [25]	0.81 (75) [50]	0.4 (145) [25]	NA	1.0 (89) [40]
Hearing loss													
Self-noticed (subjective)[a]	+ [NA]	NA	+ [NA]	-	+ [NA]	- [43]	- [44]	NA	NA	NA	- [25]	- [45]	NA
SNHL, audiologically confirmed[a]	+ [44]	- [18], + [27]	+ [NA]	- [29]	+ [70]	- [43]	+ [44]	- [24]	+ [24]	+ [50]	- [25]	NA	- [37], + [39]
Ocular signs													
Anterior lenticonus[a]	- [50]	- [27]	NA	- [29]	- [70]	- [44]	NA	- [23]	- [29]	- [49]	- [24]	NA	- [40]
Maculopathy[a]	- [50]	- [27]	+ [NA]	- [29]	NA	- [44]	NA	+ [23]	- [29]	- [49]	- [24]	- [45]	- [40]
Cataracts[a]	- [50]	- [27]	NA	- [29]	- [70]	+ [44]	NA	- [23]	- [29]	- [49]	- [24]	NA	- [40]
Ultrastructural GBM abnormalities on kidney biopsy[c]	[40]		[41]	[12]			[41]		[21]		[23]	[34]	[28]
Thinning	+	NA	+	+	NA	NA	+	NA	+	NA	+	NA	+
Thickening	+	NA	+	NA	NA	NA	NA	NA	+	NA	NA	NA	NA
Lamellation	-	NA	+	NA	NA	NA	NA	NA	-	NA	NA	+	NA
Electrondense bodies	+	NA	NA	NA	NA	NA	NA	NA	+	NA	-	NA	NA
Family history[d]	+ (3)*	+ (1)	- (0)	+ (2)*	+ (0)*	+ (0)*	+ (1)	+ (2)	+ (1)	+ (1)*†	+ (2)*	+ (3)	+ (1)
Parental consanguinity	-	-	-	-	-	-	-	-	-	-	-	-	-
Heterozygous relatives													
Microhematuria[a]	+ [Br_2/45, SI/38]	+ [Fa/57]		+ [mGA/NA]			+ [S/21]	+ [Mo/45, SI/19]	- [Fa/61]	NA	+ [SI/18] - [Fa/51]	+ [D/24]; - [Mo/79]	+ [Fa/60]
Proteinuria[a]	+ [Br_2/18, Br_2/45, SI/38,]	NA		+ [Mo/18; mGA/NA]			+ [S/21]	+ [Mo/45, SI/19]	- [Fa/61]	NA	+ [SI/18] NA [i a]	+ [D/24]; - [Mo/79]	+ [Fa/60]
CKD stage 2 or higher[a]	+ [Br_2/36, Br_2/45]	NA		+ [Mo/NA; mGA/NA]			- [S/21]	+ [Mo/46]	- [Fa/61]	NA	- [Fa/ 51, SI/18]	- [Mo/78]	+ [Fa/60]
Hearing loss[a]	+ [Br_2/36, SI/38]	- [Fa/57]		- [Mo/NA; mGA/NA]			- [S/27]	-	NA	NA	NA	+ [mA/79] - [Mo/78]	+ [Fa/ 61]
Ultrastructural GBM abnormalities[c]	NA	NA		NA			NA	+ [Mo/45]§	NA	NA	NA	NA	NA

CKD: chronic kidney disease. pCr: plasma creatinine level, in micromol/L; to convert micromoles/L (SI units) into mg/dL divide by 88.4. eGFR: estimated glomerular filtration rate, in mL/min/1.73m^2. SNHL: sensorineural hearing loss. GBM: glomerular basement membrane. [a] In brackets, the age at diagnosis or at last screening, in years. [b] In brackets, the age at start of hemodialysis (HD) or peritoneal dialysis (PD). [c] In brackets, the age at kidney biopsy. [d] +: Family history of hematuria and/or end stage renal disease (ESRD); in parenthesis, the number of enrolled relatives who were confirmed to be heterozygotes; † Family history of ESRD; # clinical data from the mother was not available. [e] In brackets, kinship and age at diagnosis or at last screening in the relative, in years. <: Before the age indicated within brackets. Fa: father; Mo: mother; Br: brother; Si: sister; S: son; D: daughter; mA: maternal aunt; mGA: maternal grand-aunt; (if there are more than one relative of the same kinship, they are identified in subscript by the order of age). Het: Heterozygous. +: Clinical feature present. –: Clinical feature not present. § GBM changes observed in this patient: thinning and lamellation. ND: effect on splicing not determined by mRNA analysis. NA: data not available.

Chapter 4 Discussion

1. Aim 1: Molecular analysis of *COL4A5*, *COL4A4* and *COL4A3* in Portugal

1.1. The epidemiology of pathogenic COL4A5, COL4A4 and COL4A3 mutations in Portugal is different from other countries

The *COL4A5* molecular analysis was performed by direct sequencing and MLPA in 60 out of 65 probands (Supplementary figure S1 of manuscript 3). XLAS was genetically confirmed in 22 families (22/60; 37%), in whom pathogenic *COL4A5* mutations were identified. In the remaining five out of 65 probands, the inheritance pattern was not compatible with XLAS and direct sequencing of *COL4A3* and *COL4A4* was performed as the first tier molecular investigation. Pathogenic *COL4A3* or *COL4A4* mutations were detected in all five probands. Thirty-five probands without pathogenic *COL4A5* mutations were also studied by direct sequencing of *COL4A3* and/or *COL4A4,* and pathogenic mutations were detected in either one of the two genes in 20 of these families (25/40; 63%). Since the common pathogenic *COL4A3* missense mutation c.1219G>C (p.Gly407Arg) was identified by targeted mutation analysis in one proband, full mutational screening of *COL4A4* was avoided in his family.

1.1.1. High detection rate of pathogenic *COL4A5*, *COL4A4* and *COL4A3* mutations

Among 62 probands with clinical diagnosis or suspicion of AS, pathogenic *COL4A5*, *COL4A4* or *COL4A3* mutations were identified in 47 (76%). We have used PCR followed by direct sequencing as the first tier investigation in the genetic diagnosis of type IV collagen nephropathies. This laboratory approach is considered the gold standard for the analysis of genes with different mutation types, including for the detection of small mutations and the determination of breakpoints of large rearrangements (Sanger, Nicklen et al. 1977). PCR and direct sequencing of genomic DNA demonstrated a *COL4A5* mutation detection rate over 80% (Martin, Heiskari et al. 1998; Hertz, Thomassen et al. 2012). In this study, it enabled: (i) detection of point mutations in the coding region of *COL4A5*, *COL4A4* and *COL4A3*,

both in males and females; (ii) detection of large *COL4A5* deletions in males; (iii) identification of pathogenic variants in intron-exon boundary regions of the three genes.

Small *COL4A5* pathogenic mutations were previously identified using screening methods followed by confirmation of the mutation in genomic DNA, with single-stranded conformation polymorphism (SSCP) analysis being one of the most widely used (Hertz, Juncker et al. 2001). However, the relative insensitivity of some of these earlier mutation scanning methods is likely to be worsened when applied to XLAS, because of the large gene size, coupled with the unusually high GC content of most of the *COL4A5* exons and the sudden change in GC content at the exon-intron boundaries (King, Flinter et al. 2006). As the high GC content is typical of the exons coding for the proline- and hydroxyproline-rich amino acid triplet Gly-X-Y, which is a hallmark feature of collagens, the same cautionary thoughts regarding the analytical sensitivity of SSCP-based methods can be generalised to the molecular study of *COL4A3* and *COL4A4*.

A mutation detection rate of over 90% was achieved using direct sequencing of cDNA obtained from ectopically expressed *COL4A5* mRNA in peripheral blood leukocytes (Inoue, Nishio et al. 1999). However, since *COL4A5* mRNA is object of variant splicing in leukocytes (Guo, Van Damme et al. 1993), the results of these analyses require careful interpretation, particularly if no mutation is found, as the absence of an alternativelly spliced exon in the *COL4A5* mRNA of peripheral blood leucocytes, but expressed in skin or kidney cells, may hide a pathogenic mutation present in genomic DNA.

The expression of α5(IV) chains in EBM makes it possible to study cDNA obtained from cultured skin fibroblasts, with the advantages of performing immunohistochemistry for α5(IV) chain and to analyze the coding sequence in a compact form, being mutations subsequently confirmed in genomic DNA, with a high detection rate (Hertz, Juncker et al. 2008). In recent studies, analysis of *COL4A5* cDNA fragments from skin fibroblasts yielded a mutation detection rate of 83% and was particularly valuable for identification of cryptic splicing mutations (Wang, Zhao et al. 2012). However, this laboratory approach requires a skin biopsy, which is a

more invasive procedure, and for this reason it was not chosen as first tier screening method in our study.

Furthermore, MLPA was additionally used for deletion/duplication analysis of *COL4A5* (Hertz, Thomassen et al. 2012). In females, MLPA was shown to be effective in detecting large deletions/duplications as well as point mutations that coincide with the location of the probes (Uberbacher, Hyatt et al. 2004; Hertz, Juncker et al. 2008). Although not being an appropriate method to screen for unknown point mutations, in our cohort MLPA analysis confirmed the presence of a frameshift mutation resulting from deletion of a single base pair in family 1, and of a missense mutation in family 48, both affecting hybridization sites for the MLPA probes. Hence, while it has been specifically designed to identify deletions/duplications not readily detectable by sequence analysis, MLPA also allows the detection of small mutations that prevent effective hybridization of its probes to test samples.

Fifteen out of 62 probands did not carry any pathogenic *COL4A5*, *COL4A4* or *COL4A3* mutations (24%). The specific detection rate of a pathogenic *COL4A5* mutation among Portuguese families with clinical diagnosis or suspicion of AS was 37% (22/60). However, the commercial P191/P192 Alport MLPA assay does not contain probes for three of the 51 exons of *COL4A5*, leaving exons 8, 25 and 40 out of the analysis. Although isolated exon deletions are rare, DNA diagnostic laboratories should be aware of this limitation, which may be particularly relevant for families in whom the proband is a female and no pathogenic *COL4A5*, *COL4A4* or *COL4A3* mutations were identified. Of three probands in whom the mutational screening was limited to *COL4A5*, two were males (so deletion of each *COL4A5* exon was excluded by PCR), but the other was a female from a family in whom all affected males were deceased. Among the 15 probands without a pathogenic *COL4A5, COL4A4* and *COL4A3* mutation, four were males and 11 were females. One female did not have family history of hematuria, CKD or deafness; three females had only affected female relatives; five females reported family history of affected male relatives but none of them was alive; the remaining two female probands, who were enrolled through dialysis clinics, had affected male relative(s) but none of them was referred for molecular analysis. However, the major explanation for the lower

than expected proportion of families with genetically confirmed XLAS was the detection of pathogenic *COL4A4* or *COL4A3* mutations in 25 out of 40 (63%) of the families. This mutation detection rate is an estimate since the molecular study of *COL4A4* and *COL4A3* did not include deletion/duplication analysis.

Several other reasons may have concurred for not identifying pathogenic *COL4A5*, *COL4A4* or *COL4A3* mutations in 24% of the enrolled families (Flinter and Plant 1998): (i) occurrence of pathogenic mutations in non-coding regions of *COL4A5*, *COL4A4* and *COL4A3*, namely in the promoter region and in the intronic regions which were not screened; (ii) the laboratory methods employed did not enable detection of complex rearrangements (e.g. a karyotype was not performed as part of the molecular study protocol) (Hertz, Persson et al. 2005); (iii) occurrence of pathogenic mutations in other genomic locations (e.g. mutations in other genes); and (iv) enrollment of patients with diseases that have clinical manifestations in common with AS (Deltas, Pierides et al. 2013). Of note, in our cohort, the 5' UTR and the promoter region of *COL4A5* were sequenced in three families with multiexon *COL4A5* deletions, excluding point mutations or a deletion of this region (manuscript 2).

1.1.2. The prevalence of pathogenic *COL4A5* mutations is similar to that of pathogenic *COL4A4 and COL4A3* mutations in Portugal

We studied a genetically heterogeneous group of patients with clinical diagnosis or suspicion of AS: (i) 65 individuals, belonging to 22 families, had a pathogenic *COL4A5* mutation; and (ii) 59 individuals, from 25 families carried one or two pathogenic *COL4A3* or *COL4A4* mutations. The prevalences of X-linked and autosomal AS in our cohort is in contrast with the expected occurrence of X-linked inheritance in 80-85% of families and of autosomal inheritance in 15-20% of families. Indeed, according to publicly available data from Human Gene Mutation Database® (http://www.hgmd.cf.ac.uk; last accessed on March 9, 2014), 755 pathogenic mutations were published in *COL4A5* (82.2%), 67 in *COL4A4* (7.3%) and 97 in *COL4A3* (10.5%). However, a recent study of 87 Italian families with clinical suspicion of AS, which used an unbiased NGS protocol to simultaneously screen the *COL4A3*, *COL4A4* and *COL4A5* genes for disease-causing mutations, showed that,

among the 48 families with pathogenic mutations identified in one of the three genes, 65% had a *COL4A5* mutation and the remaining had *COL4A3* or *COL4A4* mutations (Fallerini, Dosa et al. 2013). In our study, the proportion of families with a pathogenic *COL4A5* mutation (22/47; 47%) was similar to the proportion of families with at least one pathogenic *COL4A3* or *COL4A4* mutation (25/47; 53%). We hypothesize that the prevalence of XLAS might have been underestimated in the Portuguese families because of the following possible reasons: (i) the use of broader inclusion criteria (less than three classic diagnostic criteria of AS) enabled the enrollment of probands with less severe renal manifestations and/or without extra-renal features and might have favoured the recruitement of patients with TBMN caused by heterozygous for pathogenic *COL4A3* or *COL4A4* mutations (Fallerini, Dosa et al. 2013); (ii) Portuguese patients with genetically confirmed XLAS who have been previously enrolled in ECASCA study were not included in this cohort; (iii) the non-participation in this study of some major Portuguese kidney transplantation centers might have also caused a selection bias against patients with more severe presentations of AS. Interestingly, however, if the 13 families in whom a single pathogenic mutation was detected in *COL4A3* or *COL4A4* are excluded from the analyses (assuming a second pathogenic *COL4A3* or *COL4A4* mutation would not have occured in the proband and that these families would have TBMN), the proportion of families with XLAS rises to 65% (22/34), which is still significantly lower than the expected 80-85% prevalence, but which is in agreement with the reported findings in the Italian population (Fallerini, Dosa et al. 2013).

Furthermore, among the Portuguese families with a clinical diagnosis of AS or TBMN and pathogenic *COL4A5*, *COL4A4* or *COL4A3* mutations, the proportion of probands with heterozygous *COL4A3* or *COL4A4* mutations (13/47; 28%) was similar to the proportion of probands with molecularly confirmed ARAS (12/47; 26%). Contrastingly, in the Italian study, the proportion of families with pathogenic heterozygous mutations in *COL4A3* or *COL4A4* was higher (31%) than the proportion of families with molecularly confirmed ARAS (4%) (Fallerini, Dosa et al. 2013). Although the latter findings are not concordant with our own results, in both studies the proportion of probands with heterozygous pathogenic *COL4A3* or *COL4A4* mutations was significantly above the expected 5%. Possible reasons for

the finding of a higher than expected proportion of heterozygous patients with pathogenic *COL4A3* or *COL4A4* mutations are the following: (i) the presence of a second pathogenic mutation in either *COL4A3* or *COL4A4* may have been missed, either because of the inability of direct sequencing to detect specific types of mutations in these genes (e.g. exon or multiexon deletions and/or duplications) or because the mutations occurred in non-scanned gene regions (e.g. promotor and deep intronic mutations); (ii) individuals with heterozygous *COL4A3* or *COL4A4* mutation may be more susceptible to the effects of major risk factors for CKD (e.g. hypertension or diabetes *mellitus*), which are highly prevalent in the population, or to the effects of other genetic variants and polymorphisms that may modulate the risk of progressive kidney disease in the population; (iii) the chance occurrence in the same patient of another genetic or non-genetic kidney disease, unrelated to GBM collagen IV.

1.1.3. Documentation of pathogenic *COL4A5*, *COL4A4* and *COL4A3* mutations

It is important to emphasize that this study not only confirmed the diagnosis in the probands, but also promoted the combined clinical and molecular study of their family members for precise genetic conselling. Regarding the molecular characterization of the 156 affected and unnafected individuals, from the 47 families who were found to carry pathogenic *COL4A5*, *COL4A4* or *COL4A3* mutations, the following results were obtained:

- Among 22 families with pathogenic *COL4A5* mutations (Table I of manuscript 1):
    - 30 patients were hemizygotes;
    - 35 patients were heterozygotes;
    - 23 participants did not carry the pathogenic *COL4A5* mutation of the family.
- Among 17 families with pathogenic *COL4A3* mutations (Table I of manuscript 3):
    - 13 patients were homozygotes or compound heterozygotes;
    - 27 patients were heterozygotes;
    - 4 participants didn’t carry the pathogenic *COL4A3* mutation of the family.
- Among 8 families with pathogenic *COL4A4* mutations (Table I of manuscript 3):
    - 2 patients were homozygotes;
    - 17 patients were heterozygotes;

- 5 participants did not carry the pathogenic *COL4A4* mutation of the family.

The parents of probands were clinically and molecularly evaluated in 23 families (23/47; 49%). Mutations were confirmed to be *de novo* in 2 out of 13 families with pathogenic *COL4A5* mutations (15%). *De novo COL4A5* mutations were identified in a female with a large *COL4A5* deletion encompassing exons 43 to 45, which is predicted to lead to a truncated protein; and in a male with a previously reported missense mutation c.2633G>T (p.Gly878Val) (Plant, Green et al. 1999). In two of the probands carrying apparently homozygous pathogenic *COL4A3* or *COL4A4* mutations, parental screening confirmed the heterozygous condition of both parents. In eight families, the proband and one of the proband's parents carried a same pathogenic *COL4A3* and *COL4A4* mutation, in heterozygosity. Among the 24 families in which parents of the proband were not evaluated, both parents were deceased in six families with pathogenic *COL4A3* or *COL4A4* mutations (6/47; 12.8%) and the at-risk mother of the proband was already deceased in four families carrying a pathogenic *COL4A5* mutation (4/47; 8.5%). In 12 families, at-risk parents of the proband were alive, but were not enrolled in the study (12/47; 25.5%). In two families (2/47; 4.2%), the available family history did not include information on the proband's parents. Among four of the ten probands with an apparently homozygous or compound heterozygous pathogenic *COL4A3* or *COL4A4* mutations of whom parents were not enrolled, at least one other adult first degree relative was found to carry one mutation, in heterozygosity. Further clinical and molecular analyses will need to be persued in at-risk relatives in the remaining families to prove the occurrence of the pathogenic mutations in distinct alleles. Of note, the parents of the genetic probands were consanguineous in only three out of the 65 families.

1.1.4. Pathogenicity of the novel *COL4A5*, *COL4A4* and *COL4A3* variants

The results of this study extended the spectrum of known disease-causing mutations in the three genes, leading to the identification of 29 novel pathogenic mutations: 12 in *COL4A5*, 10 in *COL4A3*, and 7 in *COL4A4*. Our data are also in agreement with the notion that collagen IV-related nephropathies do not result from digenic mutations (Heidet, Arrondel et al. 2001), as none of the 34 probands who were screened for both *COL4A3* and *COL4A4* mutations was a double heterozygous.

1.1.4.1. According to the mutation type

Similarly to what was observed in other populations, collagen IV-related nephropathies also show a high degree of allelic heterogeneity in Portuguese families, with a considerable proportion of private mutations. As expected, missense mutations were the most frequent type of mutation detected in the three genes, and large deletions accounted for 19% of all *COL4A5* mutations (Table IX). The proportion of large *COL4A5* deletions identified in the Portuguese families is consistent with previous reports on the genetic pathology of XLAS (Jais, Knebelmann et al. 2000).

Respectively 89.9% (196/218), 78.6% (22/28) and 87.1% (27/31) of the pathogenic *COL4A5*, *COL4A4* and *COL4A3* missense mutations reported at the HGMD® database (last accessed on September 19, 2013), occurred within the collagenous domain of the protein; of those, respectively 95.4% (187/196), 90.9% (20/22) and 96.3% (26/27) were glycine substitutions. In keeping with those data, all pathogenic *COL4A5*, *COL4A4* and *COL4A3* missense mutations detected in our cohort were located in the collagenous domains and substituted glycine residues. As pointed out by several investigators (Wang, Ding et al. 2004; King, Flinter et al. 2006), the predicted substitution of glycine residues in the collagenous domain of these proteins enables, in most cases, to classify the underlying *COL4A5, COL4A4* or *COL4A3* mutations as pathogenic. This is because glycine substitutions in collagenous domains affect the triple helical folding of the alpha chains, thereby disturbing the quaternary structure of the corresponding collagen.

Table IX. Comparison of the types of pathogenic *COL4A5*, *COL4A4* and *COL4A3* mutations deposited in HGMD® and detected in this study (last accessed on September 19, 2013).

Mutation type	*Number of mutations in HGMD®*						*Number of mutations in this study*					
	COL4A5		*COL4A4*		*COL4A3*		*COL4A5*		*COL4A4*		*COL4A3*	
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Missense	218	41.7	28	50.0	31	44.3	7	33.3	2	25.0	4	30.8
Nonsense	35	6.7	5	8.9	11	15.7	2	9.5	0	0.0	2	15.4
Splicing	88	16.8	7	12.5	11	15.7	6	28.6	2	25.0	2	15.4
Small deletions	83	15.9	11	19.6	10	14.3	2	9.5	2	25.0	3	23.0
Small insertions	29	5.5	3	5.4	4	5.7	0	0.0	2	25.0	2	15.4
Small indels	4	0.8	1	1.8	1	1.4	0	0.0	0	0.0	0	0.0
Gross deletions	58	11.1	1	1.8	2	2.9	4	19.0	0	0.0	0	0.0
Gross insertions /duplications	4	0.8	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Complex rearrangements	4	0.8	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Total	523	100	56	100	70	100	21	100	8	100	14	100

1.1.4.2. According to the protein domain

Pathogenic variants occurred dispersed throughout the coding region of the gene, with the number of point mutations in a specific domain being proportional to the size of the domain of the alpha chain. In *COL4A5*, the majority of the pathogenic point mutations (n=17) detected in this study were located in the collagenous domain (15/17; 88.2%), while a minority was identified in the NC1 domain or C-terminal domain (2/17; 11.8%). Similar findings were observed in *COL4A4*, with the majority of mutations being identified in the large collagenous domain (6/8; 75.0%), while one mutation was located in 7S domain or N-terminal domain (12.5%) and another mutation in the NC1 domain (12.5%). Finally, in *COL4A3*, all pathogenic point mutations were detected in the collagenous domain of the protein. These results overlap with data reported at the HGMD® (last accessed on September 19, 2013), showing that the majority of the pathogenic mutations identified in each of the three genes occurred in the collagenous domain of the corresponding alpha chain (Table X). Interestingly, pathogenic point mutations were dispersed throughout the coding

region of the gene, and the proportion of pathogenic mutations *per* protein domain correlates with the relative length of the domain within the protein.

Table X. Comparison of the location in different protein domains of pathogenic *COL4A5*, *COL4A4* and *COL4A3* mutations deposited in HGMD® (last accessed on September 19, 2013) with those detected in this study.

	Protein domain	***Domain size***		***Number of mutations in HGMD***		***Number of mutations in this study***	
		codons	**% of protein**	**n**	**%**	**n**	**%**
COL4A5	N-terminal domain	1 - 26	1.5	5	1.1	0	0.0
	Collagenous domain	27 - 1456	84.9	402	88.7	15	88.2
	C-terminal domain	1457 - 1685	13.6	46	10.2	2	11.8
COL4A4	N-terminal domain	1 - 38	2.2	2	3.6	1	12.5
	Collagenous domain	39 - 1459	84.1	46	83.6	6	75.0
	C-terminal domain	1460 - 1690	13.7	7	12.7	1	12.5
COL4A3	N-terminal domain	1 - 28	1.7	1	1.5	0	0.0
	Collagenous domain	29 - 1438	84.4	55	80.9	13	100.0
	C-terminal domain	1439 - 1670	13.9	12	17.6	0	0.0

Moreover, when comparing the number of base pairs per exon with the number of reported pathogenic point mutations per exon in *COL4A5*, *COL4A4* and *COL4A3*, the two distributions seem to superimpose (Figures 5 and 6). Because the majority of the larger exons (approximately 100 bp) are located between the exons 19 and 51, and the majority of the reported pathogenic mutations are also in that region, the molecular analysis of the *COL4A5* gene in every index-case was started from the 3' end of the gene. A further reason for this decision was the fact that the NC1 domain, which is critical for the assembly of collagen heterotrimers, is encoded by *COL4A5* exons 47 to 51 and, therefore, mutations affecting these exons are very likely pathogenic.

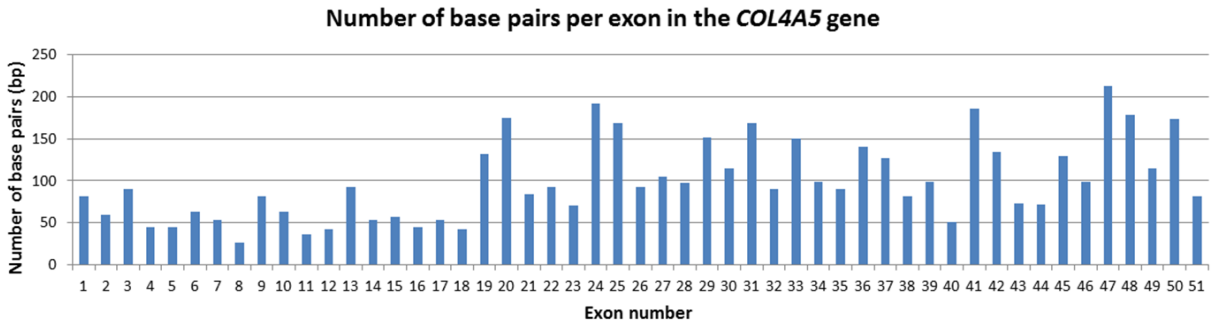


Figure 5. Number of base pairs per exon, in the *COL4A5* gene.

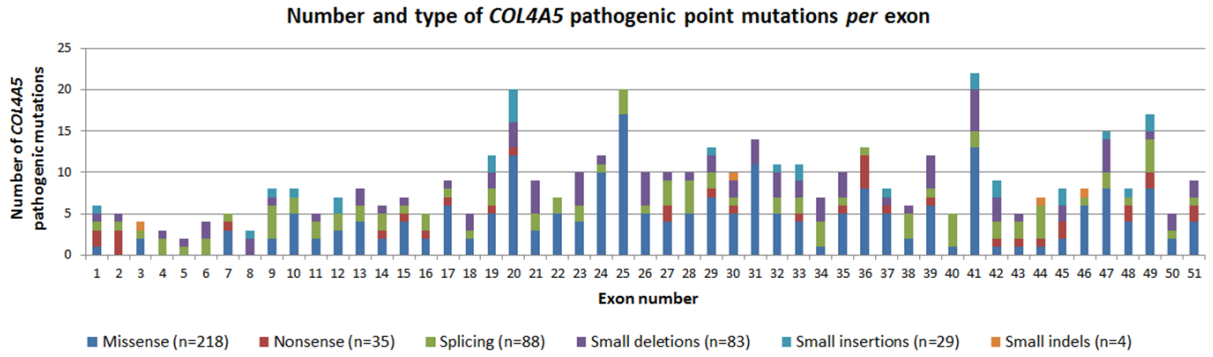


Figure 6. Number and type of pathogenic point mutations *per* exon, in the *COL4A5* gene.

1.1.4.3. Recurrent location of pathogenic mutations

Although there are no mutational hotspots described in the *COL4A5* gene, our study showed recurrent mutation locations in families with XLAS: (i) in exon 47, which harbored two distinct point mutations – c.4342G>C, p.(Gly1448Arg), and c.4444C>T, p.(Gln1482*); (ii) adjacent to or involving the acceptor splice site of intron 30, namely two distinct deletions – c.2510_2554del (45bp) and c.2510(-1)_2525del (15bp); (iii) and in intron 1, which included one of the breakpoints of two distinct multiexonic deletions. Furthermore, the pathogenic p.(Gly1448Arg) mutation in exon 47 was identified in two apparently unrelated families, but the results of the haplotype analysis were suggestive that they are related by a common ancestor. Interestingly, a mutation at the same nucleotide position, but changing a glycine to serine, was originally reported in a 39-year-old Chinese male diagnosed with AS (Wang, Wang et al. 2005). With 213 base pairs, exon 47 is the largest of the *COL4A5* exons, partially

encoding both the collagenous and the NC1 domains; for this reason, it should be prioritized in the molecular analysis of *COL4A5* by direct sequencing.

Although we did not perform RNA studies to prove the pathogenicity of the two microdeletions identified in the boundary of intron 30 with exon 31, their presumed effects at the protein level are severe enough to justify classification as disease-causing. Indeed, a much smaller c.2510delG frameshifting deletion had already been described in a male diagnosed with AS (Flinter, Cameron et al. 1988) and was classified as pathogenic (King, Flinter et al. 2006). Molecular genetics studies on patient's cell lines (e.g. skin fibroblasts), including with reverse transcriptase-PCR and quantitative real-time PCR, would be of help to characterize the consequences of those microdeletions at the RNA level.

Regarding the two gross *COL4A5* deletions with proximal breakpoints located in intron 1, both occurred within LINE1 repeats, more precisely 9.5 kb proximal to exon 2 (chrX:107,773,340) at the end of a MA3-LINE1 repeat, in the ATS-DL family (*COL4A5* deletion of exons 2 to 51), and in a 700 bp region 24.5 kb distal to *COL4A5* exon 1 (chrX:107,707,634-107,708,329) in a ME1-LINE1 repeat, in the family with the *COL4A5* deletion of exons 2 to 29. The occurrence of microdeletion breakpoints within LINE1 repeats in intron 1 of *COL4A5* has been previously reported, which may predispose to non-allelic homologous recombination (NAHR), if significant homology exists between the subfamilies of LINE1 repeats in both, proximal and distal, breakpoints (Segal, Peissel et al. 1999).

The previously reported pathogenic *COL4A3* missense mutation c.1219G>C p.(Gly407Arg), was identified in eight apparently unrelated families, suggesting that it is a common cause of collagen IV-related nephropathies in Portugal. This mutation was originally described in compound heterozygosity in a family with two affected siblings, who both needed to start dialysis treatment early in the third decade of life (Heidet, Arrondel et al. 2001). As the clinical report does not give details about the renal phenotype of their mother, who was heterozygous for the same mutation, it may be assumed that, at least, she was not severely affected. However, our study demonstrates that heterozygosity for this particular *COL4A3* mutation is associated with a broad spectrum of clinical phenotypes, including TBMN with and without progression to ESRD.

1.1.4.4. According to the population screening and bioinformatic tools

The splice site mutation c.1339+6C>G, detected in the male proband of family 3, is a novel variant, which occurs within a nucleotide sequence that is immediately adjacent to the highly conserved splice junction consensus sequence. Pathogenic mutations were not identified in *COL4A3* and *COL4A4* genes in the proband, after performing direct sequencing. Due to the fact that the proband's mother, already deceased, also had reached ESRD and had severe high tone SNHL, an autosomal inheritance pattern was less likely than the X-linked pattern. This previously unpublished variant segregates with the disease in the family, as it was also detected in heterozygosity in the proband's sister who has XLAS, but not in a maternal aunt of the proband, who didn't have microscopic hematuria, high tone SNHL, nor ocular signs of AS. This *COL4A5* variant was also not identified in 150 X-chromosomes of healthy controls, nor in 85 X-chromosomes from Portuguese AS probands. A new donnor splice site is predicted to occur in the position of the variant, within intron 20 of *COL4A5*, according to NNSplice (http://www.fruitfly.org/seq_tools/splice.html, last accessed on October 26, 2013) and Human Splicing Finder (http://www.umd.be/HSF/, last accessed on October 26, 2013). The consistency of this bioinformatic tool results was used to predict the pathogenic effect of genomic DNA variants into proteins. This option was justified given the accuracy of the publicly available bioinformatic programs in the prediction of the pathogenicity of the change based upon the effect of the change on the specific functional domains of the protein (Fairbrother, Yeh et al. 2002; Calabrese, Capriotti et al. 2009; Adzhubei, Schmidt et al. 2010; Thusberg, Olatubosun et al. 2011; Adzhubei, Jordan et al. 2013). Additionally, the bioinformatic approach has overcome the refusal of collaboration of patients for biopsy of skin and saved time and financial resources required for the implementation and accomplishment of the analysis of mRNA.

2. Aim 2: Clinical characterization of Portuguese patients

2.1. Phenotype of the 65 genetic probands

2.1.1. Microscopic hematuria and thinning of the GBM were hallmarks

Clinical information was available on 166 participants regarding renal function, renal structure, audiologic and ophthalmologic evaluation. The evaluation of the four typical clinical criteria of AS was completed in 24 participants (24/166; 14%); three of these evaluations were performed in 56 participants (56/166; 34%); the evaluation of two parameters was achieved in 42 participants (42/166; 25%); and in 44 participants only one out of the four parameters was investigated or reported in this study (44/166; 27%). Clinical evaluation information was not available for 12 individuals (five autosomal recessive AS carriers and seven non-carrier unaffected family members).

Regarding the 65 probands, the median age at enrollment was 38 years (interquartile range 20 years) (Table XI). Only one male proband was enrolled before reaching 18 years old, but he had manifested microscopic hematuria and hearing loss and was known to carry a pathogenic *COL4A5* mutation. Interestingly, besides microscopic hematuria being a renal impairment sign common to the vast majority of index cases (98%), the occurrence of thin GBM was observed in all kidney biopsies (100%), which were performed at a median age of 21 years old (adult age). Although an excess of males was expected due to the occurrence of XLAS in approximately 80-85% of the families and the increased severity of the disease in male patients with XLAS, the frequency of male and female patients among probands was similar, which prompted us to persue the molecular study of the *COL4A3* and *COL4A4* genes.

Table XI. Descriptive characteristics of the probands (n=65).

	Index cases (n=65)	
		N
Phenotype		
Male gender (%)	52.3	34/65
Age at enrollment (median (interquartile range)	38 (20)	65
Renal anomalies		
History of macroscopic hematuria (%) Age at diagnosis (median (interquartile range))	45.8 6 (10.5)	22/48 17
History of microscopic hematuria (%) Age at diagnosis (median (interquartile range))	98 17 (24)	50/51 43
History of proteinuria (%) Age at diagnosis (median (interquartile range))	98.2 14.2 (18)	54/55 49
History of hypertension (%) Age at diagnosis (median (interquartile range))	66.7 25 (20)	40/60 29
History of CKD stage 2 or higher (%) Age at diagnosis (median (interquartile range)) eGFR at diagnosis (mean (SD))	75.4 24.5 (20.5) 56.5 (25.5)	46/61 38 24
History of renal replacement therapy (%) Age at onset (median (interquartile range)) eGFR at onset (median (interquartile range))	61.9 23 (17.25) 9.5 (6)	39/63 38 14
Hearing loss		
Self-noticed or subjective (%) Age at self-noticed or subjective (median (interquartile range))	69.8 14(15.75)	44/63 24
Audiogram (%) Age at diagnosis of hearing loss by audiogram (median (interquartile range))	93.9 32 (24.5)	46/49 29
Ocular anomalies		
Anterior Lenticonus (%) Age at diagnosis (median (interquartile range))	19.5 27 (17)	8/41 6
Maculopathy (%) Age at diagnosis (median (interquartile range))	31.8 26 (20.75)	14/44 10
Cataracts (%) Age at diagnosis (median (interquartile range))	25 34 (25)	9/36 7
Ultrastuctural anomalies Age at diagnosis (median (interquartile range))	 20.5 (16)	 26
Thinning (%)	100	23/23
Thickening (%)	94.4	17/18
Lamellation (%)	73.7	14/19
Electrondense bodies (%)	89.5	17/19
Leiomyomatosis (%)	1.5	1/65
Intellectual disability and facial dysmorphy (%)	0	0/65
Family History		
Family history of hematuria (%)	78.4	40/51
Family history of CKD (%)	79	49/62
Family history of hearing loss (%)	67.2	39/58
Family history of ocular abnormalities (%)	42.1	8/19

SD: standard deviation. Age is expressed in years; eGFR is expressed in ml/min/1.73m^2.

Historically, affected males have been preferably selected as genetic probands of AS families, with the aim of optimizing the mutational analysis of the *COL4A5* gene due to the interpretation of sequencing results (Plant, Green et al. 1999). In our study, all patients with suspected or clinical diagnosis of collagen IV-related nephropathy, irrespective of their gender, were accepted as genetic probands. The current availability of molecular analysis techniques that make possible the diagnosis of the different types of variants in the *COL4A5* gene, both in males (DNA amplification by PCR followed by direct sequencing) and females (DNA amplification by PCR followed by direct sequencing and MLPA), also allowed the gender-unbiased selection of probands. By assessing the information provided in the clinical protocol of this study, several other reasons were found for the enrollment of females as genetic probands, including: (i) absence of family history of AS manifestations (n = 5); (ii) males with suspected or clinical diagnosis of AS were deceased (n = 8); (iii) suspected or clinical diagnosis of AS only in females (n = 6); and (iv) males and females equally affected, i.e., renal disease with similar severity in both genders (n = 2). In a minority of genetic probands, the reason that led to the recruitment of a female, instead of a male relative is not traceable (n = 5).

In families where a pathogenic mutation was identified in the *COL4A5*, *COL4A4* or *COL4A3* genes, the probable reasons for the enrollment of a female proband are summarized in Table XII. This distribution shows that 71% of the pathogenic mutations identified in female probands were in *COL4A3* or *COL4A4* (10/14), confirming an autosomal inheritance pattern in those families. Among the 22 families carrying pathogenic *COL4A5* mutations, only in four the genetic proband was a female. It may additionally be hypothesized that the putative genetic defect underlying the AS-like phenotype in the remaider families (i.e., those with no pathogenic mutation identified in *COL4A5*, *COL4A4* or *COL4A3*), results in an autosomal inheritance pattern.

Table XII. Probable reasons to study a female as a genetic proband, instead of a male, and result of the *COL4A5*, *COL4A4* and *COL4A3* molecular analysis (n=26).

	Absence of family history (n=5)	Deceased affected males (n=8)	Only affected females (n=6)	Males and females equally affected (n=2)	Unknown reason (n=5)
Pathogenic variant in *COL4A5*	1	0	1	0	2
Pathogenic variant in *COL4A3* or *COL4A4*	3	2	2	1	2
Without pathogenic variants in *COL4A5, COL4A4* or *COL4A3*	1	6	3	1	1

All 65 probands had microscopic hematuria and/or CKD (stage 2 or higher). Among the probands, 13 (20%) had one known diagnostic clinical criterion of AS, 28 (43%) had two known criteria, 19 (29%) had three known criteria and only a minority (8%) had four diagnosed clinical criteria. Among the 113 enrolled family members of these probands, 75 (66%) were enrolled due to family history of hematuria and/or CKD, 26 (23%) had two clinical diagnostic criteria of AS and 12 (11%) had three known diagnostic criteria of AS. Twenty family members did not have microscopic hematuria or CKD (stage 2 or higher): 17 unaffected (non-carrier) family members and three heterozygous relatives of probands with ARAS.

The retrospective, cross-sectional design and the relatively small size of the patient cohort, which is an inherent problem with rare diseases, are major limitations of this study. Careful review of all the available patient's clinical records allowed minimizing the recall bias. The age-dependent and the sex-dependent penetrance of renal, audiological and ocular manifestations also tried to be minimized by enrolling adult individuals and by specialist referral for screening of the ophthalmological and audiological manifestations of AS, which however was not systematically done. For these reasons, generalization of the genotype-phenotype correlations observed in our cohort, as well as of the phenotypic descriptions should be cautious.

3. Aim 3: A novel genotype-phenotype correlation among other correlations

3.1. *Deletion of the common* COL4A5/COL4A6 *promotor and 5' exons of* COL4A6 *is not needed for development of diffuse leiomyomatosis in X-linked Alport syndrome*

A large deletion involving *COL4A5* but not *COL4A6* was detected in three patients of the same family with ATS-DL. Previously, all patients reported with ATS-DL had deletions involving 5' exons of both *COL4A5* and *COL4A6* genes. Our results show that the deletion of the common promoter and the 5' exons of *COL4A6* is not required for the development of diffuse leiomyomatosis. The particularity of the novel mutation also resides in the fact that its breakpoints in intron 1 and beyond exon 51 in COL4A5 are the most 3' proximal and distal described to date in patients with ATS-DL. By performing high-resolution X-oligo-array, it was demonstrated that all other exons (and genes) on the entire X chromosome are neither deleted nor duplicated in the proband of the ATS-DL family. By cloning the junction sequence, a complex recombination between the *COL4A5* and the *COL4A6* genes was excluded and it was shown that the large *COL4A5* deletion identified in this family is the result of a single recombination event with no inversion or insertion at the deletion breakpoints. By direct sequencing the *COL4A5/COL4A6* common promoter and the entire first four exons of *COL4A6* in the X chromosome of the ATS-DL male proband, it was shown that these genomic regions are present and non-mutated. Additionally, by performing qPCR at several locations within *COL4A6* intron 2, it was shown that all the segments of *COL4A6* intron 2 that were probed by qPCR are present at the expected copy number. In conclusion, the deletion involving *COL4A5* but not *COL4A6* was the only genomic defect detected in this family. It is possible that the reason for leiomyoma predisposition in ATS-DL is aberrant regulation, for example due to the disruption of a regulatory domain within the large *COL4A5* intron 1 or due to the potential role of the 3' UTR of *COL4A5*.

3.2. An unknown family history of hematuria, CKD and hearing loss is not synonymous of a negative family history

When family history of different AS manifestations was analyzed separately, history of hematuria was not identified in 3 families (family 18, 26 and 41); history of CKD was not identified in 3 families (family 18, 41 and 50); and history of hearing loss was not identified in 3 families (18, 26 and 41). Two probands had unknown family history of hematuria, CKD and hearing loss (one from family 18 and one from family 41). In family 18, the negative family history of AS was confirmed by the negative results of the genetic screening of both parents of the female proband. In family 41, the apparently negative family history of AS was not confirmed by the molecular analysis, since the male proband had inherited the pathogenic mutation from his asymptomatic mother. Comprehensive clinical evaluation of the proband's mother, following the molecular genetic screening, showed that she had microhematuria, proteinuria and stage 3 CKD. Regarding family 26, in which only females were affected, hematuria and hearing loss were reported, but there was no family history of CKD. These observations demonstrate that all the at-risk relatives of patients with collagen IV-related nephropathies should be systematically screened for possible disease manifestations. Regarding family 50, in which the proband's mother was known to have hematuria and hearing loss, the absence of the mutation in *COL4A5* gene has shown that she does not have AS. Two consequences arise from this molecular result: 1) no other family members need to be screened for this mutation; 2) other causes for the mother health problems should be pursued. Taken together, these are examples of what may be faced during a consultation: 1) "true" negative family history, both phenotypically and molecularly; 2) "false" negative family history, when no manifestations of the syndrome are reported (e.g. due to variable expressivity of the disease), but the pathogenic mutation is detected in other family members; 3) "false" positive family history, when manifestations compatible with AS are reported, but the familial mutation is not present in the presumably affected relative.

3.3. Kidney biopsy with ultrastructure examination is recommended to clarify the etiology of urinary abnormalities

In total, photographs from 25 kidney biopsies performed prior to the molecular study of AS in 24 individuals, from 22 unrelated families, were collected and reviewed, corresponding to 34% of the families (22/65). Twenty three of those photographs were collected at Hospital São João, Porto, and two at Hospital Curry Cabral, Lisbon. From five patients – two index cases and three family members – only the clinical report was available. Three of these reports (including one from an index case) were excluded from the following analysis, either because they did not specify the GBM changes that led to the conclusion of a probable diagnosis of AS, or because the number of glomeruli was insufficient for a reliable diagnosis.

Ultrastructural findings suggestive of AS were present in 97% (29/30) of the kidney biopsies. All ten patients carrying hemizygous pathogenic *COL4A5* mutations or homozygous or compound heterozygous pathogenic *COL4A3* and *COL4A4* mutations manifested GBM thinning, thickening and lamellation with electrondense bodies surrounded by a halo, which are the typical GBM abnormalities of AS. Interestingly, four kidney biopsies from three females with XLAS, carrying heterozygous pathogenic *COL4A5* mutations, also showed the four typical ultrastructural signs of AS. Among four patients with heterozygous pathogenic *COL4A3* and *COL4A4* mutations, GBM thinning was observed in all four biopsies, GBM thickening and lamellation was found in 2/3 of the biopsies and electron-dense bodies were present in 2/2 of the biopsies. Among kidney biopsies of eight patients in whom pathogenic *COL4A3*, *COL4A4* and *COL4A5* mutations were not identified, GBM thinning was found in all biopsies, thickened GBM in 7/8 biopsies, lamellation in 2/8 biopsies (two patients from the same family) and electron dense bodies in 4/8, but none on these biopsies showed the characteristic electron-dense bodies surrounded by an electrolucent halo. The only kidney biopsy on which none of the typical ultrastructural signs of AS was found was from a patient who had a negative *COL4A5* genetic analysis and in whom the molecular study of *COL4A3* and *COL4A4* was decided not to be performed due to absence of typical ultrastructural criteria of AS. Moreover, immunohistochemistry analysis of the GBM from the proband of family

51 showed that both expression of α3(IV) and α5(IV) chains are present in the glomeruli of the kidney, supporting the absence of a pathogenic mutation in *COL4A3*, *COL4A4* and *COL4A5*.

3.4. Skin biopsy is helpful in determining the inheritance pattern

Immunohistochemical analysis of the α5 chain of type IV collagen in the EBM from nine patients with clinical diagnosis of AS who underwent skin biopsy showed: (i) complete absence of staining of EBM in the proband and his brother from family 15, both carrying the pathogenic *COL4A5* frameshift mutation c.590delC, p.(Pro197Glnfs*6); (ii) complete absence of staining of EBM in the proband from family 57, carrying the pathogenic *COL4A5* missense mutation c.1718G>A, p.(Gly573Asp); (iii) discontinuous staining of EBM in five out of seven females from family 57, carrying the same heterozygous pathogenic missense mutation in *COL4A5*. These results suggest that immunohistochemical analysis of the α5 chain of type IV collagen in EBM correlates better with the genotype in males than in females with XLAS.

3.5. The mutation status influences the severity of the phenotype

When the severity of the clinical phenotypes was compared according to the patients' mutation status – males versus females with XLAS and patients with ARAS versus patients with TBMN – consistent statistically significant differences were found between the groups under comparison. History of macroscopic hematuria, proteinuria, progressive CRF needing RRT and of extra-renal manifestations was more prevalent in males than in females with XLAS (Table II of manuscript 1). Nonetheless, CKD stage 2 or higher was diagnosed in 62% of the heterozygous females, not rarely at young adult age. Likewise, a higher frequency of renal, audiologic and ocular manifestations was observed in patients with homozygous or compound heterozygous pathogenic mutations in *COL4A3* or *COL4A4*, when compared with the apparently heterozygous patients (Table V of manuscript 3). Noteworthy, while diagnosis of CKD stage 2 or higher was performed at a median age of 20 years in patients carrying two pathogenic mutations in *COL4A3* or

COL4A4, it was performed medianly in the fourth decade of life in apparent heterozygote patients. Interestingly, the occurrence of microscopic hematuria was not statistically different among groups with different mutation status.

3.6. The type of mutation influences the severity of the phenotype

Male patients with truncating pathogenic *COL4A5* mutations reported higher proportion and younger ages at diagnosis of renal and extra-renal manifestations than patients with non-truncating mutations, although these results were not statistically significant (Supplementary table S11 of manuscript 1). Patients with hemizygous truncating *COL4A5* mutations were significantly younger when they underwent kidney biopsy for diagnosis of GBM ultrastructural changes and when tonal audiogram was performed. The need of undergoing these diagnostic investigations at younger ages may also indicate an earlier age of expression of renal and audiologic signs. Female patients with truncating pathogenic *COL4A5* mutations developed more often CKD (≥ stage 2) and hearing loss than females with non-truncating mutations (Supplementary table S12 of manuscript 1). Tonal audiogram was also performed at a significantly younger age.

3.7. Deliniation of the natural history of the disease in our sample

When comparing the phenotype of patients with hemizygous pathogenic *COL4A5* mutations and patients with homozygous or compound heterozygous *COL4A3* or *COL4A4* mutations, no significant differences were found (Table XIII). This finding suggests that it is feasible to use the diagnostic criteria defined for XLAS by Flinter *et al.* (1988) to clinically diagnose patients with ARAS. When comparing the phenotype of patients carrying heterozygous *COL4A5* and *COL4A3/COL4A4* mutations, differences were not statistically significant, except for the proportion of patients who reported microscopic hematuria, which was statistically higher in the group of heterozygotes with *COL4A5* mutations than in the group of heterozygotes with *COL4A3* and *COL4A4* mutations (Table XIV).

Since the overall phenotype of males with XLAS did not vary significantly from patients with ARAS, and the phenotype of females with XLAS did not deviate

considerably from patients with TMBN, these two larger groups of patients were assembled with the purpose of comparing the natural history of the disease in patients with hemizygous pathogenic *COL4A5* mutations, or homozygous or compound heterozygous pathogenic *COL4A3* and *COL4A4* mutations, with subjects with heterozygous pathogenic *COL4A3*, *COL4A4* and *COL4A5* mutations (Table XV). Microscopic hematuria was confirmed to be the hallmark of collagen IV-related nephropathies, as it manifested in both groups of patients, besides being diagnosed at significantly different median ages in both groups. Typical ultrastructural changes of the GBM were also observed at a similar proportion in both groups of patients. Altogether, these clinical findings may suggest a collagen IV-related nephropathy and may be indicative of the molecular analysis of the *COL4A3*, *COL4A4* or *COL4A5* genes.

The proportion of all other renal manifestations and all extra-renal manifestations was significantly higher in patients with a hemizygous pathogenic *COL4A5* mutation and homozygous or compound heterozygous pathogenic *COL4A3* or *COL4A4* mutations than in patients with a heterozygous pathogenic mutation in any of the three genes. When present, ocular manifestations corresponded to those previously reported. Diagnosis of microscopic hematuria, hypertension, CKD (≥ stage 2), characteristic GBM ultrastructural changes and hearing loss was made at a significantly younger age in the group of male patients with XLAS and patients with ARAS.

Table XIII. Comparison of clinical characteristics between male patients with pathogenic *COL4A5* mutations and patients with pathogenic homozygous and compound heterozygous *COL4A3* or *COL4A4* mutations.

	Male XLAS (n=30)		ARAS (n=15)		p-value
		N		N	
Phenotype					
Index-cases (%)	60.0	18/30	80.0	12/15	0.180
Male gender (%)	100.0	30/30	53.3	8/15	0.000
Age at enrollment (median (interquartile range))	33.5 (18.5)	30	42 (17)	15	0.123
Renal abnormalities					
History of macroscopic hematuria (%) Age at diagnosis (median (interquartile range))	61.9 5 (4)	13/21 11	54.5 8 (15*)	6/11 3	0.721 0.156
History of microscopic hematuria (%) Age at diagnosis (median (interquartile range))	100.0 12.5 (20.3)	25/25 22	100.0 12 (14.3)	9/9 6	a 0.625
History of proteinuria (%) Age at diagnosis (median (interquartile range))	96.2 18 (22.5)	25/26 21	100.0 16 (13.5)	12/12 9	1.000 0.217
History of hypertension (%) Age at diagnosis (median (interquartile range))	73.1 20 (11.5)	19/26 12	92.3 18 (14.8)	12/13 8	0.229 0.324
History of CKD stage 2 or higher (%) Age at diagnosis (median (interquartile range)) eGFR at diagnosis (median (interquartile range))	86.2 20.5 (10) 40 (35)	25/29 16 13	100.0 20 (10) 59 (15*)	14/14 9 3	0.286 0.563 0.092
History of renal replacement therapy (%) Age at onset (median (interquartile range)) eGFR at onset (median (interquartile range))	76.7 23 (19.5) 9 (6)	23/30 21 12	100.0 23 (12) 9.5 (3*)	15/15 15 2	0.077 0.301 0.783
GBM ultrastuctural changes Age at diagnosis (median (interquartile range))	 20 (14.5)	 6	 17.5 (7.8)	 6	 0.235
Thinning (%)	100.0	6/6	100.0	5/5	a
Thickening (%)	100.0	6/6	100.0	4/4	a
Lamellation (%)	100.0	6/6	100.0	5/5	a
Electrondense bodies (%)	100.0	6/6	100.0	5/5	a
Hearing loss					
Self-noticed or subjective (%) Age at self-noticed or subjective (median (interquartile range))	86.7 15 (30)	26/30 15	93.3 9 (19)	14/15 9	0.651 0.575
Audiogram (%) Age at diagnosis of hearing loss by audiogram (median (interquartile range))	94.1 27.5 (18)	16/17 16	90.0 32 (22)	9/10 7	1.000 0.898
Ocular abnormalities					
Anterior Lenticonus (%) Age at diagnosis (median (interquartile range))	33.3 19 (12)	5/15 5	33.3 30 (-)	3/9 1	1.000 0.665
Maculopathy (%) Age at diagnosis (median (interquartile range))	50.0 25 (19.3)	10/20 8	62.5 34 (25*)	5/8 3	0.686 0.213
Cataracts (%) Age at diagnosis (median (interquartile range))	46.2 29.5 (35.8)	6/13 6	83.3 42.5 (17*)	5/6 2	0.177 0.676
Leiomyomatosis (%) Age at diagnosis (median (interquartile range))	3.3 24 (-)	1/30 1	0.0 -	0/15 0	1.000 -
Family History					
Family history of hematuria (%)	96.6	28/29	72.7	8/11	0.056
Family history of CKD (%)	90.0	27/30	71.4	10/14	0.184
Family history of hearing loss (%)	93.1	27/29	76.9	10/13	0.162

SD: standard deviation. Age is expressed in years. eGFR is expressed in ml/min/1.73m^2. *The range is indicated instead of the interquartile range. [a] Statistics cannot be computed for this variable because the proportion in the two groups is a constant.

Table XIV. Comparison of clinical characteristics between female patients with pathogenic *COL4A5* mutations and patients with pathogenic heterozygous *COL4A3* and *COL4A4* mutations.

	Female XLAS (n=35)		TBMN (n=44)		p-value
		N		N	
Phenotype					
Index-cases (%)	11.4	4/35	29.5	13/44	0.052
Male gender (%)	0.0	0/35	31.8	14/44	0.000
Age at enrollment (mean (SD))	42 (14.8)	35	45.5 (26.8)	45	0.348
Renal abnormalities					
History of macroscopic hematuria (%) Age at diagnosis (median (interquartile range))	17.4 3 (11.5*)	4/23 3	12.0 17 (12*)	3/25 3	0.696 0.025
History of microscopic hematuria (%) Age at diagnosis (median (interquartile range))	100.0 21 (26.5)	35/35 28	77.1 26 (20.5)	27/35 26	0.005 0.216
History of proteinuria (%) Age at diagnosis (median (interquartile range))	78.8 24 (19)	26/33 23	69.7 24.5 (19.5)	23/33 22	0.574 0.315
History of hypertension (%) Age at diagnosis (median (interquartile range))	52.9 34 (22.3)	18/34 16	51.5 42 (18.5)	17/33 13	0.907 0.692
History of CKD stage 2 or higher (%) Age at diagnosis (median (interquartile range)) eGFR at diagnosis (median (interquartile range))	61.8 31.5(17.5) 61 (26.3)	21/34 20 20	51.4 46 (19.5) 67 (20)	18/35 17 15	0.387 0.269 0.257
History of renal replacement therapy (%) Age at onset (median (interquartile range)) eGFR at onset (median (interquartile range))	17.1 40 (14) 6 (1)	6/35 6 5	13.9 36 (22) 14 (-)	5/36 5 1	0.705 0.541 0.120
GBM ultrastuctural changes Age at diagnosis (median (interquartile range))	 24 (11*)	 3	 34 (19)	 9	 0.149
Thinning (%)	100	3/3	100	8/8	a
Thickening (%)	100	3/3	100	3/3	a
Lamellation (%)	100	3/3	60	3/5	0.464
Electrondense bodies (%)	100	3/3	67	2/3	1.000
Hearing loss					
Self-noticed or subjective (%) Age at self-noticed or subjective (median (interquartile range))	45.5 39 (19)	15/33 9	35.3 36 (32*)	12/34 3	0.397 0.557
Confirmed by audiogram (%) Age at diagnosis of hearing loss by audiogram (median (interquartile range))	66.7 41 (10)	14/21 13	52.6 44 (26.5)	10/19 9	0.366 0.567
Ocular abnormalities					
Anterior Lenticonus (%) Age at diagnosis (median (interquartile range))	0.0 -	0/18 0	0.0 -	0/25 0	a b
Maculopathy (%) Age at diagnosis (median (interquartile range))	26.3 44 (22)	5/19 5	9.1 40.5 (35*)	3/33 2	0.124 0.419
Cataracts (%) Age at diagnosis (median (interquartile range))	15.0 56 (56*)	3/20 3	12.0 58 (15*)	3/25 3	1.000 0.832
Leiomyomatosis (%) Age at diagnosis (median (interquartile range))	5.7 18 (-)	2/35 1	0.0 -	0/44 0	0.193 -
Family History					
Family history of hematuria (%)	94.3	33/35	97.6	41/42	0.588
Family history of CKD (%)	97.1	34/35	95.5	42/44	1.000
Family history of hearing loss (%)	94.3	33/35	73.2	30/41	0.015

SD: standard deviation. Age is expressed in years. eGFR is expressed in ml/min/1.73m^2. *The range is indicated instead of the interquartile range. [a] Statistics cannot be computed for this variable because the proportion in the two groups is a constant. [b] No analysis was performed because all cases are missing.

Table XV. Comparison of the phenotype of patients with hemizygous pathogenic *COL4A5* mutations or homozygous and compound heterozygous pathogenic *COL4A3* and COL4A4 mutations (Male XLAS and ARAS) and patients with heterozygous pathogenic *COL4A3*, *COL4A4* and *COL4A5* mutations (Female XLAS and TBMN).

	Male XLAS and ARAS (n=45)		Female XLAS and TBMN (n=79)		p-value
		N		N	
Phenotype					
Index-cases (%)	66.7	30/45	21.5	17/79	0.000
Male gender (%)	84.4	38/45	17.7	14/79	0.000
Age at enrollment (mean (SD))	37.9 (14.9)	45	43.9 (23)	79	0.040
Renal abnormalities					
History of macroscopic hematuria (%)	59.4	19/32	14.6	7/48	0.000
Age at diagnosis (median (interquartile range))	5 (6)	14	12.5 (16.6)	6	0.323
History of microscopic hematuria (%)	100	34/34	88.6	62/70	0.051
Age at diagnosis (median (interquartile range))	12.5 (18)	28	24 (24.3)	54	0.006
History of proteinuria (%)	97.4	37/38	74.2	49/66	0.002
Age at diagnosis (median (interquartile range))	18 (17.5)	30	24 (19.5)	45	0.059
History of hypertension (%)	79.5	31/39	52.2	35/67	0.005
Age at diagnosis (median (interquartile range))	19.5 (12.3)	20	40 (21.5)	29	0.000
History of CKD stage 2 or higher (%)	90.7	39/43	56.5	39/69	0.000
Age at diagnosis (median (interquartile range))	20 (10)	25	39 (20)	37	0.000
eGFR at diagnosis (median (interquartile range))	46 (35.5)	16	65 (24)	35	0.001
History of renal replacement therapy (%)	84.4	38/45	15.5	11/71	0.000
Age at onset (median (interquartile range))	23 (12)	36	39 (15)	11	0.065
eGFR at onset (median (interquartile range))	9 (4.75)	14	6 (3)	6	0.062
GBM ultrastuctural changes					
Age at diagnosis (median (interquartile range))	19 (8.75)	12	28 (19.25)	12	0.006
Thinning (%)	100	11/11	100	11/11	a
Thickening (%)	100	10/10	100	6/6	a
Lamellation (%)	100	11/11	75	6/8	0.164
Electrondense bodies (%)	100	11/11	83.3	5/6	0.353
Hearing loss					
Self-noticed or subjective (%)	88.9	40/45	40.3	27/67	0.000
Age at self-noticed or subjective (median (interquartile range))	14 (21.5)	24	38.5 (18)	12	0.029
Confirmed by audiogram (%)	92.6	25/27	60.0	24/40	0.003
Age at diagnosis of hearing loss by audiogram (median (interquartile range))	30 (19)	23	44 (15)	22	0.001
Ocular abnormalities					
Anterior Lenticonus (%)	33.3	8/24	0	0/43	0.000
Age at diagnosis (median (interquartile range))	21.5 (13)	6	-	-	-
Maculopathy (%)	53.6	15/28	15.4	8/52	0.000
Age at diagnosis (median (interquartile range))	26 (21)	11	44 (26)	7	0.258
Cataracts (%)	57.9	11/19	13.3	6/45	0.000
Age at diagnosis (median (interquartile range))	34 (29.3)	8	57 (25.3)	6	0.159
Leiomyomatosis (%)	2.2	1/45	2.5	2/79	1.000
Age at diagnosis (median (interquartile range))	24 (0*)	1	18 (0*)	1	0.317
Family History					
Family history of hematuria (%)	90.0	36/40	96.1	74/77	0.228
Family history of CKD (%)	84.1	37/44	96.2	76/79	0.034
Family history of hearing loss (%)	88.1	37/42	82.9	63/76	0.452

SD: standard deviation. Age is expressed in years. eGFR is expressed in ml/min/1.73m^2. *The range is indicated instead of the interquartile range. [a] Statistics annot be computed for this variable because the proportion in the two groups is a constant.

3.8. Patients without known mutations: differences and similarities

The group of probands in whom no pathogenic *COL4A3*, *COL4A4* or *COL4A5* mutation was identified is heterogeneous (Tables XVI and XVII). The median age at enrolment was 37 years old (interquartile range 23 years) and there was a preponderance of female probands (73%). Microscopic hematuria and proteinuria were detected in all probands without known pathogenic *COL4A3*, *COL4A4* or *COL4A5* mutations. The proportion of patients who manifested CKD (≥ stage 2) was significantly lower in the group of patients without *COL4A3*, *COL4A4* or *COL4A5* mutations than in the group of male XLAS and ARAS, although the proportion of patients who needed RRT was not significantly different. Extra-renal manifestations were less frequent in the group of patients without *COL4A3*, *COL4A4* or *COL4A5* mutations than in the group of male XLAS and ARAS. Could other disorders cause these manifestations or is this an Alport-like subgroup of patients? Although thinning, thickening and electron-dense bodies were observed in the GBM of the majority of kidney biopsies of patients without a known pathogenic mutation, lamellation was only observed in a minority of patients. This might point to an AS-like phenotype. This phenotypic heterogeneity might be caused by locus heterogeneity (known or unknown genes), modifier genes (predisposing/high-risk versus protecting genetic variants) and/or environmental factors.

In this Portuguese cohort, the clinical screening for AS based in three out of four clinical criteria in a patient with unexplained hematuria and/or CKD had a reasonably high clinical sensitivity, but a low specificity. It was effective in selecting individuals for *COL4A5*, *COL4A4* and *COL4A3* molecular analysis, but it also selected a relatively high number of individuals with a similar phenotype who did not have detectable mutations in any of the three genes. Using the cut-off of ≥ two clinical criteria, the clinical sensitivity and specificity are similar. In conclusion, the characteristics, particularly the low specificity, of the clinical criteria proposed by Flinter and colleagues (1988) to diagnose patients with AS and that was used to recruit participants for this study, may be the reason for not identifying pathogenic mutations in any of the three genes in 15 out of 62 probands (24%) enrolled in this research study. These sensitivity and specificity figures are reasonable and indicate

that a cut-off of two clinical diagnostic criteria may also perform fairly well as a clinical test in the selection of patients for molecular analysis. Despite the large overall sample size (n=146), the clinical manifestations were age- and gender-dependent and many values were missing. The wide confidence intervals demonstrate that the estimates of sensitivity based on small numbers were imprecise. In other words, since there were very few ‘positives’, the precision of the sensitivity is low. However, the lower specificity of the clinical screening test fitted an additional aim of the study by enabling the recruitment of a further group of patients with a phenotype similar to AS for future etiologic investigation.

Table XVI. Comparison of the phenotype of probands without known pathogenic *COL4A3*, *COL4A4* and *COL4A5* mutations (Probands without known mutations) and patients with hemizygous pathogenic *COL4A5* mutations or homozygous and compound heterozygous pathogenic *COL4A3* and *COL4A4* mutations (Male XLAS and ARAS).

	Probands without known mutations (n=15)		Male XLAS and ARAS (n=45)		p-value
		N		N	
Phenotype					
Index-cases (%)	100.0	15/15	66.7	30/45	0.013
Male gender (%)	26.7	4/15	84.4	38/45	0.000
Age at enrollment (median (interquartile range))	37 (23)	15	37.9 (14.9)	45	0.585
Renal abnormalities					
History of macroscopic hematuria (%) Age at diagnosis (median (interquartile range))	27.3 11 (8*)	3/11 2	59.4 5 (6)	19/32 14	0.066 0.584
History of microscopic hematuria (%) Age at diagnosis (median (interquartile range))	100.0 28 (38)	9/9 7	100.0 12.5 (18)	34/34 28	a 0.229
History of proteinuria (%) Age at diagnosis (median (interquartile range))	100 26 (26)	11/11 9	97.4 18 (17.5)	37/38 30	1.000 0.366
History of hypertension (%) Age at diagnosis (median (interquartile range))	66.7 31 (26)	10/15 9	79.5 19.5 (12.3)	31/39 20	0.478 0.084
History of CKD stage 2 or higher (%) Age at diagnosis (median (interquartile range)) eGFR at diagnosis (mean (SD))	60.0 23 (31) 33 (65.3)	9/15 9 4	90.7 20 (10) 46 (35.5)	39/43 25 16	0.013 0.246 0.539
History of renal replacement therapy (%) Age at onset (median (interquartile range)) eGFR at onset (median (interquartile range))	53.8 35 (29) 7 (2*)	7/13 7 2	84.4 23 (12) 9 (4.75)	38/45 36 14	0.053 0.171 0.299
GBM ultrastuctural changes Age at diagnosis (median (interquartile range))	 20.5 (20.8)	 6	 19 (8.75)	 12	 0.876
Thinning (%)	100.0	5/5	100.0	11/11	a
Thickening (%)	80.0	4/5	100.0	10/10	0.333
Lamellation (%)	25.0	1/4	100.0	11/11	0.009
Electrondense bodies (%)	80.0	4/5	100.0	11/11	0.313
Hearing loss					
Self-noticed or subjective (%) Age at self-noticed or subjective (median (interquartile range))	38.5 29 (29*)	5/13 3	88.9 14 (21.5)	40/45 24	0.001 0.448
Confirmed by audiogram (%) Age at diagnosis of hearing loss by audiogram (median (interquartile range))	90.9 52.5 (32.3)	10/11 4	92.6 30 (19)	25/27 23	1.000 0.025
Ocular abnormalities					
Anterior Lenticonus (%) Age at diagnosis (median (interquartile range))	0.0 -	0/8 -	33.3 21.5 (13)	8/24 6	0.081 -
Maculopathy (%) Age at diagnosis (median (interquartile range))	11.1 -	1/9 -	53.6 26 (21)	15/28 11	0.050 -
Cataracts (%) Age at diagnosis (median (interquartile range))	0.0 -	0/9 -	57.9 34 (29.3)	11/19 8	0.004 -
Leiomyomatosis (%) Age at diagnosis (median (interquartile range))	0.0 -	0/15 -	2.2 24 (0*)	1/45 1	1.000 -
Family History					
Family history of hematuria (%)	85.7	6/7	90.0	36/40	0.571
Family history of CKD (%)	76.9	10/13	84.1	37/44	0.680
Family history of hearing loss (%)	53.8	7/13	88.1	37/42	0.014

SD: standard deviation. Age is expressed in years. eGFR is expressed in ml/min/1.73m^2.
*The range is indicated instead of the interquartile range. [a] Statistics cannot be computed for this variable because the proportion in the two groups is a constant.

Table XVII. Comparison of phenotype between probands without *COL4A3*, *COL4A4* and *COL4A5* mutations (Probands without known mutations) and patients with heterozygous mutations in the *COL4A3*, *COL4A4* or *COL4A5* genes.

	Probands without mutations (n=15)		Female XLAS and TBMN (n=79)		p-value
		N		N	
Phenotype					
Index-cases (%)	100	15/15	21.5	17/79	0.000
Male gender (%)	26.7	4/15	17.7	14/79	0.476
Age at enrollment (median (interquartile range))	37 (23)	15	43.9 (23)	79	0.451
Renal abnormalities					
History of macroscopic hematuria (%) Age at diagnosis (median (interquartile range))	27.3 11 (8*)	3/11 2	14.6 12.5 (16.6)	7/48 6	0.376 0.749
History of microscopic hematuria (%) Age at diagnosis (median (interquartile range))	100.0 28 (38)	9/9 7	88.6 24 (24.3)	62/70 54	0.586 0.789
History of proteinuria (%) Age at diagnosis (median (interquartile range))	100 26 (26)	11/11 9	74.2 24 (19.5)	49/66 45	0.109 0.982
History of hypertension (%) Age at diagnosis (median (interquartile range))	66.7 31 (26)	10/15 9	52.2 40 (21.5)	35/67 29	0.310 0.088
History of CKD stage 2 or higher (%) Age at diagnosis (median (interquartile range)) eGFR at diagnosis (mean (SD))	60.0 23 (31) 33 (65.3)	9/15 9 4	56.5 39 (20) 65 (24)	39/69 37 35	0.805 0.465 0.115
History of renal replacement therapy (%) Age at onset (median (interquartile range)) eGFR at onset (median (interquartile range))	53.8 35 (29) 7 (2*)	7/13 7 2	15.5 39 (15) 6 (3)	11/71 11 6	0.005 0.884 0.864
GBM ultrastuctural changes Age at diagnosis (median (interquartile range))	 20.5 (20.8)	 6	 28 (19.25)	 12	 0.038
Thinning (%)	100.0	5/5	100.0	11/11	a
Thickening (%)	80.0	4/5	100.0	6/6	0.455
Lamellation (%)	25.0	1/4	75.0	6/8	0.222
Electrondense bodies (%)	80.0	4/5	83.3	5/6	1.000
Hearing loss					
Self-noticed or subjective (%) Age at self-noticed or subjective (median (interquartile range))	38.5 29 (29*)	5/13 3	40.3 38.5 (18)	27/67 12	0.902 0.296
Audiogram (%) Age at diagnosis of hearing loss by audiogram (median (interquartile range))	90.9 52.5 (32.3)	10/11 4	60.0 44 (15)	24/40 22	0.075 0.412
Ocular abnormalities					
Anterior Lenticonus (%) Age at diagnosis (median (interquartile range))	0.0 -	0/8 -	0 -	0/43 -	a a
Maculopathy (%) Age at diagnosis (median (interquartile range))	11.1 -	1/9 -	15.4 44 (26)	8/52 7	1.000 -
Cataracts (%) Age at diagnosis (median (interquartile range))	0.0 -	0/9 -	13.3 57 (25.3)	6/45 6	0.574 -
Leiomyomatosis (%) Age at diagnosis (median (interquartile range))	0.0 -	0/15 -	2.5 18 (0*)	2/79 1	1.000 -
Family History					
Family history of hematuria (%)	85.7	6/7	96.1	74/77	0.299
Family history of CKD (%)	76.9	10/13	96.2	76/79	0.035
Family history of hearing loss (%)	53.8	7/13	82.9	63/76	0.029

SD: standard deviation. Age is expressed in years. eGFR is expressed in ml/min/1.73m^2.
*The range is indicated instead of the interquartile range. [a] Statistics cannot be computed for this variable because the proportion in the two groups is a constant.

3.9. Diagnostic criteria defined for XLAS by Flinter and colleagues (1988) also enable the diagnosis of patients with ARAS

A complete clinical evaluation (assessment of the four classic criteria of AS) was achieved in 5 probands (5/65; 8%) (Table XVIII). Among the probands in whom the four typical clinical criteria were diagnosed, four had a hemizygous pathogenic *COL4A5* mutation (4/5; 80%) and one had pathogenic compound heterozygous *COL4A3* mutations (1/5; 20%). Among 19 patients with at least three clinical criteria, four had a pathogenic *COL4A5* mutation (4/19; 21%) and 11 had pathogenic *COL4A3* or *COL4A4* mutations (4/19; 58%). Among 28 probands patients with at least two clinical criteria, 10 had a pathogenic *COL4A5* mutation (10/28; 36%) and 10 had pathogenic *COL4A3* or *COL4A4* mutations (10/28; 36%). Among 13 probands patients with one clinical criterion, four had a pathogenic *COL4A5* mutation (4/13; 31%) and 4 had pathogenic *COL4A3* or *COL4A4* mutations (4/13; 31%). The global detection mutation rate of pathogenic *COL4A5*, *COL4A4* and *COL4A3* mutations among probands who met one, two, three or four clinical criteria was 62%, 71%, 79% and 100%, respectively.

The clinical diagnosis of collagen IV-related nephropathies is amongst the biggest challenges in clinical genetics, as it is not always unambiguous (Flinter, Cameron et al. 1988). Precise and strict diagnostic criteria have been used to best identify those patients most likely to harbour *COL4A5* mutations. Additionally, the clinical criteria defined by Flinter and colleagues (1988) to diagnose patients with XLAS proved also useful in the diagnosis of patients with ARAS. This was an expected finding, since the clinical course of severe collagen IV-related nephropathies is similar. More interesting was the fact that these criteria also work for the diagnosis of TBMN, if not used so strictly (Frasca, Onetti-Muda et al. 2005).

Table XVIII. Mutation detection rate according to number of diagnostic criteria met in probands (n=65).

Clinical Criteria							Number of probands with and without pathogenic mutations					Detection mutation rate	
			FH	AC	OC	UC	*COL4A5* mutation	*COL4A4* mutation	*COL4A3* mutation	Without mutation	Total number of probands	Per combination of clinical criteria	Per number of criteria met
1 criterion	FH		10				4	1	2	3	10	70%	62% (8/13)
	AC			1			0	0	1	0	1	100%	
	OC				0						0		
	UC					2	0	0	0	2	2	0%	
2 criteria	FH			15			8	2	2	3	15	80%	71% (20/28)
					1		0	1	0	0	1	100%	
						7	1	2	1	3	7	57%	
	AC				1		0	0	0	1	1	0%	
						4	1	0	2	1	4	75%	
	OC					0					0		
3 criteria	FH	AC			9		2	1	3	3	9	67%	79% (15/19)
						7	1	2	3	1	7	86%	
	FH	OC				0					0		
		AC OC				3	1	0	2	0	3	100%	
4 criteria	FH	AC OC				5	4	0	1	0	5	100%	100% (5/5)

FH: Family history. AC: Hearing loss (Audiologic criteria), including hearing loss with and without confirmation by audiogram.
OC: Ocular signs (Ophthalmologic criteria), including anterior lenticonus, cataracts and retinopathy. UC: Characteristic glomerular basement membrane (GBM) changes (Ultrastructural criteria), including thinning, thickening, lamellation and presence of electron-dense bodies in the GBM.

4. Short term impact of the study: Massive parallel sequencing is recommended to simultaneously study the COL4A5*,* COL4A4 *and* COL4A3 *genes*

The involvement of the *COL4A5*, *COL4A4* and *COL4A3* genes in the etiology of a majority of families with clinical diagnosis or suspicion of AS and in an estimated high proportion of TBMN families poses a major challenge in the identification of the underlying molecular cause of this spectrum of glomerulopathies. Sanger sequencing of these genes for each proband requires a great amount of repetitive, hard work and has time and cost limitations. Notably, new sequencing technologies termed Massive Parallel Sequencing (MPS) – also called Next-Generation Sequencing (NGS) – were developed as methods to overtake the disadvantages of laborious, time-consuming and expensive methodologies. NGS offer the opportunity to detect single nucleotide variants (SNVs), copy number variations (CNVs), insertions and deletions (indels) and other structural variations in a single test (Harismendy, Ng et al. 2009; Ng, Turner et al. 2009; Gullapalli, Lyons-Weiler et al. 2012). Particularly in the case of genetic diseases with high phenotypic and genotypic heterogeneity such as collagen IV-related nephropathies, the simultaneous analysis of the *COL4A5*, *COL4A4* and *COL4A3* genes by NGS has been successfully applied to identify the disease-causing genetic defect (Artuso, Fallerini et al. 2012; Fallerini, Dosa et al. 2013).

NGS has advantages over direct Sanger sequencing to be used in the identification of pathogenic variants in *COL4A5*, *COL4A4* and *COL4A3*. Firstly, the molecular etiology of collagen type IV nephropathy proved to be heterogeneous in the Portuguese population of patients with AS and TBMN. More than 70% of participants had mutations in one of these three genes, being the participant's ratio with pathogenic *COL4A5* mutations to *COL4A3* and *COL4A4* mutations of approximately 1:1. Secondly, as expected, collagen IV-related nephropathies in this Portuguese cohort were caused by different types of mutations, located throughout any of the three genes, which are 48 to 51 exons long. Thirdly, low grade mosaicism is very difficult to detect by Sanger sequencing, but more readily by NGS (Beicht, Strobl-Wildemann et al. 2013). The use of NGS directed to the study of these three genes will avoid the detection of incidental findings, which raise ethical problems

posed by the returning of potential health or reproductive importance of genetic variants information to the patients. In the end, the time (turn-around-time) and cost of the molecular study of collagen IV-related nephropathies would be reduced, while improving the quality of health care. Clinical reasons would also justify performing genetic diagnosis of collagen type IV nephropathies by NGS (Deltas, Pierides et al. 2013). Firstly, collagen IV-related nephropathies have a wide spectrum of manifestations, which are age-dependent and may be gender-dependent. Additionally, a number of XLAS patients follow a milder course reminiscent of that of patients with heterozygous *COL4A3* or *COL4A4* mutations and TBMN, while at the same time a significant subset of patients with TBMN and familial microhematuria progress to CKD or ESRD (Deltas, Pierides et al. 2012). Secondly, in particular if the proband is the only known affected individual in the family, an inheritance pattern would only be confirmed by molecular analysis of the proband. Even when a complete phenotypic characterization of AS patients was performed, our results showed that, comparing the phenotype of XLAS males versus ARAS and XLAS females versus TBMN, there were very few statistically significant differences between the compared groups. If family history is not clear, the three genes would eventually be tested. Thirdly, clinical non-invasive differential diagnosis with other glomerulopathies may be difficult. The challenges of NGS would be the complexity of the workflow of this methodology, which result in high implementation and maintenance costs, and the posterior classification of the pathogenicity of each novel variant. Overall, it is recommended to adjust the molecular genetic analysis strategy of collagen IV-related nephropathies to the Portuguese population for its use in clinical practice.

Whenever *COL4A3* and *COL4A4* direct sequencing is considered in a Portuguese family, we recommend screening for the c.1219G>C (p.Gly407Arg) mutation in the *COL4A3* gene, prior to the screening of the complete *COL4A3* and *COL4A4* coding regions, since it is rapid and cost-effective. This mutation was found in heterozygosity, as well as in compound heterozygosity, in patients who manifested progressive or non-progressive renal disease, with or without extra-renal signs of AS, warranting its study in patients with clinical diagnosis of TBMN and AS, in particular if

an autosomal dominant or recessive inheritance pattern is suggested by the family history.

In the end, the molecular characterization of collagen type IV-related nephropathies, using NGS to simultaneously analyse the *COL4A5*, *COL4A4* and *COL4A3* genes, has the potential to improve the quality of the healthcare system at three levels: prevention (upstream intervention), diagnosis and treatment (downstream intervention). At the diagnosis level, molecular testing enables: (1) confirmation of the clinical diagnosis in the affected individuals, at a younger age and avoiding an extensive, and eventually invasive, etiologic investigation; (2) identification of at-risk family members and determination of their carrier status, by a geneticist. At the treatment level, the molecular diagnosis supports the decision of an early start of effective medical treatment for the prevention of the progression of the glomerulopathy, since very young ages (Massella, Muda et al. 2010; Gross, Friede et al. 2012; Gross, Licht et al. 2012). At the prevention level, the molecular diagnosis of collagen IV-related nephropathies enables an accurate genetic counseling, providing insight into the natural history of the disease (including genotype-phenotype correlations), and contributing to the prevention of severe manifestations in family members by two means: (1) preventing kidney donation between affected family members; (2) enabling PND and PGD for couples at-risk of having severely affected offspring, specifically in a time shortage context of a preconceptional counseling.

5. Long term impact of the study: Future perspectives

Looking into the long term impact of our studies, different lines of investigation may be followed, both directed to basic science and to clinical practice, aiming for a better understanding of genotype-phenotype correlations. A first research question that should be posed is if patients with ARAS and severe TBMN have a similar genetic background. In other words, are patients with severe TBMN indeed patients with ARAS, in whom only one pathogenic mutation was found? A first option to investigate this question could be to search for pathogenic exonic *COL4A3* or *COL4A4* deletions or duplications in patients with a pathogenic heterozygous

mutation in one of these genes, who developed proteinuria and CKD ≥ stage 3, eventually reaching ESRD. Secondly, may known polymorphisms in *COL4A5*, *COL4A4* and *COL4A3* confer a higher risk for progressive CKD? Or, thirdly, may yet unknown genetic variants influence the progression of CKD to ESRD in patients with a collagen IV-related nephropathy? It may be possible that different genetic backgrounds, other than solely pathogenic variants in the *COL4A5*, *COL4A4* and *COL4A3* genes, influence the progression of the kidney disease. These variables should be searched for and tested. In this cohort, around 25% of patients did not have identified pathogenic mutations in *COL4A5*, *COL4A4* or *COL4A3*. May mutations in other genes may be responsible for Alport- and TBMN-like nephropathies? NGS, including exome sequencing and/or whole genome sequencing, may enable identification of these variants.

In this study, renal data from patients with heterozygous mutations in the *COL4A3*, *COL4A4* and *COL4A5* genes showed that proteinuria and progressive CKD were more frequent than expected in this sample of the population, in particular in patients with the mutation c.1219G>C (p.Gly407Arg) that is a common variant in the Portuguese population. Previous studies have shown that microscopic hematuria is not rare. Patients with microscopic hematuria tend to have a higher risk of developing CKD and this occurs at an earlier age when compared with people without microscopic hematuria (Vivante, Afek et al. 2011; Vivante, Calderon-Margalit et al. 2013). Consequently, unexplained microscopic hematuria is not as benign as previously thought. So, future research studies of patients with isolated microscopic hematuria are recommended to determine the proportion of individuals in the Portuguese population with TBMN caused by pathogenic mutations in the *COL4A3* and *COL4A4* genes or in early stages of AS caused by pathogenic *COL4A5* mutations, since it would be a non-invasive method to confirm the diagnosis at an early age and counsel the patients and family members of patients affected by an heterogeneous group of disorders, as the collagen type-IV nephropathies.

Regarding ATS-DL, the development of diffuse leiomyomatosis in three patients of a family with XLAS carrying a *COL4A5* deletion, suggested that the deletion of a previously proposed critical region involving the 5' of *COL4A5* and *COL4A6* genes, as well as the common promotor, is not needed for the development

of diffuse leiomyomatosis. Identifying further patients with a similar genotype-phenotype correlation and mapping deletion breakpoints would help clarifying the molecular pathogenic mechanisms underlying the development of diffuse leiomyomatosis accompanying XLAS. Besides studying *COL4A5* and *COL4A6* in ATS-DL patients, transcriptome (gene expression) analysis in these patients would be crucial to demonstrate the effect of the genomic *COL4A5* deletion in the mRNA synthesis and how it is related with the expression of other genes in the leiomyomatosis pathway. In this case, we hypothesize that the reason for leiomyoma predisposition in ATS-DL is aberrant regulation.

6. Main recommendations

The following practical aspects resulting from the clinical and molecular studies performed in this project are worth to emphasize:

- Microscopic hematuria is not always benign; so, early diagnosis of collagen type IV-related nephropathy as its cause, using reliable molecular genetic methods or renal biopsy should be persued, in order to increase the rate of ethiologic diagnosis, to individualize genetic counseling and to personalize the treatment, since early nephroprotective pharmacological treatment is effective.
- AS should be suspected in any patient with an isolated microscopic hematuria of unknown etiology, as this is the earliest and most consistent manifestation of collagen type IV-related nephropathies across all ages and gender.
- The multidisciplinary study of probands and family members may be extremely informative in the investigation of collagen type IV-related glomerulopathies: screening of extra-renal manifestations should be systematically pursued, namely by audiometric and ophthalmologic evaluation.
- The referral of patients with collagen IV-related nephropathies and their at-risk family members to Medical Genetics consultations is recommended for clinical and molecular diagnosis of collagen IV-related nephropathies, as well as for genetic couselling and ultimately prevention by PND and PGD.

- Patients with collagen IV-related nephropathies should regularly perform renal, audiologic and ophthalmologic evaluation for detection and surveillance of progressive renal and extra-renal manifestations.
- Since earlier stages of CKD can be detected through routine laboratory measurements and adverse outcomes of CKD can be delayed through early detection and effective treatment, the following annual renal screening protocol is suggested for at-risk family members: plasma creatinine, urea, uric acid and ionogram; summary analysis of urine (three occasional urin samples), including the determination of the albumin-to-creatinine ratio (ACR) or the protein-to-creatinine ratio, assessment of blood pressure and renal ultrasound.
- In patients with hematuria and without proteinuria, the basis for annual follow-up is to survey or rule out the appearance of proteinuria, a risk factor for CKD.
- When AS is suspected, an ophthalmologic examination directed to the screening of anterior lenticonus and retinopathy should include at least: 1) biomicroscopy; 2) funduscopy and/or retinography; 3) Optical Coherence Tomography (OCT).
- Since patients with molecularly confirmed XLAS or ARAS had typical kidney biopsies, ie, they manifested the four characteristic ultrastructural changes in the GBM, a more widespread and larger renal biopsy policy, including ultrastructural examination, is recommended, even in cases with moderate urinary abnormalities (isolated persistent hematuria with or without proteinuria), in order to increase the number of patients with an etiology for identified urinary abnormalities.
- The evaluation of ultrastructural changes of the GBM on a kidney biopsy should result in a medical report discriminating the presence or absence of GBM thinning, GBM thickening, GBM lamellation and GBM with electrodense bodies (EB) surrounded by a halo, as the presence of four typical GBM changes correlates with a higher mutation detection rate.
- If a patient (particularly a female or a child) presents with microscopic persistent or intermittent microscopic hematuria, and audiologic and ophthalmologic criteria were excluded, performing a kidney biopsy with electron microscopy analysis may be the only method to clinically diagnose a collagen IV-related nephropathy.
- Consider performing a kidney biopsy in at least one affected member in the family, as it may assist in the interpretation of molecular genetic analysis and may

obviate the need for kidney biopsies in other relatives; however, in a well-studied population with known common mutations, specific mutation analysis may avoid a kidney biopsy.

- The molecular study of *COL4A5* and *COL4A6* is recommended in patients with diffuse leiomyomatosis (namely of the gastrointestinal and female reproductive tracts), particularly in females even without any clinical or laboratory evidence of nephropathy, due to the risk of males with ATS-DL in the offspring.
- Conversely, diffuse leiomyomatosis should be successively investigated in patients with XLAS caused by large *COL4A5* deletions, since deletion of the common *COL4A5/COL4A6* promoter and the 5'exons of *COL4A6* is not needed for the development of diffuse leiomyomatosis.
- Molecular genetics analysis may allow the diagnosis of patients not satisfying all Flinter's criteria.
- Finding the exact gene and the mutation is important for risk assessment, genetic counselling and family programming.
- DNA sequencing remains the gold-standard for the final diagnosis, especially in diseases of genetic heterogeneity; but NGS is expected to enhance the analysis process and lead to robust characterization of more patients.
- MLPA assays should include probes for *COL4A5* exons 8, 25 and 40, as well as for the exon 3 of *COL4A6*.
- The establishment of a national AS database of individuals with *COL4A5*, *COL4A4* and *COL4A3* variants is helpful in the long-term surveillance of clinical manifestations and the treatment of affected individuals (including to provide access to multicentric clinical trials) and effective in the study of family members; its creation will help maintaining good quality clinical archives and a biobank from these patients.

Conclusion

The studies in this thesis contributed to an increase of knowledge on collagen IV-related nephropathies. Pathogenic mutations in the protein coding region of *COL4A5*, *COL4A4* or *COL4A3* were detected in 76% (47/62) of families with suspected or clinical diagnosis of AS. NGS of these genes, including the promotor region and introns, may increase this yield even more. Finding novel genes in patients with an AS-like phenotype will shed light on the molecular foundations of the development and functioning of the kidney through the identification of pathways that lead to glomerulopathies.

References

Abrahamson, D. R., B. G. Hudson, et al. (2009). "Cellular origins of type IV collagen networks in developing glomeruli." J Am Soc Nephrol **20**(7): 1471-1479.

Adzhubei, I., D. M. Jordan, et al. (2013). "Predicting functional effect of human missense mutations using PolyPhen-2." Curr Protoc Hum Genet **Chapter 7**: Unit7 20.

Adzhubei, I. A., S. Schmidt, et al. (2010). "A method and server for predicting damaging missense mutations." Nat Methods **7**(4): 248-249.

Agresti, A. and B. A. Coull (1998). "Approximate is better than "exact" for interval estimation of binomial proportions." American Statistician **52**(2): 119-126.

Agresti, A. and B. A. Coull (1998). "Approximate is better than "exact" for interval estimation of binomial proportions." Am Stat **52**(2): 119-126.

Ahmed, F., K. K. Kamae, et al. (2013). "Temporal macular thinning associated with X-linked Alport syndrome." JAMA Ophthalmol **131**(6): 777-782.

Al-Mahmood, A. M., S. A. Al-Swailem, et al. (2010). "Progressive posterior lenticonus in a patient with alport syndrome." Middle East Afr J Ophthalmol **17**(4): 379-381.

Alport, A. C. (1927). "Hereditary familial congenital haemorrhagic nephritis." Brit. Med. J. (1): 504-506.

Alves, F. R. and A. Q. R. F. de (2005). "Revision about hearing loss in the Alport's syndrome, analyzing the clinical, genetic and bio-molecular aspects." Braz J Otorhinolaryngol **71**(6): 813-819.

Alves, F. R. and A. Ribeiro Fde (2008). "Clinical data and hearing of individuals with Alport syndrome." Braz J Otorhinolaryngol **74**(6): 807-814.

Anker, M. C., J. Arnemann, et al. (2003). "Alport syndrome with diffuse leiomyomatosis." Am J Med Genet A **119A**(3): 381-385.

Antignac, C., B. Knebelmann, et al. (1994). "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**(3): 1195-1207.

Antignac, C., J. Zhou, et al. (1992). "Alport syndrome and diffuse leiomyomatosis: deletions in the 5' end of the COL4A5 collagen gene." Kidney Int **42**(5): 1178-1183.

Arrondel, C., N. Vodovar, et al. (2002). "Expression of the nonmuscle myosin heavy chain IIA in the human kidney and screening for MYH9 mutations in Epstein and Fechtner syndromes." J Am Soc Nephrol **13**(1): 65-74.

Artuso, R., C. Fallerini, et al. (2012). "Advances in Alport syndrome diagnosis using next-generation sequencing." Eur J Hum Genet **20**(1): 50-57.

Aslanzadeh, G. A., D. Gharabaghi, et al. (2008). "Clear lens phacoemulsification in the anterior lenticonus due to Alport Syndrome: two case reports." J Med Case Reports **2**: 178.

Atkin, C. L., S. J. Hasstedt, et al. (1988). "Mapping of Alport syndrome to the long arm of the X chromosome." Am J Hum Genet **42**(2): 249-255.

Badenas, C., M. Praga, et al. (2002). "Mutations in theCOL4A4 and COL4A3 genes cause familial benign hematuria." J Am Soc Nephrol **13**(5): 1248-1254.

Barker, D. F., S. L. Hostikka, et al. (1990). "Identification of mutations in the COL4A5 collagen gene in Alport syndrome." Science **248**(4960): 1224-1227.

Barker, D. F., C. J. Pruchno, et al. (1996). "A mutation causing Alport syndrome with tardive hearing loss is common in the western United States." Am J Hum Genet **58**(6): 1157-1165.

Basta-Jovanovic, G., V. S. Venkataseshan, et al. (1990). "Correlation of glomerular basement membrane alterations with clinical data in progressive hereditary nephritis (Alport's syndrome)." Am J Kidney Dis **16**(1): 51-56.

Beicht, S., G. Strobl-Wildemann, et al. (2013). "Next Generation Sequencing as a Useful Tool in the Diagnostics of Mosaicism in Alport Syndrome." Gene.

Bekheirnia, M. R., B. Reed, et al. (2010). "Genotype-Phenotype Correlation in X-Linked Alport Syndrome." J Am Soc Nephrol.

Bhatnagar, R., A. Kumar, et al. (1990). "Alport's syndrome--ocular manifestations and unusual features." Acta Ophthalmol (Copenh) **68**(3): 347-349.

Blaise, P., P. Delanaye, et al. (2003). "[Anterior lenticonus: diagnostic aid in Alport syndrome]." J Fr Ophtalmol **26**(10): 1075-1082.

Borza, D. B., O. Bondar, et al. (2001). "The NC1 domain of collagen IV encodes a novel network composed of the alpha 1, alpha 2, alpha 5, and alpha 6 chains in smooth muscle basement membranes." J Biol Chem **276**(30): 28532-28540.

Bourque, M. D., N. Spigland, et al. (1989). "Esophageal leiomyoma in children: two case reports and review of the literature." J Pediatr Surg **24**(10): 1103-1107.

Boutaud, A., D. B. Borza, et al. (2000). "Type IV collagen of the glomerular basement membrane. Evidence that the chain specificity of network assembly is encoded by the noncollagenous NC1 domains." J Biol Chem **275**(39): 30716-30724.

Bower, K. S., J. D. Edwards, et al. (2009). "Novel corneal phenotype in a patient with alport syndrome." Cornea **28**(5): 599-606.

Boye, E., G. Mollet, et al. (1998). "Determination of the genomic structure of the COL4A4 gene and of novel mutations causing autosomal recessive Alport syndrome." Am J Hum Genet **63**(5): 1329-1340.

Brainwood, D., C. Kashtan, et al. (1998). "Targets of alloantibodies in Alport anti-glomerular basement membrane disease after renal transplantation." Kidney Int **53**(3): 762-766.

Brunner, H., C. Schroder, et al. (1988). "Alport's syndrome: localization of the X-chromosomal gene and consequences for future investigations." Contrib Nephrol **67**: 200-205.

Bruttini, M., F. Vitelli, et al. (2000). "Mosaicism in Alport syndrome with genetic counselling." J Med Genet **37**(9): 717-719.

Butkowski, R. J., J. Wieslander, et al. (1989). "Basement membrane collagen in the kidney: regional localization of novel chains related to collagen IV." Kidney Int **35**(5): 1195-1202.

Buzza, M., H. Dagher, et al. (2003). "Mutations in the COL4A4 gene in thin basement membrane disease." Kidney Int **63**(2): 447-453.

Byrne, M. C., M. N. Budisavljevic, et al. (2002). "Renal transplant in patients with Alport's syndrome." Am J Kidney Dis **39**(4): 769-775.

Calabrese, R., E. Capriotti, et al. (2009). "Functional annotations improve the predictive score of human disease-related mutations in proteins." Hum Mutat **30**(8): 1237-1244.

Cangiotti, A. M., A. Sessa, et al. (1996). "Evolution of glomerular basement membrane lesions in a male patient with Alport syndrome: ultrastructural and morphometric study." Nephrol Dial Transplant **11**(9): 1829-1834.

Charbit, M., M. C. Gubler, et al. (2007). "Cyclosporin therapy in patients with Alport syndrome." Pediatr Nephrol **22**(1): 57-63.

Cheong, H. I., C. E. Kashtan, et al. (1994). "Immunohistologic studies of type IV collagen in anterior lens capsules of patients with Alport syndrome." Lab Invest **70**(4): 553-557.

Chung, P. C., K. K. Lin, et al. (2007). "Alport syndrome with recurrent herpes simplex virus keratitis." Cornea **26**(10): 1279-1281.

Churg, J. and R. L. Sherman (1973). "Pathologic characteristics of hereditary nephritis." Arch Pathol **95**(6): 374-379.

Claus, F., K. Geboes, et al. (2008). "Clinical Challenges and Images in GI: image 2. Anorectal leiomyomatosis in Alport(-like) syndrome." Gastroenterology **135**(4): 1053, 1432.

Cochat, P., P. Guibaud, et al. (1988). "Diffuse leiomyomatosis in Alport syndrome." J Pediatr **113**(2): 339-343.
Colville, D., J. Savige, et al. (1997). "Ocular manifestations of autosomal recessive Alport syndrome." Ophthalmic Genet **18**(3): 119-128.
Colville, D., Y. Y. Wang, et al. (2009). "The retinal "lozenge" or "dull macular reflex" in Alport syndrome may be associated with a severe retinopathy and early-onset renal failure." Br J Ophthalmol **93**(3): 383-386.
Colville, D. J. and J. Savige (1997). "Alport syndrome. A review of the ocular manifestations." Ophthalmic Genet **18**(4): 161-173.
Cooper, J., G. A. Patterson, et al. (1999). "Clinical-pathologic conference in thoracic surgery: Alport syndrome with diffuse leiomyomatosis." J Thorac Cardiovasc Surg **118**(4): 760-764.
Cosgrove, D. (2012). "Glomerular pathology in Alport syndrome: a molecular perspective." Pediatr Nephrol **27**(6): 885-890.
Crawfurd, M. A. (1988). "Alport's syndrome." J Med Genet **25**(9): 623-627.
Crockett, D. K., G. Pont-Kingdon, et al. (2010). "The Alport syndrome COL4A5 variant database." Hum Mutat **31**(8): E1652-1657.
Dagher, H., M. Buzza, et al. (2001). "A comparison of the clinical, histopathologic, and ultrastructural phenotypes in carriers of X-linked and autosomal recessive Alport's syndrome." Am J Kidney Dis **38**(6): 1217-1228.
Dahan, K., L. Heidet, et al. (1995). "Smooth muscle tumors associated with X-linked Alport syndrome: carrier detection in females." Kidney Int **48**(6): 1900-1906.
Deltas, C., A. Pierides, et al. (2012). "The role of molecular genetics in diagnosing familial hematuria(s)." Pediatr Nephrol **27**(8): 1221-1231.
Deltas, C., A. Pierides, et al. (2013). "Molecular genetics of familial hematuric diseases." Nephrol Dial Transplant **28**(12): 2946-2960.
den Dunnen, J. T. and S. E. Antonarakis (2000). "Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion." Hum Mutat **15**(1): 7-12.
Eason J, S. G. L. M., Buchanan G (1924). "Hereditary and familial nephritis." Lancet **204**(5274): 639-646.
Exome Variant Server, N. G. E. S. P. E., Seattle, WA (URL: http://evs.gs.washington.edu/EVS/) [date (month, yr) accessed].
Fairbrother, W. G., R. F. Yeh, et al. (2002). "Predictive identification of exonic splicing enhancers in human genes." Science **297**(5583): 1007-1013.
Fallerini, C., L. Dosa, et al. (2013). "Unbiased next generation sequencing analysis confirms the existence of autosomal dominant Alport syndrome in a relevant fraction of cases." Clin Genet.
Fehmers, G. A. and R. Croughs (1964). "[ALPORT'S SYNDROME.]." Folia Med Neerl **81**: 183-184.
Firth, H. V. and J. A. Hurst, Eds. (2007). Oxford Desk Reference - Clinical Genetics, Oxford University Press.
Fleming, S. J., C. O. Savage, et al. (1988). "Anti-glomerular basement membrane antibody-mediated nephritis complicating transplantation in a patient with Alport's syndrome." Transplantation **46**(6): 857-859.
Flinter, F. (1997). "Alport's syndrome." J Med Genet **34**(4): 326-330.
Flinter, F., E. R. Maher, et al., Eds. (2003). The Genetics of Renal Disease. New York, Oxford University Press Inc.
Flinter, F. and K. Plant (1998). "Why are mutations in COL4A5 not detectable in all patients with Alport's syndrome?" Nephrol Dial Transplant **13**(6): 1348-1351.
Flinter, F. A., J. S. Cameron, et al. (1988). "Genetics of classic Alport's syndrome." Lancet **2**(8618): 1005-1007.

Frasca, G. M., E. Balestra, et al. (2008). "[Thin glomerular basement membrane disease]." G Ital Nefrol **25**(1): 49-56.

Frasca, G. M., A. Onetti-Muda, et al. (2005). "Thin glomerular basement membrane disease: clinical significance of a morphological diagnosis--a collaborative study of the Italian Renal Immunopathology Group." Nephrol Dial Transplant **20**(3): 545-551.

Froyen, G., H. Van Esch, et al. (2007). "Detection of genomic copy number changes in patients with idiopathic mental retardation by high-resolution X-array-CGH: important role for increased gene dosage of XLMR genes." Hum Mutat **28**(10): 1034-1042.

Garcia-Torres, R. and L. Orozco (1993). "Alport-leiomyomatosis syndrome: an update." Am J Kidney Dis **22**(5): 641-648.

Garcia Torres, R. and V. Guarner (1983). "[Leiomyomatosis of the esophagus, tracheo-bronchi and genitals associated with Alport type hereditary nephropathy: a new syndrome]." Rev Gastroenterol Mex **48**(3): 163-170.

Gehrs, K. M., S. C. Pollock, et al. (1995). "Clinical features and pathogenesis of Alport retinopathy." Retina **15**(4): 305-311.

Gobel, J., C. J. Olbricht, et al. (1992). "Kidney transplantation in Alport's syndrome: long-term outcome and allograft anti-GBM nephritis." Clin Nephrol **38**(6): 299-304.

Gregory, M. C., D. A. Terreros, et al. (1996). "Alport syndrome--clinical phenotypes, incidence, and pathology." Contrib Nephrol **117**: 1-28.

Grodecki, K. M., M. J. Gains, et al. (1997). "Treatment of X-linked hereditary nephritis in Samoyed dogs with angiotensin converting enzyme (ACE) inhibitor." J Comp Pathol **117**(3): 209-225.

Grompe, M. (1993). "The rapid detection of unknown mutations in nucleic acids." Nat Genet **5**(2): 111-117.

Gross, O., B. Beirowski, et al. (2003). "Preemptive ramipril therapy delays renal failure and reduces renal fibrosis in COL4A3-knockout mice with Alport syndrome." Kidney Int **63**(2): 438-446.

Gross, O., T. Friede, et al. (2012). "Safety and Efficacy of the ACE-Inhibitor Ramipril in Alport Syndrome: The Double-Blind, Randomized, Placebo-Controlled, Multicenter Phase III EARLY PRO-TECT Alport Trial in Pediatric Patients." ISRN Pediatr **2012**: 436046.

Gross, O. and C. E. Kashtan (2009). "Treatment of Alport syndrome: beyond animal models." Kidney Int **76**(6): 599-603.

Gross, O., C. Licht, et al. (2012). "Early angiotensin-converting enzyme inhibition in Alport syndrome delays renal failure and improves life expectancy." Kidney Int **81**(5): 494-501.

Gross, O., K. O. Netzer, et al. (2002). "Meta-analysis of genotype-phenotype correlation in X-linked Alport syndrome: impact on clinical counselling." Nephrol Dial Transplant **17**(7): 1218-1227.

Gross, O., E. Schulze-Lohoff, et al. (2004). "Antifibrotic, nephroprotective potential of ACE inhibitor vs AT1 antagonist in a murine model of renal fibrosis." Nephrol Dial Transplant **19**(7): 1716-1723.

Gross, O., M. Weber, et al. (2009). "Living donor kidney transplantation from relatives with mild urinary abnormalities in Alport syndrome: long-term risk, benefit and outcome." Nephrol Dial Transplant **24**(5): 1626-1630.

Grunfeld, J. P. (1985). "The clinical spectrum of hereditary nephritis." Kidney Int **27**(1): 83-92.

Gubler, M. C., B. Knebelmann, et al. (1995). "Autosomal recessive Alport syndrome: immunohistochemical study of type IV collagen chain distribution." Kidney Int **47**(4): 1142-1147.

Guillem, P., F. Delcambre, et al. (2001). "Diffuse esophageal leiomyomatosis with perirectal involvement mimicking Hirschsprung disease." Gastroenterology **120**(1): 216-220.

Gullapalli, R. R., M. Lyons-Weiler, et al. (2012). "Clinical integration of next-generation sequencing technology." Clin Lab Med **32**(4): 585-599.

Gunwar, S., F. Ballester, et al. (1998). "Glomerular basement membrane. Identification of a novel disulfide-cross-linked network of alpha3, alpha4, and alpha5 chains of type IV collagen and its implications for the pathogenesis of Alport syndrome." J Biol Chem **273**(15): 8767-8775.

Guo, C., B. Van Damme, et al. (1993). "Differential splicing of COL4A5 mRNA in kidney and white blood cells: a complex mutation in the COL4A5 gene of an Alport patient deletes the NC1 domain." Kidney Int **44**(6): 1316-1321.

Guthrie, L. B. (1902). "Idiopathic,' or congenital, hereditary and familial haematuria." Lancet **159**(4105): 1243-1246.

Haas, M. (2009). "Alport syndrome and thin glomerular basement membrane nephropathy: a practical approach to diagnosis." Arch Pathol Lab Med **133**(2): 224-232.

Haberal, M., R. Emiroglu, et al. (2001). "Alport's syndrome and renal transplantation." Transplant Proc **33**(5): 2702-2703.

Hanson, H., H. Storey, et al. (2010). "The Value of Clinical Criteria in Identifying Patients with X-Linked Alport Syndrome." Clin J Am Soc Nephrol.

Hanson, H., H. Storey, et al. (2011). "The value of clinical criteria in identifying patients with X-linked Alport syndrome." Clin J Am Soc Nephrol **6**(1): 198-203.

Harismendy, O., P. C. Ng, et al. (2009). "Evaluation of next generation sequencing platforms for population targeted sequencing studies." Genome Biol **10**(3): R32.

Harvey, S. J., K. Zheng, et al. (1998). "Role of distinct type IV collagen networks in glomerular development and function." Kidney Int **54**(6): 1857-1866.

Hasstedt, S. J. and C. L. Atkin (1983). "X-linked inheritance of Alport syndrome: family P revisited." Am J Hum Genet **35**(6): 1241-1251.

Heidet, L., C. Arrondel, et al. (2001). "Structure of the human type IV collagen gene COL4A3 and mutations in autosomal Alport syndrome." J Am Soc Nephrol **12**(1): 97-106.

Heidet, L., E. Boye, et al. (1998). "Somatic deletion of the 5' ends of both the COL4A5 and COL4A6 genes in a sporadic leiomyoma of the esophagus." Am J Pathol **152**(3): 673-678.

Heidet, L., Y. Cai, et al. (1997). "Diffuse leiomyomatosis associated with X-linked Alport syndrome: extracellular matrix study using immunohistochemistry and in situ hybridization." Lab Invest **76**(2): 233-243.

Heidet, L., L. Cohen-Solal, et al. (1997). "Novel COL4A5/COL4A6 deletions and further characterization of the diffuse leiomyomatosis-Alport syndrome (DL-AS) locus define the DL critical region." Cytogenet Cell Genet **78**(3-4): 240-246.

Heidet, L., K. Dahan, et al. (1995). "Deletions of both alpha 5(IV) and alpha 6(IV) collagen genes in Alport syndrome and in Alport syndrome associated with smooth muscle tumours." Hum Mol Genet **4**(1): 99-108.

Heikkila, P. and R. Soininen (1996). "The type IV collagen gene family." Contrib Nephrol **117**: 105-129.

Heikkila, P., A. Tibell, et al. (2001). "Adenovirus-mediated transfer of type IV collagen alpha5 chain cDNA into swine kidney in vivo: deposition of the protein into the glomerular basement membrane." Gene Ther **8**(11): 882-890.

Heikkila, P., K. Tryggvason, et al. (2000). "Animal models of Alport syndrome: advancing the prospects for effective human gene therapy." Exp Nephrol **8**(1): 1-7.

Hentati, N., D. Sellami, et al. (2008). "[Ocular findings in Alport syndrome: 32 case studies]." J Fr Ophtalmol **31**(6 Pt 1): 597-604.

Hertz, J. M. (2009). "Alport syndrome. Molecular genetic aspects." Dan Med Bull **56**(3): 105-152.

Hertz, J. M., I. Juncker, et al. (2008). "MLPA and cDNA analysis improves COL4A5 mutation detection in X-linked Alport syndrome." Clin Genet **74**(6): 522-530.

Hertz, J. M., I. Juncker, et al. (2001). "Detection of mutations in the COL4A5 gene by SSCP in X-linked Alport syndrome." Hum Mutat **18**(2): 141-148.

Hertz, J. M., U. Persson, et al. (2005). "Alport syndrome caused by inversion of a 21 Mb fragment of the long arm of the X-chromosome comprising exon 9 through 51 of the COL4A5 gene." Hum Genet **118**(1): 23-28.

Hertz, J. M., M. Thomassen, et al. (2012). "Clinical utility gene card for: Alport syndrome." Eur J Hum Genet **20**(6).

Herwig, M. C., N. Eter, et al. (2011). "Corneal clouding in Alport syndrome." Cornea **30**(3): 367-370.

Hild, M., H. S. Walter, et al. (2009). "[Unclear deterioration of vision after renal transplant]." Ophthalmologe **106**(11): 1024-1028.

Hinglais, N., J. P. Grunfeld, et al. (1972). "Characteristic ultrastructural lesion of the glomerular basement membrane in progressive hereditary nephritis (Alport's syndrome)." Lab Invest **27**(5): 473-487.

Hino, S., T. Takemura, et al. (1996). "Absence of alpha 6(IV) collagen in kidney and skin of X-linked Alport syndrome patients." Pediatr Nephrol **10**(6): 742-744.

Hoefele, J., B. Lange-Sperandio, et al. (2010). "Novel heterozygous COL4A3 mutation in a family with late-onset ESRD." Pediatr Nephrol **25**(8): 1539-1542.

Hostikka, S. L., R. L. Eddy, et al. (1990). "Identification of a distinct type IV collagen alpha chain with restricted kidney distribution and assignment of its gene to the locus of X chromosome-linked Alport syndrome." Proc Natl Acad Sci U S A **87**(4): 1606-1610.

Hudson, B. G. (2004). "The molecular basis of Goodpasture and Alport syndromes: beacons for the discovery of the collagen IV family." J Am Soc Nephrol **15**(10): 2514-2527.

Hudson, B. G., R. Kalluri, et al. (1994). "Structure and organization of type IV collagen of renal glomerular basement membrane." Contrib Nephrol **107**: 163-167.

Hudson, B. G., R. Kalluri, et al. (1992). "The pathogenesis of Alport syndrome involves type IV collagen molecules containing the alpha 3(IV) chain: evidence from anti-GBM nephritis after renal transplantation." Kidney Int **42**(1): 179-187.

Hudson, B. G., K. Tryggvason, et al. (2003). "Alport's syndrome, Goodpasture's syndrome, and type IV collagen." N Engl J Med **348**(25): 2543-2556.

Hurst, A. (1923). "Hereditary familial congenital haemorrhagic nephritis: occurring in sixteen individuals in three generations." Guy's Hospital Reports(3): 368-370.

Inoue, Y., H. Nishio, et al. (1999). "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**(5): 854-862.

Jacobs, K. and F. M. Meire (2000). "Lenticonus." Bull Soc Belge Ophtalmol(277): 65-70.

Jais, J. P., B. Knebelmann, et al. (2003). "X-linked Alport syndrome: natural history and genotype-phenotype correlations in girls and women belonging to 195 families: a "European Community Alport Syndrome Concerted Action" study." J Am Soc Nephrol **14**(10): 2603-2610.

Jais, J. P., B. Knebelmann, et al. (2000). "X-linked Alport syndrome: natural history in 195 families and genotype- phenotype correlations in males." J Am Soc Nephrol **11**(4): 649-657.

Johnston, J. B., O. T. Clagett, et al. (1953). "Smooth-muscle tumours of the oesophagus." Thorax **8**(4): 251-265.

Jonsson, J. J., A. Renieri, et al. (1998). "Alport syndrome, mental retardation, midface hypoplasia, and elliptocytosis: a new X linked contiguous gene deletion syndrome?" J Med Genet **35**(4): 273-278.

Junk, A. K., F. H. Stefani, et al. (2000). "Bilateral anterior lenticonus: Scheimpflug imaging system documentation and ultrastructural confirmation of Alport syndrome in the lens capsule." Arch Ophthalmol **118**(7): 895-897.

Kalluri, R., C. F. Shield, et al. (1997). "Isoform switching of type IV collagen is developmentally arrested in X-linked Alport syndrome leading to increased susceptibility of renal basement membranes to endoproteolysis." J Clin Invest **99**(10): 2470-2478.

Karakayali, H., S. Pehlivan, et al. (2008). "Acute rejection rates and survival of renal transplant recipients with Alport's Syndrome." Transplant Proc **40**(1): 120-122.
Kashtan, C., A. J. Fish, et al. (1986). "Nephritogenic antigen determinants in epidermal and renal basement membranes of kindreds with Alport-type familial nephritis." J Clin Invest **78**(4): 1035-1044.
Kashtan, C. E. (1990). "Should women who are known or potential carriers of the Alport gene be accepted as kidney donors?" Pediatr Nephrol **4**(3): 248.
Kashtan, C. E. (1993). Alport Syndrome and Thin Basement Membrane Nephropathy. GeneReviews. R. A. Pagon, M. P. Adam, T. D. Birdet al. Seattle (WA).
Kashtan, C. E. (1993). "Collagen IV-Related Nephropathies (Alport Syndrome and Thin Basement Membrane Nephropathy)."
Kashtan, C. E. (1993-2013). "Alport Syndrome and Thin Basement Membrane Nephropathy." GeneReviews™ [Internet].
Kashtan, C. E. (1999). "Alport syndrome. An inherited disorder of renal, ocular, and cochlear basement membranes." Medicine (Baltimore) **78**(5): 338-360.
Kashtan, C. E. (1999). "Alport syndrome: is diagnosis only skin-deep?" Kidney Int **55**(4): 1575-1576.
Kashtan, C. E. (2001). Collagen IV-related nephropathies (Alport Syndrome and Thin Basement Membrane Nephropathy). GeneReviews™. R. A. Pagon, [Editor-in-chief], T. D. Birdet al. Seattle (WA), USA, University of Washington, Seattle; 1993-.
Kashtan, C. E. (2001). Collagen IV-related nephropathies (Alport Syndrome and Thin Basement Membrane Nephropathy). GeneReviews™. R. A. Pagon, [Editor-in-chief], T. D. Birdet al. Seattle (WA), USA, University of Washington, Seattle; 1993-.
Kashtan, C. E. (2001). "Familial hematuric syndromes--Alport syndrome, thin glomerular basement membrane disease and Fechtner/Epstein syndromes." Contrib Nephrol(136): 79-99.
Kashtan, C. E. (2009). "Women with Alport syndrome: risks and rewards of kidney donation." Nephrol Dial Transplant **24**(5): 1369-1370.
Kashtan, C. E., R. J. Butkowski, et al. (1990). "Posttransplant anti-glomerular basement membrane nephritis in related males with Alport syndrome." J Lab Clin Med **116**(4): 508-515.
Kashtan, C. E., P. T. McEnery, et al. (1995). "Renal allograft survival according to primary diagnosis: a report of the North American Pediatric Renal Transplant Cooperative Study." Pediatr Nephrol **9**(6): 679-684.
Kashtan, C. E. and A. F. Michael (1996). "Alport syndrome." Kidney Int **50**(5): 1445-1463.
Kefalides, N. A. (1966). "A collagen of unusual composition and a glycoprotein isolated from canine glomerular basement membrane." Biochem Biophys Res Commun **22**(1): 26-32.
Kefalides, N. A. (1968). "Isolation and characterization of the collagen from glomerular basement membrane." Biochemistry **7**(9): 3103-3112.
Kefalides, N. A. (1971). "Isolation of a collagen from basement membranes containing three identical alpha chains." Biochem Biophys Res Commun **45**(1): 226-234.
Kendall, G. a. H., A.F. (1912). "Hereditary familial congenital haemorrhagic nephritis." Guy's Hosp. Rep. **66**: 137-141.
Kenney, L. J. (1953). "Giant intramural leiomyoma of esophagus; a case report." J Thorac Surg **26**(1): 93-100.
Khoshnoodi, J., J. P. Cartailler, et al. (2006). "Molecular recognition in the assembly of collagens: terminal noncollagenous domains are key recognition modules in the formation of triple helical protomers." J Biol Chem **281**(50): 38117-38121.
Khoshnoodi, J., V. Pedchenko, et al. (2008). "Mammalian collagen IV." Microsc Res Tech **71**(5): 357-370.
King, K., F. A. Flinter, et al. (2006). "A two-tier approach to mutation detection in the COL4A5 gene for Alport syndrome." Hum Mutat **27**(10): 1061.

King, K., F. A. Flinter, et al. (2002). "Unusual deep intronic mutations in the COL4A5 gene cause X linked Alport syndrome." Hum Genet **111**(6): 548-554.

Kinoshita, Y., G. Osawa, et al. (1969). "Hereditary chronic nephritis (alport) complicated by nephrotic syndrome. Light, Fluorescent and electronmicroscopic studies of renal biopsy specimens." Acta Med Biol (Niigata) **17**(2): 101-117.

Kleppel, M. M., C. Kashtan, et al. (1989). "Distribution of familial nephritis antigen in normal tissue and renal basement membranes of patients with homozygous and heterozygous Alport familial nephritis. Relationship of familial nephritis and Goodpasture antigens to novel collagen chains and type IV collagen." Lab Invest **61**(3): 278-289.

Kleppel, M. M., P. A. Santi, et al. (1989). "Human tissue distribution of novel basement membrane collagen." Am J Pathol **134**(4): 813-825.

Knebelmann, B., C. Breillat, et al. (1996). "Spectrum of mutations in the COL4A5 collagen gene in X-linked Alport syndrome." Am J Hum Genet **59**(6): 1221-1232.

Kruegel, J., D. Rubel, et al. (2013). "Alport syndrome--insights from basic and clinical research." Nat Rev Nephrol **9**(3): 170-178.

Kruegel, J., D. Rubel, et al. (2013). "Alport syndrome-insights from basic and clinical research." Nat Rev Nephrol **9**(3): 170-178.

Le Bras, A., A. David, et al. (1998). "[Diffuse leiomyomatosis with genital involvement and Alport syndrome. Report of two cases]." J Gynecol Obstet Biol Reprod (Paris) **27**(5): 523-528.

LeBleu, V., H. Sugimoto, et al. (2009). "Stem cell therapies benefit Alport syndrome." J Am Soc Nephrol **20**(11): 2359-2370.

Leborgne, J., J. C. Le Neel, et al. (1989). "[Diffuse esophageal leiomyomatosis. Apropos of 5 cases with 2 familial cases]." Chirurgie **115**(4-5): 277-285; discussion 286.

Legius, E., W. Proesmans, et al. (1990). "Muscular hypertrophy of the oesophagus and "Alport-like" glomerular lesions in a boy." Eur J Pediatr **149**(9): 623-627.

Lemmink, H. H., T. Mochizuki, et al. (1994). "Mutations in the type IV collagen alpha 3 (COL4A3) gene in autosomal recessive Alport syndrome." Hum Mol Genet **3**(8): 1269-1273.

Lemmink, H. H., W. N. Nillesen, et al. (1996). "Benign familial hematuria due to mutation of the type IV collagen alpha4 gene." J Clin Invest **98**(5): 1114-1118.

Lemmink, H. H., C. H. Schroder, et al. (1997). "The clinical spectrum of type IV collagen mutations." Hum Mutat **9**(6): 477-499.

Lerone, M., P. Dodero, et al. (1991). "Leiomyomatosis of oesophagus, congenital cataracts and hematuria. Report of a case with rectal involvement." Pediatr Radiol **21**(8): 578-579.

Levey, A. S., J. P. Bosch, et al. (1999). "A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group." Ann Intern Med **130**(6): 461-470.

Levey, A. S., L. A. Stevens, et al. (2009). "A new equation to estimate glomerular filtration rate." Ann Intern Med **150**(9): 604-612.

Levy, M. and J. Feingold (2000). "Estimating prevalence in single-gene kidney diseases progressing to renal failure." Kidney Int **58**(3): 925-943.

Liu, Y. B., S. J. Tan, et al. (2008). "Clear lens phacoemulsification with continuous curvilinear capsulorhexis and foldable intraocular lens implantation for the treatment of a patient with bilateral anterior lenticonus due to Alport syndrome." J Int Med Res **36**(6): 1440-1444.

Longo, I., P. Porcedda, et al. (2002). "COL4A3/COL4A4 mutations: from familial hematuria to autosomal-dominant or recessive Alport syndrome." Kidney Int **61**(6): 1947-1956.

Marcocci, E., V. Uliana, et al. (2009). "Autosomal dominant Alport syndrome: molecular analysis of the COL4A4 gene and clinical outcome." Nephrol Dial Transplant **24**(5): 1464-1471.

Martin, P., N. Heiskari, et al. (1998). "High mutation detection rate in the COL4A5 collagen gene in suspected Alport syndrome using PCR and direct DNA sequencing." J Am Soc Nephrol **9**(12): 2291-2301.

Martin, P. H. and K. Tryggvason (2001). "Two novel alternatively spliced 9-bp exons in the COL4A5 gene." Pediatr Nephrol **16**(1): 41-44.

Massella, L., A. O. Muda, et al. (2010). "Cyclosporine A treatment in patients with Alport syndrome: a single-center experience." Pediatr Nephrol.

Mazzucco, G., P. Barsotti, et al. (1998). "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**(6): 1023-1031.

Mazzucco, G., P. Barsotti, et al. (1997). "Expression of alpha (IV) chains in Alport's syndrome and its correlation with ultrastructural and genetic data." Contrib Nephrol **122**: 129-131.

Mazzucco, G., M. De Marchi, et al. (2002). "Renal biopsy interpretation in Alport Syndrome." Semin Diagn Pathol **19**(3): 133-145.

McCarthy, P. A. and D. M. Maino (2000). "Alport syndrome: a review." Clin Eye Vis Care **12**(3-4): 139-150.

McCoy, R. C., H. K. Johnson, et al. (1982). "Absence of nephritogenic GBM antigen(s) in some patients with hereditary nephritis." Kidney Int **21**(4): 642-652.

McDonald, T. J., H. Zincke, et al. (1978). "Reversal of deafness after renal transplantation in Alport's syndrome." Laryngoscope **88**(1 Pt 1): 38-42.

Meloni, I., F. Vitelli, et al. (2002). "Alport syndrome and mental retardation: clinical and genetic dissection of the contiguous gene deletion syndrome in Xq22.3 (ATS-MR)." J Med Genet **39**(5): 359-365.

Milliner, D. S., A. M. Pierides, et al. (1982). "Renal transplantation in Alport's syndrome: anti-glomerular basement membrane glomerulonephritis in the allograft." Mayo Clin Proc **57**(1): 35-43.

Miner, J. H. (1999). "Alport syndrome with diffuse leiomyomatosis. When and when not?" Am J Pathol **154**(6): 1633-1635.

Miner, J. H. and J. R. Sanes (1994). "Collagen IV alpha 3, alpha 4, and alpha 5 chains in rodent basal laminae: sequence, distribution, association with laminins, and developmental switches." J Cell Biol **127**(3): 879-891.

Mitschke, H., P. Schmidt, et al. (1975). "Reversible uremic deafness after successful renal transplantation." N Engl J Med **292**(20): 1062-1063.

Mitschke, H., P. Schmidt, et al. (1977). "Effect of renal transplantation on uremic deafness: a long-term study." Audiology **16**(6): 530-534.

Mochizuki, T., H. H. Lemmink, et al. (1994). "Identification of mutations in the alpha 3(IV) and alpha 4(IV) collagen genes in autosomal recessive Alport syndrome." Nat Genet **8**(1): 77-81.

Mothes, H., L. Heidet, et al. (2002). "Alport syndrome associated with diffuse leiomyomatosis: COL4A5-COL4A6 deletion associated with a mild form of Alport nephropathy." Nephrol Dial Transplant **17**(1): 70-74.

Myers, G. J. and H. R. Tyler (1972). "The etiology of deafness in Alport's syndrome." Arch Otolaryngol **96**(4): 333-340.

Myers, J. C., T. A. Jones, et al. (1990). "Molecular cloning of alpha 5(IV) collagen and assignment of the gene to the region of the X chromosome containing the Alport syndrome locus." Am J Hum Genet **46**(6): 1024-1033.

Naito, I., S. Kawai, et al. (1996). "Relationship between COL4A5 gene mutation and distribution of type IV collagen in male X-linked Alport syndrome. Japanese Alport Network." Kidney Int **50**(1): 304-311.

Naito, I., S. Nomura, et al. (1997). "X-linked Alport syndrome with normal distribution of collagen IV alpha chains in epidermal basement membrane." Contrib Nephrol **122**: 134-139.

Ng, S. B., E. H. Turner, et al. (2009). "Targeted capture and massively parallel sequencing of 12 human exomes." Nature **461**(7261): 272-276.

Ninichuk, V., O. Gross, et al. (2006). "Multipotent mesenchymal stem cells reduce interstitial fibrosis but do not delay progression of chronic kidney disease in collagen4A3-deficient mice." Kidney Int **70**(1): 121-129.

Ninomiya, Y., M. Kagawa, et al. (1995). "Differential expression of two basement membrane collagen genes, COL4A6 and COL4A5, demonstrated by immunofluorescence staining using peptide-specific monoclonal antibodies." J Cell Biol **130**(5): 1219-1229.

Norby, S. M. and F. G. Cosio (2005). "Thin basement membrane nephropathy associated with other glomerular diseases." Semin Nephrol **25**(3): 176-179.

Olitsky, S. E., W. R. Waz, et al. (1999). "Rupture of the anterior lens capsule in Alport syndrome." J AAPOS **3**(6): 381-382.

Olson, D. L., S. K. Anand, et al. (1980). "Diagnosis of hereditary nephritis by failure of glomeruli to bind anti-glomerular basement membrane antibodies." J Pediatr **96**(4): 697-699.

Oohashi, T., I. Naito, et al. (2011). "Clonal overgrowth of esophageal smooth muscle cells in diffuse leiomyomatosis-Alport syndrome caused by partial deletion in COL4A5 and COL4A6 genes." Matrix Biol **30**(1): 3-8.

Pajari, H., H. Kaariainen, et al. (1996). "Alport's syndrome in 78 patients: epidemiological and clinical study." Acta Paediatr **85**(11): 1300-1306.

Peissel, B., L. Geng, et al. (1995). "Comparative distribution of the alpha 1(IV), alpha 5(IV), and alpha 6(IV) collagen chains in normal human adult and fetal tissues and in kidneys from X-linked Alport syndrome patients." J Clin Invest **96**(4): 1948-1957.

Perkoff, G. T., C. A. Nugent, Jr., et al. (1958). "A follow-up study of hereditary chronic nephritis." AMA Arch Intern Med **102**(5): 733-746.

Perkoff, G. T., F. E. Stephens, et al. (1951). "A clinical study of hereditary interstitial pyelonephritis." AMA Arch Intern Med **88**(2): 191-200.

Persson, U., J. M. Hertz, et al. (2005). "Alport syndrome in southern Sweden." Clin Nephrol **64**(2): 85-90.

Pierides, A., K. Voskarides, et al. (2009). "Clinico-pathological correlations in 127 patients in 11 large pedigrees, segregating one of three heterozygous mutations in the COL4A3/ COL4A4 genes associated with familial haematuria and significant late progression to proteinuria and chronic kidney disease from focal segmental glomerulosclerosis." Nephrol Dial Transplant **24**(9): 2721-2729.

Pihlajaniemi, T., E. R. Pohjolainen, et al. (1990). "Complete primary structure of the triple-helical region and the carboxyl-terminal domain of a new type IV collagen chain, alpha 5(IV)." J Biol Chem **265**(23): 13758-13766.

Pirson, Y. (1999). "Making the diagnosis of Alport's syndrome." Kidney Int **56**(2): 760-775.

Plant, K. E., E. Boye, et al. (2000). "Somatic mosaicism associated with a mild Alport syndrome phenotype." J Med Genet **37**(3): 238-239.

Plant, K. E., P. M. Green, et al. (1999). "Detection of mutations in COL4A5 in patients with Alport syndrome." Hum Mutat **13**(2): 124-132.

Prodromidi, E. I., R. Poulsom, et al. (2006). "Bone marrow-derived cells contribute to podocyte regeneration and amelioration of renal disease in a mouse model of Alport syndrome." Stem Cells **24**(11): 2448-2455.

Rabushka, L. S., E. K. Fishman, et al. (1991). "Diffuse esophageal leiomyomatosis in a patient with Alport syndrome: CT demonstration." Radiology **179**(1): 176-178.

Renieri, A., M. T. Bassi, et al. (1994). "Deletion spanning the 5' ends of both the COL4A5 and COL4A6 genes in a patient with Alport's syndrome and leiomyomatosis." Hum Mutat **4**(3): 195-198.

Renieri, A., M. Bruttini, et al. (1996). "X-linked Alport syndrome: an SSCP-based mutation survey over all 51 exons of the COL4A5 gene." Am J Hum Genet **58**(6): 1192-1204.

Reyersbach, G. C. and A. M. Butler (1954). "Congenital hereditary hematuria." N Engl J Med **251**(10): 377-380.

Reynolds, D. R. (1964). "HEREDITARY NEPHRITIS AND ITS OCULAR DEFECTS." Trans Pac Coast Otoophthalmol Soc Annu Meet **45**: 273-280.

Rheault, M. N. (2012). "Women and Alport syndrome." Pediatr Nephrol **27**(1): 41-46.

Rhys, C., B. Snyers, et al. (1997). "Recurrent corneal erosion associated with Alport's syndrome. Rapid communication." Kidney Int **52**(1): 208-211.

Rigden, S. P., O. Mehls, et al. (1996). "Report on management of renale failure in Europe, XXVI, 1995. The child-adult interface: a report on Alport's syndrome, 1975-1993. The ERA-EDTA Registry." Nephrol Dial Transplant **11 Suppl 7**: 21-27.

Risteli, J., H. P. Bachinger, et al. (1980). "7-S collagen: characterization of an unusual basement membrane structure." Eur J Biochem **108**(1): 239-250.

Rodriguez, J. D., S. S. Bhat, et al. (2010). "Intellectual disability, midface hypoplasia, facial hypotonia, and alport syndrome are associated with a deletion in Xq22.3." Am J Med Genet A **152A**(3): 713-717.

Rogers, P. W., N. A. Kurtzman, et al. (1973). "Familial benign essential hematuria." Arch Intern Med **131**(2): 257-262.

Rumpelt, H. J. (1980). "Hereditary nephropathy (Alport syndrome): correlation of clinical data with glomerular basement membrane alterations." Clin Nephrol **13**(5): 203-207.

Sa, M. J., N. Fieremans, et al. (2013). "Deletion of the 5'exons of COL4A6 is not needed for the development of diffuse leiomyomatosis in patients with Alport syndrome." J Med Genet **50**(11): 745-753.

Sabates, R., J. H. Krachmer, et al. (1983). "Ocular findings in Alport's syndrome." Ophthalmologica **186**(4): 204-210.

Sanger, F., S. Nicklen, et al. (1977). "DNA sequencing with chain-terminating inhibitors." Proc Natl Acad Sci U S A **74**(12): 5463-5467.

Sasaki, S., B. Zhou, et al. (1998). "Expression of mRNA for type IV collagen alpha1, alpha5 and alpha6 chains by cultured dermal fibroblasts from patients with X-linked Alport syndrome." Matrix Biol **17**(4): 279-291.

Savage, C. O., L. H. Noel, et al. (1989). "Hereditary nephritis: immunoblotting studies of the glomerular basement membrane." Lab Invest **60**(5): 613-618.

Savige, J., E. Ars, et al. (2013). "DNA variant databases improve test accuracy and phenotype prediction in Alport syndrome." Pediatr Nephrol.

Savige, J. and D. Colville (2009). "Opinion: Ocular features aid the diagnosis of Alport syndrome." Nat Rev Nephrol **5**(6): 356-360.

Savige, J., M. Gregory, et al. (2013). "Expert guidelines for the management of Alport syndrome and thin basement membrane nephropathy." J Am Soc Nephrol **24**(3): 364-375.

Savige, J., K. Rana, et al. (2003). "Thin basement membrane nephropathy." Kidney Int **64**(4): 1169-1178.

Schneider, R. G. (1963). "Congenital Hereditary Nephritis with Nerve Deafness." N Y State J Med **63**: 2644-2648.

Segal, Y., B. Peissel, et al. (1999). "LINE-1 elements at the sites of molecular rearrangements in Alport syndrome-diffuse leiomyomatosis." Am J Hum Genet **64**(1): 62-69.

Seremetis, M. G., W. S. Lyons, et al. (1976). "Leiomyomata of the esophagus. An analysis of 838 cases." Cancer **38**(5): 2166-2177.

Sessa, A., A. Pietrucci, et al. (1995). "Renal transplantation from living donor parents in two brothers with Alport syndrome. Can asymptomatic female carriers of the Alport gene be accepted as kidney donors?" Nephron **70**(1): 106-109.

Seymenoglu, G. and E. F. Baser (2009). "Ocular manifestations and surgical results in patients with Alport syndrome." J Cataract Refract Surg **35**(7): 1302-1306.

Shaw, E. A., D. Colville, et al. (2007). "Characterization of the peripheral retinopathy in X-linked and autosomal recessive Alport syndrome." Nephrol Dial Transplant **22**(1): 104-108.

Shoulders, M. D. and R. T. Raines (2009). "Collagen structure and stability." Annu Rev Biochem **78**: 929-958.

Slajpah, M., B. Gorinsek, et al. (2007). "Sixteen novel mutations identified in COL4A3, COL4A4, and COL4A5 genes in Slovenian families with Alport syndrome and benign familial hematuria." Kidney Int **71**(12): 1287-1295.

Snyers, B., C. Rhys, et al. (1996). "[Recurrent corneal erosions: new manifestation of juvenile alport syndrome]." Bull Soc Belge Ophtalmol **263**: 15-18.

Sohar, E. (1954). "A heredo-familial syndrome characterized by renal disease, inner ear deafness, and ocular changes." Harefuah **47**(8): 161-162.

Sohar, E. (1956). "Renal disease, inner ear deafness, and ocular changes; a new heredofamilial syndrome." AMA Arch Intern Med **97**(5): 627-630.

Sousa, R. G., P. C. Figueiredo, et al. (2013). "An unusual cause of pseudoachalasia: the Alport syndrome-diffuse leiomyomatosis association." Eur J Gastroenterol Hepatol.

Spear, G. S. (1973). "Editorial: Alport's syndrome: a consideration of pathogenesis." Clin Nephrol **1**(6): 336-337.

Spear, G. S. and R. J. Slusser (1972). "Alport's syndrome. Emphasizing electron microscopic studies of the glomerulus." Am J Pathol **69**(2): 213-224.

Spraul, C. W. and G. E. Lang (2000). "[Cone dystrophy associated with Alport syndrome]." Klin Monbl Augenheilkd **217**(3): 194-197.

Srivastava, A. K., S. McMillan, et al. (1999). "Integrated STS/YAC physical, genetic, and transcript map of human Xq21.3 to q23/q24 (DXS1203-DXS1059)." Genomics **58**(2): 188-201.

Stenson, P. D., M. Mort, et al. (2013). "The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine." Hum Genet.

Stephens, F. E., G. T. Perkoff, et al. (1951). "Partially sex-linked cominant inheritance of interstitial pyelonephritis." Am J Hum Genet **3**(4): 303-313.

Storey, H., J. Savige, et al. (2013). "COL4A3/COL4A4 Mutations and Features in Individuals with Autosomal Recessive Alport Syndrome." J Am Soc Nephrol **24**(12): 1945-1954.

Sugimoto, H., T. M. Mundel, et al. (2006). "Bone-marrow-derived stem cells repair basement membrane collagen defects and reverse genetic kidney disease." Proc Natl Acad Sci U S A **103**(19): 7321-7326.

Sugimoto, K., H. Yanagida, et al. (2005). "A Japanese family with Alport syndrome associated with esophageal leiomyomatosis: genetic analysis of COL4A5 to COL4A6 and immunostaining for type IV collagen subtypes." Clin Nephrol **64**(2): 144-150.

Sugimoto, M., T. Oohashi, et al. (1994). "The genes COL4A5 and COL4A6, coding for basement membrane collagen chains alpha 5(IV) and alpha 6(IV), are located head-to-head in close proximity on human chromosome Xq22 and COL4A6 is transcribed from two alternative promoters." Proc Natl Acad Sci U S A **91**(24): 11679-11683.

Sukhija, J., J. S. Saini, et al. (2003). "Phacoemulsification and intraocular lens implantation in an Alport's syndrome patient with bilateral anterior and posterior lenticonus." J Cataract Refract Surg **29**(9): 1834-1836.

Sund, M., Y. Maeshima, et al. (2005). "Bifunctional promoter of type IV collagen COL4A5 and COL4A6 genes regulates the expression of alpha5 and alpha6 chains in a distinct cell-specific fashion." Biochem J **387**(Pt 3): 755-761.

System, U. S. R. D., N. I. o. Health, et al. (2010). USRDS 2010 Annual Data Report: Atlas of Chronic Kidney Disease and End-Stage Renal Disease in the United States, Bethesda, MD.

Szpiro-Tapia, S., G. Bobrie, et al. (1988). "Linkage studies in X-linked Alport's syndrome." Hum Genet **81**(1): 85-87.

Tazon-Vega, B., E. Ars, et al. (2007). "Genetic testing for X-linked Alport syndrome by direct sequencing of COL4A5 cDNA from hair root RNA samples." Am J Kidney Dis **50**(2): 257 e251-214.

Tazon Vega, B., C. Badenas, et al. (2003). "Autosomal recessive Alport's syndrome and benign familial hematuria are collagen type IV diseases." Am J Kidney Dis **42**(5): 952-959.

Teekhasaenee, C., S. Nimmanit, et al. (1991). "Posterior polymorphous dystrophy and Alport syndrome." Ophthalmology **98**(8): 1207-1215.

Temme, J., F. Peters, et al. (2012). "Incidence of renal failure and nephroprotection by RAAS inhibition in heterozygous carriers of X-chromosomal and autosomal recessive Alport mutations." Kidney Int **81**(8): 779-783.

Thielen, B. K., D. F. Barker, et al. (2003). "Deletion mapping in Alport syndrome and Alport syndrome-diffuse leiomyomatosis reveals potential mechanisms of visceral smooth muscle overgrowth." Hum Mutat **22**(5): 419.

Thorner, P., L. Heidet, et al. (1998). "Diffuse leiomyomatosis of the esophagus: disorder of cell-matrix interaction?" Pediatr Dev Pathol **1**(6): 543-549.

Thorner, P. S., K. Zheng, et al. (1996). "Coordinate gene expression of the alpha3, alpha4, and alpha5 chains of collagen type IV. Evidence from a canine model of X-linked nephritis with a COL4A5 gene mutation." J Biol Chem **271**(23): 13821-13828.

Thusberg, J., A. Olatubosun, et al. (2011). "Performance of mutation pathogenicity prediction methods on missense variants." Hum Mutat **32**(4): 358-368.

Tiebosch, A. T., P. M. Frederik, et al. (1989). "Thin-basement-membrane nephropathy in adults with persistent hematuria." N Engl J Med **320**(1): 14-18.

Timpl, R., J. Risteli, et al. (1979). "Identification of a new basement membrane collagen by the aid of a large fragment resistant to bacterial collagenase." FEBS Lett **101**(2): 265-268.

Tryggvason, K., P. Heikkila, et al. (1997). "Can Alport syndrome be treated by gene therapy?" Kidney Int **51**(5): 1493-1499.

Tryggvason, K. and J. Patrakka (2009). Alport's disease and thin basement membrane nephropathy. Genetic Diseases of the Kidney. R. P. Lifton, S. Somlo, G. H. Giebisch and D. W. Seldin. Burlington (MA), USA, Academic Press, Elsevier Inc.**:** 77-96.

Uberbacher, E. C., D. Hyatt, et al. (2004). "GrailEXP and Genome Analysis Pipeline for genome annotation." Curr Protoc Hum Genet **Chapter 6**: Unit 6 5.

Ueki, Y., I. Naito, et al. (1998). "Topoisomerase I and II consensus sequences in a 17-kb deletion junction of the COL4A5 and COL4A6 genes and immunohistochemical analysis of esophageal leiomyomatosis associated with Alport syndrome." Am J Hum Genet **62**(2): 253-261.

Uliana, V., E. Marcocci, et al. (2011). "Alport syndrome and leiomyomatosis: the first deletion extending beyond COL4A6 intron 2." Pediatr Nephrol **26**(5): 717-724.

Unger, H. H. and K. Rother (1964). "[RETINAL CHANGES IN ALPORT'S SYNDROME (HEREDITARY NEPHROPATHY WITH INNER EAR DEAFNESS).]." Ber Zusammenkunft Dtsch Ophthalmol Ges **65**: 293-298.

Van Agtmael, T. and L. Bruckner-Tuderman (2010). "Basement membranes and human disease." Cell Tissue Res **339**(1): 167-188.

van der Loop, F. T., L. Heidet, et al. (2000). "Autosomal dominant Alport syndrome caused by a COL4A3 splice site mutation." Kidney Int **58**(5): 1870-1875.

Van Loo, A., R. Vanholder, et al. (1997). "Alport syndrome and diffuse leiomyomatosis with major morbid events presenting at adult age." Nephrol Dial Transplant **12**(4): 776-780.

van Paassen, P., P. J. van Breda Vriesman, et al. (2004). "Signs and symptoms of thin basement membrane nephropathy: a prospective regional study on primary glomerular disease-The Limburg Renal Registry." Kidney Int **66**(3): 909-913.

Vandenberg, P., A. Kern, et al. (1991). "Characterization of a type IV collagen major cell binding site with affinity to the alpha 1 beta 1 and the alpha 2 beta 1 integrins." J Cell Biol **113**(6): 1475-1483.

Vandewalle, J., H. Van Esch, et al. (2009). "Dosage-dependent severity of the phenotype in patients with mental retardation due to a recurrent copy-number gain at Xq28 mediated by an unusual recombination." Am J Hum Genet **85**(6): 809-822.

vd Heuvel, L. P., C. H. Schroder, et al. (1989). "The development of anti-glomerular basement membrane nephritis in two children with Alport's syndrome after renal transplantation: characterization of the antibody target." Pediatr Nephrol **3**(4): 406-413.

Vedantham, V., J. Rajagopal, et al. (2005). "Bilateral simultaneous anterior and posterior lenticonus in Alport's syndrome." Indian J Ophthalmol **53**(3): 212-213.

Vetrie, D., E. Boye, et al. (1992). "DNA rearrangements in the alpha 5(IV) collagen gene (COL4A5) of individuals with Alport syndrome: further refinement using pulsed-field gel electrophoresis." Genomics **14**(3): 624-633.

Vivante, A., A. Afek, et al. (2011). "Persistent asymptomatic isolated microscopic hematuria in Israeli adolescents and young adults and risk for end-stage renal disease." JAMA **306**(7): 729-736.

Vivante, A., R. Calderon-Margalit, et al. (2013). "Hematuria and risk for end-stage kidney disease." Curr Opin Nephrol Hypertens **22**(3): 325-330.

Voskarides, K., L. Damianou, et al. (2007). "COL4A3/COL4A4 mutations producing focal segmental glomerulosclerosis and renal failure in thin basement membrane nephropathy." J Am Soc Nephrol **18**(11): 3004-3016.

Wang, F., J. Ding, et al. (2002). "Phenotypic and genotypic features of Alport syndrome in Chinese children." Pediatr Nephrol **17**(12): 1013-1020.

Wang, F., Y. Wang, et al. (2005). "Detection of mutations in the COL4A5 gene by analyzing cDNA of skin fibroblasts." Kidney Int **67**(4): 1268-1274.

Wang, F., D. Zhao, et al. (2012). "Skin biopsy is a practical approach for the clinical diagnosis and molecular genetic analysis of X-linked Alport's syndrome." J Mol Diagn **14**(6): 586-593.

Wang, Y. F., J. Ding, et al. (2004). "Effect of glycine substitutions on alpha5(IV) chain structure and structure-phenotype correlations in Alport syndrome." Biochem Biophys Res Commun **316**(4): 1143-1149.

Wang, Y. Y., K. Rana, et al. (2004). "COL4A3 mutations and their clinical consequences in thin basement membrane nephropathy (TBMN)." Kidney Int **65**(3): 786-790.

Wieslander, J., J. F. Barr, et al. (1984). "Goodpasture antigen of the glomerular basement membrane: localization to noncollagenous regions of type IV collagen." Proc Natl Acad Sci U S A **81**(12): 3838-3842.

Williamson, D. A. (1961). "Alport's syndrome of hereditary nephritis with deafness." Lancet **2**(7216): 1321-1323.

Wuhl, E., O. Mehls, et al. (2004). "Antihypertensive and antiproteinuric efficacy of ramipril in children with chronic renal failure." Kidney Int **66**(2): 768-776.

Wuhl, E., A. Trivelli, et al. (2009). "Strict blood-pressure control and progression of renal failure in children." N Engl J Med **361**(17): 1639-1650.

Xu, J. M., S. S. Zhang, et al. (2010). "Ocular manifestations of Alport syndrome." Int J Ophthalmol **3**(2): 149-151.

Yoshikawa, N., S. Matsuyama, et al. (1988). "Benign familial hematuria." Arch Pathol Lab Med **112**(8): 794-797.

Yoshioka, K., S. Hino, et al. (1994). "Type IV collagen alpha 5 chain. Normal distribution and abnormalities in X-linked Alport syndrome revealed by monoclonal antibody." Am J Pathol **144**(5): 986-996.

Zare, M. A., M. T. Rajabi, et al. (2007). "Phacoemulsification and intraocular lens implantation in Alport syndrome with anterior lenticonus." J Cataract Refract Surg **33**(6): 1127-1130.

Zhang, K. W., D. Colville, et al. (2008). "The use of ocular abnormalities to diagnose X-linked Alport syndrome in children." Pediatr Nephrol **23**(8): 1245-1250.

Zheng, K., S. Harvey, et al. (1999). "Absence of the alpha6(IV) chain of collagen type IV in Alport syndrome is related to a failure at the protein assembly level and does not result in diffuse leiomyomatosis." Am J Pathol **154**(6): 1883-1891.

Zhou, J., M. Ding, et al. (1994). "Complete primary structure of the sixth chain of human basement membrane collagen, alpha 6(IV). Isolation of the cDNAs for alpha 6(IV) and comparison with five other type IV collagen chains." J Biol Chem **269**(18): 13193-13199.

Zhou, J., T. Mochizuki, et al. (1993). "Deletion of the paired alpha 5(IV) and alpha 6(IV) collagen genes in inherited smooth muscle tumors." Science **261**(5125): 1167-1169.

Protocol P1 *Clinical protocol*

Médico: _____ **Contacto(s):**

_____

Hospital, Serviço / Unidade de Saúde:

_____

Identificação do paciente:

Caso-índice ☐

Familiar afectado ☐

Familiar não-afectado ☐

Nome: _____

Data de Nascimento: ____ / ____ / _____

Sexo: Masculino / Feminino

(ou colar etiqueta)

Manifestações de Síndrome de Alport e sua cronologia:

Assinale "**[S]**" para indicar as manifestações de Síndrome de Alport observadas no doente e a idade aproximada em que foram identificadas; "**[N]**" para indicar as manifestações que foram excluídas por rastreio específico; "**[I]**" para indicar as manifestações que ignore que o doente apresente ou que não tenham sido especificamente rastreadas.

1 - Hematúria: macroscópica **[S]** (idade) / **[N]** / **[I]** ; microscópica **[S]** (idade) / **[N]** / **[I]** .

2 - Proteinúria: **[S]** (idade) / **[N]** / **[I]** ; grau máximo de proteinúria: ≤0,5g / 0,5-1,0g / 1,0-3,5g / ≥3,5g (data: mês / ano) .
[Proteinúria definida como ≥0,3g/dia, expressa em g/dia ou $g/g_{creatinina}$].

3 - Hipertensão arterial: **[S]** (idade) / **[N]** / **[I]** ; medicação anti-hipertensiva: **[S]** (idade) / **[N]** / **[I]** .
[Hipertensão arterial definida como ≥140/90 mmHg].

4 - Insuficiência renal crónica (IRC): **[S]** (idade) / **[N]** / **[I]** ; doença renal crónica, estádios 3-5: **[S]** (idade) / **[N]** / **[I]** .
[Insuficiência renal crónica definida como creatinina sérica ≥1,2 mg/dl no homem e ≥0,9 mg/dl na mulher; doença renal crónica, estádios 3-5, definida como creatinina sérica >1,5 mg/dl no homem e >1,2 mg/dl na mulher].

5 - Tratamento de substituição de função renal: **[S]** (idade) / **[N]** / **[I]** ; modalidade inicial: HD / DP / TR .
[HD – Hemodiálise / DR – Diálise Peritoneal / TR – Transplante Renal].

6 - Surdez neuro-sensorial bilateral: **[S]** (idade) / **[N]** / **[I]** ; audiograma mais recente (data: mês / ano) .

7 - Sinais oculares: maculopatia **[S]** (idade) / **[N]** / **[I]** ; lenticone anterior **[S]** (idade) / **[N]** / **[I]** ; outro(s): _____ ; exame oftalmológico mais recente (data: mês / ano) .

8 - Leiomiomatose: **[S]** (idade) / **[N]** / **[I]** ; localização: _____

9 - Dificuldades de aprendizagem / atraso mental: **[S]** (idade) / **[N]** / **[I]** ; dismorfia facial **[S]** (idade) / **[N]** / **[I]** .

10 - Biopsia renal: **[S]** (idade) / **[N]** ; microscopia óptica: _____

microscopia electrónica: **[S]** / **[N]** ; alterações da membrana basal glomerular – estreitamento: **[S]** / **[N]** / **[I]**; espessamento: **[S]** / **[N]** / **[I]**; lamelação: **[S]** / **[N]** / **[I]**; corpos electrodensos intralacunares **[S]** / **[N]** / **[I]**.

História Familiar:

História familiar de hematúria: **[S]** / **[N]** / **[I]** Existe mutação familiar identificada? **[S]** / **[N]**

História familiar de IRC : **[S]** / **[N]** / **[I]** Em que gene? ☐ *COL4A5* ☐ *COL4A*☐ *COL4A3*

História familiar de surdez: **[S]** / **[N]** / **[I]** Descrição da mutação: _____

Árvore Familiar (inclua, pelo menos, 3 geraç ões)

Legenda:

Requisição de Análise Genética

Produto biológico enviado:

Sangue ☐ ADN ☐ Data da Colheita:_____/_____/_____

Assinatura do Médico: _____ (Cédula Profissional: _____)

Data: _____ / _____ / _____

Protocol P2 *Molecular protocol*

SEQUENCIAÇÃO DIRECTA DO GENE *COL4A5*

A sequenciação directa do gene *COL4A5* tem como objectivo identificar mutações pontuais e grandes delecções, em doentes do sexo masculino. Nos doentes do sexo feminino, não é tão informativa, permitindo apenas detectar mutações pontuais. Nestes casos, deve ser adicionalmente efectuada a técnica *Multiplex-Ligation dependent Probe Amplification* (MLPA), que permite a identificação de delecções e duplicações.

O estudo de um probando inclui a sequenciação directa da região codificante de 53 exões do gene *COL4A5* (incluindo os exões sujeitos a *splicing* alternativo 41A e 41B), assim como das regiões de transição exão-intrão. Este estudo não inclui a sequenciação da região promotora comum aos genes *COL4A5* e *COL4A6*. As sequências de *primers* (iniciadores) e condições de amplificação e sequenciação do ADN genómico foram adaptadas do estudo efectuado por Martin *et al.* (1998) (*Supplementary table S2* do manuscript 1). No caso de mutações pontuais, a variante é confirmada em dois PCR independentes, incluindo, sempre que possível um controlo positivo (familiar afectado) e um controlo negativo (controlo saudável). No caso das delecções, a variante é confirmada em indivíduos do sexo masculino através de duas técnicas (sequenciação e MLPA). O estudo de uma variante patogénica em familiares do probando deverá ser efectuado através da análise do exão (ou exões) relevante(s), por sequenciação directa e/ou MLPA.

I. Colheita e envio da amostra

<u>Instruções para colheita da amostra:</u>

Programar a colheita da amostra de sangue de modo a que possa ser enviada por correio e chegar ao Laboratório do Departamento de Genética dentro de 48 horas.

Colher 5-10 ml de sangue venoso em tubo(s) de plástico com K-EDTA (Vacutainer® ou análogo – tampa roxa).

Se o doente estiver em programa de hemodiálise, a colheita deverá ser feita antes da heparinização inicial.

Instruções para envio da amostra:

Manter e enviar o tubo à temperatura ambiente.

Colocar o(s) tubo(s) de plástico, devidamente identificados, no interior de envelope almofadado, endereçado a:

Departamento de Genética
Faculdade de Medicina da Universidade do Porto
Alameda Prof. Hernâni Monteiro
4200 – 319 Porto

II. Extracção do ADN e armazenamento da amostra

Extrair ADN genómico a partir de linfócitos de sangue venoso, usando um método de *salting-out* (Citomed, Lisboa), e armazenar a 4ºC. Armazenar parte do pellet de linfócitos a -80ºC, que pode ser utilizado para extrair ADN ou mARN.

III. Quantificação da concentração de DNA de cadeia dupla da amostra

As amostras de ADN foram quantificadas num espetrofotómetro NanoDrop2000C (ThermoScientific). A medição da absorvância no comprimento de onda de 260 nm permitiu calcular a concentração de ADN em ng/µL.

IV. Amplificação do ADN genómico por *Polymerase Chain Reaction* (PCR)

A mistura para a reação de amplificação do ADN genómino por PCR tem um volume de 50 µl, contendo: tampão *Fermentas*® (1x), $MgCl_2$ *Fermentas*® (1,5mM), dNTPs (0,2mM), primers *forward* e *reverse* (5 pmol/µl), Taq DNA polymerase *Fermentas*® (5U/µl), água destilada e ADN (100ng/µl). O ADN genómico é amplificado *in vitro,* usando os termocicladores GeneAmp PCR System 9700 ou Veriti Thermal Cycler (Applied Biosystems), do seguinte modo:

1. Desnaturação inicial a 94ºC, durante 5 minutos.
2. Repetição de 35 ciclos com as seguintes condições:
    i. Desnaturação da dupla cadeia de ADN a 94ºC, durante 1 minuto.

ii. Emparelhamento dos *primers* à temperatura definida, durante 1 minuto.
iii. Extensão da dupla cadeia emergente de ADN a 72ºC, durante 1 minuto e 30 segundos.

3. Extensão final das duplas cadeia de DNA a 72ºC, durante 10 minutos.
4. Manutenção do produto da amplificação a 4ºC, após a conclusão do número total de ciclos de PCR.

A confirmação e a determinação dos tamanhos dos produtos amplificados foram realizadas através de eletroforese capilar no aparelho QIAxcel, com o *kit* QIAxel DNA *Screening* e com o *software* BioCalculator™ (Qiagen).

V. Sequenciação directa de produtos de PCR

A. Purificação do ADN

O ADN amplificado pode ser purificado usando o kit comercial *illustra GFX PCR DNA and Gel Band Purification Kit* (GE Healthcare®), do seguinte modo:

1. Captura do produto de PCR
    a. Colocar a coluna num tubo de recolha (2 ml) e identificá-lo.
    b. Pipetar 500 µl de Capture buffer type 3 para a coluna.
    c. Transferir o produto de PCR (50 µl) para a coluna, misturando o produto com o Capture buffer type 3 vigorosamente até obter uma solução homogénea.
2. Ligação do produto de PCR
    a. Centrifugar a mistura durante 30 segundos, a 14000 r.p.m.
    b. Desperdiçar o centrifugado e reutilizar o tubo de recolha com a coluna.
3. Lavar e secar
    a. Pipetar 500 µl de Wash buffer para a coluna.
    b. Centrifugar a mistura durante 30 segundos, a 14000 r.p.m.
    c. Transferir a coluna para um tubo de 1,5 ml.
4. Eluição
    a. Pipetar 30 µl de Elution buffer type 6 para a coluna.
    b. Incubar o produto de PCR na coluna à temperatura ambiente, durante 1 minuto.

c. Centrifugar a mistura durante 1 minuto, a 14000 r.p.m.
d. Desperdiçar a coluna e guardar o ADN purificado a – 20ºC.

O ADN purificado pode ser utilizado para sequenciação directa.

B. Reacção de Sequenciação:

A preparação da mistura para a reacção de sequenciação é realizada com um suporte colocado sobre gelo e usando o BigDye® Terminator v3.1 Cycle Sequencing kit (Life Technologies; Carlsbad, CA, USA), seguindo o seguinte procedimento:

.

1. Preparar uma mistura para cada amostra, adicionando 2µl ADN purificado, 2µl Premix, 1µl Buffer, 0,2µl primer *foward* ou *reverse* (10pmol/µl) e 4,8µl com água esterilizada.
2. Realizar a reacção de sequenciação com as seguintes condições:
    a. Desnaturação inicial a 96ºC, durante 3 minutos.
    b. Repetir 24 ciclos:
        i. Desnaturação da dupla cadeia a 96ºC, durante 10 segundos.
        ii. Emparelhamento do primer a 50ºC, durante 5 segundos.
        iii. Extensão da cadeia simples emergente a 60ºC, durante 4 minutos.
    c. Manter as amostras a 4ºC, após a conclusão do número total de ciclos de PCR.

C. Precipitação dos produtos:

A preparação da mistura é realizada à temperatura ambiente, sem gelo, do seguinte modo:

1. Preparar uma mix para cada amostra, adicionando 25µl de etanol 100%, 1µl EDTA (125mM) e 1µl acetato de sódio (3M), todos previamente armazenados a 4ºC.
2. Adicionar 27µl da mix a cada amostra a sequenciar e incubar a 4ºC durante 25 minutos ou à temperatura ambiente durante 30 minutos.
3. Centrifugar a 14.000rpm, a 4ºC, durante 30 minutos.

4. Retirar o sobrenadante sem tocar no *pellet*.
5. Adicionar 100µl etanol 70%.
6. Centrifugar a 14.000rpm, a 4ºC, durante 20 minutos.
7. Retirar o sobrenadante sem tocar no *pellet* e secar o *pellet*, a 90ºC, durante 1 minuto.
8. Guardar a -20ºC.

D. Antes de sequenciar:

1. Adicionar 20µl formamida ao *pellet*.
2. Incubar à temperatura ambiente, protegido da luz com papel de alumínio, durante 30 minutos.
3. Colocar 21µl do produto a sequenciar em tubos de sequenciação, após efectuar vortex e short spin do produto incubado.
4. Desnaturar o ADN a sequenciar a 95ºC, durante 5 minutos, seguidos de conservação da amostra em gelo, durante 5 minutos.

E. Sequenciação automática:

1. Efectuar a sequenciação automática dos produtos amplificados e purificados, nas direcções *foward* e/ou *reverse*, através de um serviço de *outsourcing* (STABVIDA®) ou no Departamento de Genética, da Faculdade de Medicina da Universidade do Porto, usando os sequenciadores automáticos ABI Prism® 310 ou 3500 Genetic Analyzer (Applied Biosystems; Foster City, CA, USA).
2. Analisar os electroferogramas visualmente, usando o so*ftware* ABI s*equenci*ng Analysis v.5.0 (Applied *Biosystems,* USA).
3. Comparar a sequência da amostra sequenciada com controlos, recorrendo ao software Blast®, para detecção de variantes patogénicas (http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastSearch&VIEW_SEARCH=on&UNIQ_SEARCH_NAME=A_SearchOptions_1PvEPK_37vW_DP1lHXE04Xw_GTR6V_2FOw0G#).

Index

List of figures

List of tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Appendix I *Clinical and molecular characterization of patients (CD)*