American Journal of Kidney Disease

COL4A5 Founder Mutation Identified in Three Families Leads to an Unusual X-linked Glomerulopathy

--Manuscript Draft--

Manuscript Number:	
Article Type:	Case Report
Full Title:	COL4A5 Founder Mutation Identified in Three Families Leads to an Unusual X-linked Glomerulopathy
Order of Authors:	Moumita Barua, M.D.
	Weili Li
	Lorenzo Stella
	Nicole Roslin
	Saidah Hack
	Bedra Sharif
	Ginette Lajoie-Starkell
	Rohan John
	Andrew Schwaderer
	Matthias Wuttke
	Anna Kottgen
	Daniel Cattran
	Andrew Paterson
	York Pei
Suggested Reviewers:	Jeff Miner
Opposed Reviewers:	

COL4A5 Founder Mutation Identified in Three Families Leads to an Unusual X-linked Glomerulopathy

Moumita Barua MD[1,2,3,4], Weili Li PhD[5], Lorenzo Stella PhD[6], Rohan John MD[7], Nicole Roslin PhD[5], Bedra Sharif PhD[1], Saidah Hack BSc[1], Ginette Lajoie-Starkell MD[8], Andrew L. Schwaderer MD[9], Brian Becknell MD PhD[9], Matthias Wuttke MD[10, 11], Anna Kottgen MD[10, 11], Daniel Cattran MD[1,2,3], *Andrew D. Paterson MD[4,5,12], *York Pei MD[1,2,3,4]

[1]Division of Nephrology, University Health Network, Toronto, Canada

[2]Toronto General Research Institute, Toronto General Hospital, Toronto, Canada

[3]Department of Medicine, University of Toronto, Canada

[4]Institute of Medical Sciences, University of Toronto, Canada

[5]Hospital for Sick Children's, Toronto, Canada

[6]Department of Chemical Science and Technology, University of Rome Tor Vergata, Rome, Italy

[7]Department of Pathology, University Health Network, Toronto, Canada

[8]William Osler Health Center, Toronto, Canada

[9]Nephrology Section, Nationwide Children's Hospital Department of Pediatrics, Ohio State University Center for Clinical and Translational Research, The Research Institute at Nationwide Children's Hospital

[10]Division of Genetic Epidemiology, Institute for Medical Biometry and Statistics, Medical Center – University of Freiburg, Faculty of Medicine, Freiburg, Germany

[11]Division of Nephrology, University of Freiburg, Faculty of Medicine, Freiburg, Germany

[12]Epidemiology & Biostatistics, Dalla Lana School of Public Health, University of Toronto

[*]These authors contributed equally to this work

Correspondence:

Dr. Moumita Barua

200 Elizabeth Street, 8NU-855

Toronto General Hospital

Toronto, ON

M5G 2C4

moumita.barua@uhn.ca

Abstract word count: 179

Manuscript word count: 1494

Running header: *COL4A5* Founder Mutation Leads to Unusual Glomerulopathy

Abstract:

Alport syndrome (AS) is a rare hereditary disorder caused by mutations in one of three genes encoding for type IV collagen. Mutations in *COL4A5* on chr.Xq22 cause X-linked AS (XLAS), which accounts for ~80% of the cases. AS has a variable clinical presentation including progressive renal failure, hearing loss and ocular defects. Exome sequencing performed in two affected related males with an undefined X-linked glomerulopathy, characterized by global and segmental glomerulosclerosis, mesangial hypercellularity and vague basement membrane immune complex deposition, revealed a *COL4A5* sequence variant, c.T665G, p.Phe222Cys (NM_000495, rs281874761), not seen in databases cataloguing natural human genetic variation including dbSNP138, 1000 Genomes Project release version 01-11-2004, Exome Sequencing Project 21-06-2014 or ExAC 01-11-2014. Review of the literature identified two additional families with the same *COL4A5* variant leading to similar atypical histopathological features, suggesting a unique pathologic mechanism initiated by this specific mutation. Homology modeling suggests that the substitution will alter the structural and dynamic properties of the COL4A trimer. Genetic analysis comparing members of the three families indicated a distant relationship with a shared haplotype implying a founder effect.

Index words: *COL4A5* mutations, X-link inheritance, founder mutation, atypical phenotype

Introduction

Alport syndrome (AS) is a rare hereditary disorder, with an estimated prevalence of 1 in 50,000, caused by mutations in one of three genes coding for Type IV collagen.[1] AS has a variable clinical presentation including progressive renal failure, hearing loss and ocular defects. It is characterized by thinning basement membranes and a hallmark degenerative glomerular BM splitting. Type IV collagen (COL4) is composed of six genetically distinct α-chains (α1-α6) arranged in three distinct triplet helical protomers (i.e. α1-α1-α2, α3-α4-α5, and α5-α5-α6). In the kidney, the α5α5α6 network is found in Bowman's capsule.[2]

Mutations in *COL4A5* on chrXq22 cause X-linked AS (XLAS) which accounts for ~80% of the cases, whereas recessive and dominant mutations of *COL4A3* or *COL4A4* on chr2q36-q37 and chr2q35-37, respectively, account for ~15% and ~5% of the remaining cases.[1] Certain specific variants of any of the above three genes interfere with the assembly of the α3-α4-α5 (IV) network in the glomerular BM (GBM) and arrest the normal developmental switch from the immature α1-α1-α2 (IV) network to the α3-α4-α5 (IV) network.[3, 4] The persistence of an immature α1-α1-α2 (IV) network is more susceptible to proteolysis and eventually leads to the "hallmark" degenerative glomerular BM splitting on pathology.

We report a family with an X-linked recessive inheritance pattern of renal disease but with features atypical for any single diagnosis. Genetic analysis reveals a mutation in *COL4A5*, p.Phe222Cys (NM_000495, rs281874761), which was previously reported in two families with similar atypical histopathological characteristics.[5, 6] Homology modeling predicts that the substitution will alter the structural and dynamic properties of the COL4A trimer. Further

investigations of these families demonstrate a distant relationship, suggesting that the mutations arose on the same haplotype background indicative of a founder effect.

Case Report

This study was approved by the University Health Network Research Ethics Board. The clinical course of the proband (6442) is characterized by recurrent tonsillitis, hematuria and proteinuria up to 10 g/day since he was 20 years of age (Figure 1). A renal biopsy was performed revealing global and segmental glomerulosclerosis with mesangial hypercellularity and glomerular basement membrane (GBM) thickening by light microscopy (LM) and a homogeneously thickened GBM with scattered intramembranous densities by electron microscopy (EM) (Figure 1). Immunofluorescent (IF) microscopy showed some linear GBM staining for IgG and IgA and granular GBM and mesangial staining for IgM (Figure 1). The proband was treated with an ACE inhibitor and a course of steroids with partial remission of his proteinuria (~2 g/day) at the last follow-up at age 22.

Interestingly, review of the renal biopsy report from the proband's maternal uncle, 6444, performed in 1969 at the age of 18 years for proteinuria described similar finding on light microscopy – GBM thickening with some fuchsinophilic outer capillary wall deposits. Immunofluorescence or electron microscopy was not done. His progressed to end-stage renal disease (ESRD) by his late 40s.

The proband's maternal grandmother, 6465, did not have a significant renal history until she was diagnosed with high-grade urothelial cancer at age 78. Her father was reported to have passed

away from kidney failure in his early 20s, and had a brother also affected by renal disease. Two male grandchildren of this latter individual developed kidney failure: 6471, was the recipient of a live donor renal transplant at the age of 25 years from his sister. His brother, 6470, had up to approximately 800 mg of protein excretion in a 24 hour period and developed ESRD at the age of 48 years.

The proband's mother and aunt, 6443 and 6533, had normal urinalysis in 2013 and 2014, respectively, in the 4^{th} and 5^{th} decades of life. Their sisters, 6596 and 6599, were reported to not have kidney disease. The son to 6596 (7435) was identified to have microscopic hematuria and proteinuria in his teenage years.

The absence of disease in four putative obligate female carriers (i.e. 6443, 6465, 6596, and 6466) in this family suggested an X-linked recessive inheritance. We performed parametric multipoint linkage analysis on 7 informative individuals in the pedigree, using genotype data generated with the Illumina HumanCoreExome-12 v1.0 array. The linkage analysis was conducted under a rare X-linked recessive model, and demonstrated a 14.35 cM region on the X-chromosome with a peak LOD score of 1.8, from rs10126713 to rs5911135 (Figure 1). Simulations based on the same family structure, inheritance model and linkage parameter values suggested that the maximum attainable X-chromosome LOD score was 1.8.

Genotyped individuals clustered with HapMap Caucasian European (CEU) samples.[7] A shared missense variant within the *COL4A5* gene, c.T665G, p.F222C, within the linked region was identified in exome sequence data (Figure 1). This mutation was not found in dbSNP138, 1000

Genomes Project release version 01-11-2004, Exome Sequencing Project 21-06-2014 or ExAC 01-11-2014.

In 2011, a family from the United States (U.S.) was reported as having similar renal histopathological characteristics.[6] The X-linked recessive inheritance pattern prompted Sanger sequencing of *COL4A5*, revealing the same variant, c.T665G, p.Phe222Cys.[6] In 2015, a second family from Germany with affected male cousins was reported to have the same sequence variant.[5] We genotyped an affected male (6692) from the 2011 report using Illumina HumanCoreExome-12v1-1A. Relationship checking comparing mitochondrial genotypes between 6692 and 6442/6470 from our pedigree did not reveal close maternal relatedness. However, a shared haplotype spanning physical distance of 18.5 Mb (genetic distance of 17.4 cM) between coordinates 97,025,440 (rs5921899) and 115,480,782 (rs11091077) containing 1620 SNPs on the X-chromosome between our proband (6442) and an affected individual from the U.S. family (6692) was identified.[8-10]

We obtained X chromosome exome sequence variant data belonging to the two affected males (335_2, 335_5) from the German family reported in 2015 and compared to sequence data from the proband belonging to our pedigree (6442).[5] This analysis revealed a shared 14.5 Mb haplotype around *COL4A5* between coordinates 103,267,865 (rs553509) and 117,718,760 (rs1781090). The shared haplotype amongst the four affected individuals, at least one from each of the three families (6442, 6692, 335_2, 335_5), was chr X: 103, 267, 865 and 115, 480, 782 around *COL4A5*.

To analyze the possible effect of the p.Phe222Cys substitution, we created a homology model of part of the α3α4α5 heterotrimer (Figure 2).[11-16] Collagen triple helical regions have a repetitive sequence with a Gly-Xaa-Yaa motif, Xaa and Yaa being any amino acids, but COL4A5 contains 22 interruptions, which are sites of increased flexibility, and are important for the network formed in GBM.[2, 15-19] The p.Phe222Cys mutation is located in one of them (GLN**F**QG). In the model, Phe222 forms a hydrophobic cluster with the corresponding residues in the other strands, as previously observed in other interruptions, where these inter-chain interactions stabilize the triple helix, and can favor the assembly of the trimer in the correct register.[15, 17, 20] Substitution of Cys would remove this stabilizing interaction, thus changing the structural and dynamical properties of the trimer. The mutation could also perturb protein-protein interactions (heat shock protein 47 binds at the mutation site).[2] Alternatively, the introduction of a Cys residue might cause the formation of an aberrant S-S bridge (for instance with Cys266 in α4).[21]

Discussion

Genetic analysis performed in two affected related males with an unusual X-linked glomerulopathy, characterized by global and segmental glomerulosclerosis, mesangial expansion and vague basement membrane immune complex deposition, revealed a rare *COL4A5* mutation, c.T665G, p.Phe222Cys. Review of the literature identified two families with the same *COL4A5* variant and showing similar atypical histopathological features, suggesting a unique pathologic mechanism initiated by this specific mutation. Mitochondrial haplogroups differed and no evidence for relationship between the pedigrees could be determined using autosome-wide SNP data. However, analysis of shared variants around the mutation among affected members of the

three families indicates a distant relationship and that the mutations arose on the same haplotype background suggestive of a founder effect.

We speculate that the substitution of phenylalanine with cysteine residue within the non-collagenous interruption of the α5(IV) chain may lead to a new epitope that elicits some sort of an immune response leading to vague intramembranous deposits. Interestingly, the proband in this series had a partial response to prednisone therapy. Modeling suggests that the substitution of phenylalanine with cysteine within the repeat interruption of GLY-X-Y will alter the structural and dynamic properties of the col4a trimer. Out of more than 300 missense variants reported in the Alport syndrome database (http://www.arup.utah.edu/database/alport/ALPORT_welcome.php, accessed October 4, 2016), p.Phe222Cys remains the only pathogenic missense mutation that occurs within a non-collagenous interruption.[22] It is thus conceivable that the structural and dynamic effects of this substitution might lead to a different Alport phenotype.

This report on XLAS with a *COL4A5* p.Phe222Cys variant illustrates the specific influence that gene mutation can have on phenotype. Homology modeling suggests that the substitution will change the structural and dynamic properties of the COL4 α3α4α5 heterotrimer, though how this leads to the unusual pathology observed is still unclear. While this is the third report of a family with this variant, our genetic analysis indicates that affected individuals across the pedigrees are descendants of a common founder, arguing against a mutational hotspot in the non-collagenous interruption.

ACKNOWLEDGEMENTS

M.B. and D.C. received a Health Research Grant (14-04) from Physician's Services Incorporated which funded this work. M.B. is supported by a new investigator award from the Kidney Research Scientist Core Education and National Training Program (KRESCENT). M.W. and A.K. were supported by the German Research Foundation (3598/3-1 and CRC 1140). Y.P. was supported by a grant from the Kidney Foundation of Canada. The authors wish to thank the patients and their families for their participation. The authors also wish to thank Sergio Pereira, Daniel Merico, Wilson Sung and Bhooma Thiruvahindrapuram of The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Canada for assistance with sequencing and variant calling; as well as Winnie Chan for assistance with patient recruitment.

FINANCIAL DISCLOSURES

ALS has a consulting agreement with Allena Pharmaceuticals.

REFERENCES

1. Keithi-Reddy S KR. Molecular and genetic basis of Alport Syndrome. In: Mount H PM, ed. *Molecular and Genetic Basis of Renal Disease*: W.B. Saunders; 2008:151-171.

2. Parkin JD, San Antonio JD, Pedchenko V, Hudson B, Jensen ST, Savige J. Mapping structural landmarks, ligand binding sites, and missense mutations to the collagen IV heterotrimers predicts major functional domains, novel interactions, and variation in phenotypes in inherited diseases affecting basement membranes. *Human mutation.* 2011;32:127-143.

3. Heidet L, Arrondel C, Forestier L, et al. Structure of the human type IV collagen gene COL4A3 and mutations in autosomal Alport syndrome. *Journal of the American Society of Nephrology : JASN.* 2001;12:97-106.

4. Longo I, Scala E, Mari F, et al. Autosomal recessive Alport syndrome: an in-depth clinical and molecular analysis of five families. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association.* 2006;21:665-671.

5. Wuttke M, Seidl M, Malinoc A, et al. A COL4A5 mutation with glomerular disease and signs of chronic thrombotic microangiopathy. *Clin Kidney J.* 2015;8:690-694.

6. Becknell B, Zender GA, Houston R, et al. Novel X-linked glomerulopathy is associated with a COL4A5 missense mutation in a non-collagenous interruption. *Kidney international.* 2011;79:120-127.

7. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS genetics.* 2006;2:e190.

8. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nature genetics.* 2002;30:97-101.

9. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics.* 2011;43:491-498.

10. Ott J. Computer-simulation methods in human linkage analysis. *Proceedings of the National Academy of Sciences of the United States of America.* 1989;86:4175-4178.

11. Biasini M, Bienert S, Waterhouse A, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic acids research.* 2014;42:W252-258.

12. Bella J, Liu J, Kramer R, Brodsky B, Berman HM. Conformational effects of Gly-X-Gly interruptions in the collagen triple helix. *J. Mol. Biol.* 2006;362:298-311.

13. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics.* 2011;27:343-350.

14. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry.* 2004;25:1605-1612.

15. Li Y, Brodsky B, Baum J. NMR shows hydrophobic interactions replace glycine packing in the triple helix at a natural break in the (Gly-X-Y)n repeat. *J. Biol. Chem.* 2007;282:22699-22706.

16. Zhou J, Hertz JM, Leinonen A, Tryggvason K. Complete amino acid sequence of the human alpha 5 (IV) collagen chain and identification of a single-base mutation in exon 23 converting glycine 521 in the collagenous domain to cysteine in an Alport syndrome patient. *J. Biol. Chem.* 1992;267:12475-12481.

17. Sun X, Chai Y, Wang Q, Liu H, Wang S, Xiao J. A Natural Interruption Displays Higher Global Stability and Local Conformational Flexibility than a Similar Gly Mutation Sequence in Collagen Mimic Peptides. *Biochemistry (Mosc).* 2015;54:6106-6113.

18. Mohs A, Popiel M, Li Y, Baum J, Brodsky B. Conformational features of a natural break in the type IV collagen Gly-X-Y repeat. *J. Biol. Chem.* 2006;281:17197-17202.

19. Brazel D, Oberbaumer I, Dieringer H, et al. Completion of the amino acid sequence of the alpha 1 chain of human basement membrane collagen (type IV) reveals 21 non-triplet interruptions located within the collagenous domain. *Eur. J. Biochem.* 1987;168:529-536.

20. Xiao J, Sun X, Madhan B, Brodsky B, Baum J. NMR studies demonstrate a unique AAB composition and chain register for a heterotrimeric type IV collagen model peptide containing a natural interruption site. *J. Biol. Chem.* 2015;290:24201-24209.

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

22. Crockett DK, Pont-Kingdon G, Gedge F, Sumner K, Seamons R, Lyon E. The Alport syndrome COL4A5 variant database. *Human mutation.* 2010;31:E1652-1657.

LEGENDS

Figure 1. Pedigree of X-linked recessive family with unusual glomerulopathy, pathology and genetic analysis. (a) Pedigree. Individuals where DNA was available are represented with a 4 digit number. The proband is indicated with an arrow. Affected individuals are shown in black. Whole exome sequencing was performed in individuals 6442 and 6444, revealing a shared rare *COL4A5* variant, c.T665G or p.F222C (NM_000495, rs281874761). Sanger sequencing of the *COL4A5* region was done in all individuals where DNA was available. Males hemizygous for the variant are denoted with an "M" (mutation) while females heterozygous for the variant are denoted with a "C" (carrier). Individuals where whole genome genotyping was performed for linkage analysis are indicated with a "+". (b) Renal biopsy showed thickened glomerular basement membranes, mesangial expansion by cells and matrix (PAS, 40 x), besides (not shown) global and segmental glomerulosclerosis; Immunofluorescence showed linear glomerular capillary wall staining for albumin (2+), IgA (1-2+) and IgG (1+) and segmental granular mesangial and capillary wall staining for IgM (1-2+), besides (not shown) normal staining for collagen IV alpha 1, 3 and 5 chains; Electron Microscopy revealed mostly homogeneously thickened glomerular basement membranes, rare loops with lamellations, and quite a few scattered intramembranous and partly resorbed subepithelial electron dense deposits (8000 x). (c) Linkage analysis and exome data. Analysis revealed a region on the X-chromosome from 90,000,000 to 118,700,000 (hg19 build 37) spanning 14.35 cM with a peak LOD score of 1.8 (left). Targeted exome enrichment and massively parallel sequencing was performed on genomic DNA from two affected individuals, the proband (6442) and his maternal uncle (6444), resulting in 56,702,602 and 42,709,094 reads, respectively. Following alignment, target region coverage had an average sequencing depth of 66X and 53X for the two samples. Exome sequencing

revealed a *COL4A5* variant, c.T665G, p.F222C, shared between the affected related male individuals as shown in Integrative Genomic Viewer, which was not reported in dbSNP138, 1000 Genomes Project release version 01-11-2004, Exome Sequencing Project 21-06-2014 or ExAC 01-11-2014.

Figure 2. Homology model of the α3α4α5 heterotrimer in correspondence of the mutation site. A homology model for the collagen 4 triple helix (α3, α4 and α5 chains) in the region surrounding the amino acidic substitution was obtained with the DeepView software and the SwissModel server, using the crystallographic structure of a trimer with a G1G gap as a template (PDB code 1Ei8), obtaining a favorable model quality estimate QMEAN Z-score of -0.94.[11-13] Several possibilities exist for the register between the three chains, which in the collagen trimer are staggered with respect to each other. The order α3α5α4 (in the trailing, middle, and leading chain notation) was selected for the model, based on optimization of interactions between the chains, minimization of the number of steps without Gly residues, and by analogy to the α1α1α2 trimer.[15] Molecular graphics and analyses were realized with the UCSF Chimera software package.[14] The protein backbone is reported in ribbon representation, with Gly residues colored in yellow, while all others are shown in red (α3), green (α4) or blue (α5). Interruptions in the Gly-Xaa-Yaa repetition are colored in lighter colors (pink, cyan, and light green for α3, α4 and α5, respectively. Phe222 (purple) and the hydrophobic residues interacting with it (Val222 and Ile223 in α3, Val236 and Val238 in α4 and Leu220 in α5, shown in grey) are reported in a semi-transparent space filling representation. Arg221 in in α3, corresponding to a HSP47 binding site, is shown in orange, in stick representation.