Next-Generation Sequencing for New Gene Identification and for Diagnosis in Steroid-Resistant Nephrotic Syndrome

Thesis

NEXT-GENERATION SEQUENCING FOR NEW GENE IDENTIFICATION AND FOR DIAGNOSIS IN STEROID-RESISTANT NEPHROTIC SYNDROME

Thesis submitted by

Paraskevas Iatropoulos, MD

for the degree of

Doctor of Philosophy

Discipline of Life and Biomolecular Sciences

The Open University, United Kingdom

IRCCS Istituto di Ricerche Farmacologiche Mario Negri, Italy

Director of Studies

Dr. Marina Noris, PhD

Supervisors

Prof. Robert Kleta, MD, PhD

Dr. Erica Daina, MD

30th September 2016

ABSTRACT

Nephrotic syndrome is clinically characterized by massive proteinuria, and hypoalbuminemia. It represents a heterogeneous group of glomerular disorders characterized by distinct causes and histopathologic lesions. The present thesis evaluates the contribution of genetics in conditions commonly associated with NS: the immune complex-mediated membranoproliferative glomerulonephritis (IC-MPGN), the C3 glomerulopathy (C3G) and the podocytopathies, usually characterized by steroid-resistant nephrotic syndrome (SRNS).

The first part of the thesis focused on IC-MPGN and C3G. We found likely pathogenic (LP) variants in complement alternative pathway genes in 18% of patients and describe, for the first time, thrombomodulin rare pathogenic variants in C3G patients. Interestingly, mutations alone did not increase the risk of developing IC-MPGN or C3G, but they did so when combined with common susceptibility variants.

The prevalence of alternative pathway abnormalities (mutations and/or C3 Nephritic Factors, C3NeFs) was similar in IC-MPGN (53-56%) and C3G (64-65%). To investigate for more homogeneous subgroups within IC-MPGN/C3G, an operator-independent approach was applied using genetic, biochemical, histological and clinical data. Four different groups emerged showing distinct pathogenesis, and histological and clinical features.

In the second part, the genetic causes of podocytopathies were investigated. LP variants in podocytopathy-associated genes are found in 23% of SRNS patients, prevalently in

COL4A3-5 genes. Moreover, 8% of patients carried LP variants in Congenital Anomalies of the Kidney and the Urinary Tract (CAKUT)-associated genes, which were not previously associated with podocytopathies. LP variants in CAKUT-associated and podocytopathy-associated genes frequently combined together.

Finally, possibly pathogenic variants in *EPB41L45*, a candidate gene for podocytopathies, were identified in 3 unrelated patients.

In conclusion, this thesis contributes to understand the complex genetic basis of both IC-MPGN/C3G and podocytopathies. It introduces new players in the pathogenesis of these diseases. Finally, it provides evidence of the presence of subgroups within IC-MPGN/C3G with distinct underlying mechanisms, and clinical and histopathologic features.

To Elisabetta and Lydia

ACKNOWLEDGEMENTS

I would like to express my deepest thanks to Dr. Marina Noris, my Director of Studies, for her continuous suggestions that enrich my research, and for her wise and efficient supervision; to Prof. Robert Kleta, my Supervisor, for our inspiring meetings and for the time he kindly dedicated to supervise my research; to Dr. Erica Daina, also my Supervisor, for her constant guidance and support, for teaching me each day the real aim of a researcher and for her patience; and to Dr. Elena Bresin, my Third Party monitor, for our long pleasant and fruitful collaboration.

I am grateful to Prof. Giuseppe Remuzzi, Research Coordinator of the IRCCS Istituto di Ricerche Farmacologiche Mario Negri, for giving me the opportunity to attend the PhD course and for keenly identifying the strengths and weaknesses of my research work all these years.

A heartfelt thank to my wife and collaborator, Dr. Elisabetta Valoti, for contributing to the experiments presented in this thesis. Without her constant help, support and encouragement, this thesis could not have been finished.

I would also like to thank Dr. Manuela Curreri for her precious contribution to the first part of the thesis, for the harmonious collaboration and friendship.

I am grateful to Dr. Samantha Solini. By being one of the most efficient and hard-working person I ever met, she provided an invaluable contribution to the second part of the thesis.

My sincere thanks go to Dr. Serena Bettoni, Dr. Rossella Piras, and Marta Alberti for

their precious contribution to this study; to Dr. Matteo Breno for the bioinformatic support; and to Giovanni Nattino and Annalisa Perna for the useful discussion on statistical analysis results.

I also want to thank Mrs. Sara Gamba, Eng. Laura Bottanelli, Mrs. Giusepinna Omati and Ms. Monica Lena for making my work easier with their daily efforts.

I would like to thank all the patients and their physicians; without their volunteer participation this research work would not be feasible.

If I faced the hard period of the thesis writing with a feel of happiness, it was thanks to the little Lydia.

A special thank to my parents, Chrysoula and Antonio, whose love and sacrifices have allowed me to pursue this goal.

Contribution to the Thesis by other Researchers

Part of the thesis project was carried on with the collaboration of other researchers that contributed to the research as follows:

Internal collaborations in the IRCCS Istituto di Ricerche Farmacologiche Mario Negri

- Clinical data and biological samples were collected through the MPGN and SRNS registries by Dr. Manuela Curreri, Dr. Bresin, Dr. Daina, Ms. Sara Gamba and the candidate.
- Regarding to the six gene NGS panel for IC-MPGN/C3G cohort, Dr. Caterina Mele participated in the "wet-lab" and the bioinformatic set-up. She also participated in the screening of the 6 complement genes together with Dr. Elisabetta Valoti, Dr. Rossella Piras and Marta Alberti.
- SC5b-9 assays were performed by Dr. Serena Bettoni.
- C3 and C4 assays were performed by Dr. Flavio Gaspari and the Antonio Nicola Cannata.
- Regarding to the genetic panel for proteinuric nephropathies and the screening of the podocytopathy/SRNS cohort, Dr. Samantha Solini participated in the set-up of the assay and in the screening of the patients.
- Analysis of the FSGS MASTR panel data was performed by Dr. Matteo Breno

External collaborations

- Marina Vivarelli, Francesco Emma Division (both at the Nephrology and Dialysis, Bambin Gesù Pediatric Hospital, Roma, Italy) and Luisa Murer (at the Unit of Pediatric Nephrology, Dialysis and Transplantation, Azienda Ospedaliera of Padova, Italy) participated in the recruitment of IC-MPGN/C3G patients.

- C3NeF assays were performed by Dr. Veronique Fremaux-Bacchi at the Department of Immunology, Assistance Publique-Hopitaux de Paris, Hopital Europeen George-Pompidou and INSERM UMRS 1138, Cordelier Research Center, Complement and Diseases Team, Paris, France.

- Genotyping using the Genome-Wide Human SNP Array 6.0 (Affymetrix) was performed by Dr. Chiara Magri at the Department of Biomedical Sciences, University of Brescia.

- Whole exome sequencing in SN140 was performed by the group of Prof. Massimo Delledonne at the Department of Biotechnologies, University of Verona providing a list of not annotated variants.

- Whole exome sequencing in IR35 was performed by Genomnia srl., Bresso, Milano, Italy, providing a list of not annotated variants.

- Sequencing with the FSGS MASTR panel was performed by Multiplicom France, Lyon, France, providing a list of not annotated variants. Analysis of the data was performed by Dr. Matteo Breno.

All other experiments and analyzes described in this thesis were performed by the PhD candidate, Dr Paraskevas Iatropoulos.

Research Funding

This work was supported by:

- Fondazione ART per la Ricerca sui Trapianti ART ONLUS (Milano, Italy)
- Fondazione Aiuti per la Ricerca sulle Malattie Rare ARMR ONLUS (Bergamo, Italy)
- Progetto DDD Onlus-Associazione per la lotta alla DDD, Milan, Italy.
- Telethon grant GGP09075
- Ministero della Salute - Ricerca finalizzata – Giovani ricercatori GR-2011-02347678
- European Union Seventh Framework Programme FP7-EURenOmics project number 305608
- Minigrant 2010 of Genomnia (Lainate, Milan, Italy) in collaboration with Life Technologies

TABLE OF CONTENTS

INDEX OF FIGURES

INDEX OF TABLES

ABBREVIATIONS

ACEi	Angiotensin-Converting-Enzyme inhibitors
ACTN4	α-actinin-4
ADCK4	AarF Domain Containing Kinase-4
AF	Allele frequency based on Flow Evaluator observation counts
aHUS	Atypical Hemolytic Uremic Syndrome
AII-RA	Angiotensin II Receptor Antagonists
ANP	Atrial Natriuretic Peptide
AO	Alternate allele observations
AP	Alternative Pathway (of complement system)
aPKC	Atypical Protein Kinase C
AR	Aldose reductase
ARB	Angiotensin Receptor Blockers
ARHGAP24	RhoA-activated Rac1 GTPase-activating protein
ARHGDIA	Rho GDP dissociation inhibitor-α
BAF	B-allele frequency
BRCA1	Breast Cancer 1 gene
BRCA2	Breast Cancer 2 gene
BSA	Bovine serum albumin
C1	Complement complex 1
C1-INH	C1 Inhibitor
C1q	Complement complex 1Q
C1r	Complement component 1R
C1s	Complement component 1S
C2	Complement component 2
C3	Complement component 3
C3a	Smaller of the two cleavage products of C3
C3aR	Complement C3a Receptor
C3b	Larger of the two cleavage products of C3
C3G	C3 glomerulopathy
C3GN	C3 glomerulonephritis
C3NeF	C3 Nephritic Factor
C4	Complement component 4
C4BP	C4 Binding Protein
C5	Complement component 5
C5a	Smaller of the two cleavage products of C5
C5aR	Complement C5a Receptor, also known as CD88
C5b	Larger of the two cleavage products of C5

C5b-9	Complex of proteins composed of C5b, C6, C7, C8 and multiple molecules of C9
C6	Complement component 6
C7	Complement component 7
C8	Complement component 8
C9	Complement component 9
CAKUT	Congenital Anomalies of the Kidney and the Urinary Tract
CASK	Calcium/Calmodulin dependent Serine protein Kinase
CATL	Cytoplasmic cathepsin L
CD2AP	CD2-Associated Protein
CD46	Cluster of Differentiation 46, also called MCP
CDC42	Cell Division Cycle 42
CFB	Complement Factor B gene
CFH	Complement Factor H gene
CFHR	Complement Factor H-related gene
CFI	Complement Factor I gene
CI	Confidence Interval
CNF	Congenital nephrotic syndrome of Finnish type
CNS	Congenital Nephrotic Syndrome
COL4A3	Collagen Type IV α3
COL4A4	Collagen Type IV α4
COL4A5	Collagen Type IV α5
CoQ10	Coenzyme Q10
CoQ2	Coenzyme Q2 Polyprenyltransferase
CoQ6	Coenzyme Q6 FAD-dependent Monooxygenase
CP	Classical Pathway (of complement system)
CR1	Complement Receptor 1, also known as CD35
CR3	Complement receptor 3
CR4	Complement receptor 4
CsA	Cyclosporin A
Cα1-3	Constant regions 1 to 3 of the IgA heavy chains
DAF	Decay Accelerating Factor
DAG	Diacylglycerol
DDD	Dense-deposit disease
DID domain	Diaphanous inhibitory domain (of INF2)
DMS	Diffuse Mesangial Sclerosis
DNA	Deoxyribonucleic acid
EM	Electron microscopy
EMP2	Epithelial Membrane Protein 2
EPB41L5	Erythrocyte Membrane Protein Band 4.1 Like 5 gene
ESRD	End Stage Renal Disease
ExAC	Exome Aggregation Consortium
FAT1	Fatty Acid Transporter tumor suppressor homolog-1
FAT2	Fatty Acid Transporter tumor suppressor homolog-2

FB	Complement factor B
FDP	Flow-space corrected read depth
FH	Factor H (of complement system)
FHR	Factor H-related protein
FI	Factor I (of complement system)
FRNS	Frequently relapsing nephrotic syndrome
FSGS	Focal Segmental Glomerulosclerosis
FXX	Flow Evaluator failed read ratio
Gal	Galactose
GalNAc	N-acetylgalactosamine
GBM	Glomerular Basement Membrane
Gd-IgA1	galactose-deficient O-linked glycan IgA1
GDP	Guanine Diphosphate
GFB	Glomerular Filtration Barrier
GLEPP1	Glomerular epithelial protein 1
GQ	Genotype quality score
GWAS	Genome-wide association study
HDL	High Density Lipoprotein
HRUN	Run length: the number of consecutive repeats of the alternate allele in the reference genome
HSPG	Heparan sulfate proteoglycan
ICAM-1	Intercellular Adhesion Molecule 1
IC-MPGN	Immune-complex-associated MPGN
IF	Immunofluorescence
IFT139	Retrograde intraflagellar transport protein 139
IgA	Immunoglobulin A
IgA1	Immunoglobulin A type 1
IgAN	Immunoglobulin A Nephropathy
IgG	Immunoglobulin G
IgM	Immunoglobulin M
Ig-MPGN	Immunoglobulin-associated MPGN, also called IC-MPGN
IGV	Integrative Genomics Viewer
INF2	Inverted formin 2
IP3	Inositol 1,4,5-trisphosphate
IQR	Interquartile Range
JAMA	Junctional Adhesion Molecule A
LAMB2	Laminin β2 subunit
LDL	Low Density Lipoprotein
LEN	allele length
LG	Laminin Globular domain
LM	Light microscopy
LN	Laminin N-terminal domain
LP	Likely Pathogenic
MAC	Membrane Attack Complex, also called Terminal Complement Complex (TCC)

MAF	Minor Allelic Frequency
MAGI-1 / MAGI1	Membrane Associated Guanylate Kinase, WW And PDZ Domain Containing 1
MAGI-2 / MAGI2	Membrane Associated Guanylate Kinase, WW And PDZ Domain Containing 2
MASP	Mannan-binding lectin serine protease
MASP2	Mannan-binding lectin serine protease 2
MC	Minimal Change disease
MCP	Membrane Cofactor Protein, also called CD46
MLLD	Mean log-likelihood delta per read.
MN	Membranous Nephropathy
mnp	Multi-Nucleotide Polymorphism
MPGN	Membranoproliferative Glomerulonephritis
MYO1E	Myosin 1E
NC1	Non Collagen 1 domain
NCK	NCK (Non-Catalytic Region Of Tyrosine Kinase) Adaptor Proteins
NCK1	NCK (Non-Catalytic Region Of Tyrosine Kinase) Adaptor Protein 1
NCK2	NCK (Non-Catalytic Region Of Tyrosine Kinase) Adaptor Protein 2
NEP	Neutral endopeptidase
NEPH1	Nephrin-Like Protein 1, also known as Kin Of IRRE-Like Protein 1 (KIRREL1)
NEPH2	Nephrin-Like Protein 2
NEPH3	Nephrin-Like Protein 3
NGS	Next Generation Sequencing
NOS	Not otherwise specified
NPHS1	nephrin
NPHS2	podocin
NS	Nephrotic Syndrome
OR	Odds-ratio
Par3	Partitioning defective 3
Par6	Partitioning defective 6
PCA	Principal component analysis
PDGF	Platelet-derived Growth Factor
PDSS2	Prenyl Decaprenyl Diphosphate Synthase Subunit 2
PEC	Parietal Epithelial Cell
PGM	Personal Genome Machine
PL	Genotype likelihood scores
PLA2R1	M-type phospholipase A2 receptor 1
PLCE1	Phospholipase C epsilon 1
PLRD	Variant likelihood ratio over coverage depth
QD	Quality By Depth as 4*QUAL/FDP (analogous to GATK)
QUAL	Variant quality score

RAAS	Renin-Angiotensin-Aldosterone System
RAC1	Ras-Related C3 Botulinum Toxin Substrate 1
RBI	Distance of bias parameters from zero.
REVB	Reverse strand bias in prediction.
RHOA	Ras Homolog Family Member A
RR	Relative risk
s-C3	serum C3
SC5b-9	S-protein associated C5b-9 (terminal pathway) complex
SCR	Short Consensus Repeat
SDNS	Steroid-dependent nephrotic syndrome
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
SOD2	Superoxide dismutase 2
SRNS	Steroid Resistant Nephrotic Syndrome
SSNS	Steroid sensitive nephrotic syndrome
SSSB	Strand-specific strand bias for allele.
STB	Strand bias in variant relative to reference.
TBMN	Thin basement membrane nephropathy
TCC	Terminal Complement Complex, also called Membrane Attack Complex (MAC)
TGF-β	Transforming Growth Factor β
THBD	Thrombomodulin
THSD7A	Thrombospondin type-1 domain-containing 7A
TRPC	Transient Receptor Potential cation Channel
TRPC6	Calcium channel Transient Receptor Potential cation Channel 6
TSVC	Torrent Suite Variant Caller
VARB	Variant Hypothesis bias in prediction.
VARW	Variation of the width of deletions and insertions
VCAM-1	Vascular Cell Adhesion Molecule 1
VE-cadherin	Vascular Endothelial cadherin
VEGF	Vascular Endothelial Growth Factor
V_H	Variable region of the IgA heavy chains
VQSR	Variant quality score recalibration
WAGR syndrome	Wilms tumor, Aniridia, Genitourinary abnormalities, and mental Retardation syndrome
ZO1	Zona Occludens associated protein 1
αENO	α-enolase

1. INTRODUCTION

1.1 The glomerular anatomy, physiology and molecular biology

1.1.1 The glomerulus

The kidney's key function is to filter waste products of metabolism out of the blood and pass them out of the body as urine. The functional units of the kidney are nephrons, which are further divided into filtration units, called renal corpuscles, and reabsorption units, called renal tubules. In the renal corpuscle, the afferent arteriole splits into a network of capillaries (glomerulus) that form tortuous loops and then drain to the efferent arteriole close to the afferent (vascular pole). The glomerular tuft of capillaries is enclosed in a globular sac, called Bowman's capsule, and the space that separates the capillaries and Bowman's capsule is called urinary space or Bowman's space.[1]

The blood enters in the glomerulus through the afferent arteriole, it flows into the glomerular capillaries where part of the plasma water and small solutes are forced out into the Bowman's space, then it leaves the glomerulus through the efferent arteriole. This process is called ultrafiltration and is promoted by the efferent arteriole resistance that results in high hydrostatic pressure inside the glomerular capillary. The ultrafiltrate flows into the proximal tubule at the opposite of the vascular pole (urinary pole). The majority of the ultrafiltrate is reabsorbed back to the blood stream as it flows through the tubules and the collecting ducts. The remaining is eliminated as urine.[1, 2]

The capillary wall where the filtration takes place is called the glomerular filtration barrier (GFB). It is composed of three layers: the endothelial cells that constitute the

intricately tortuous inner capillary tuft; the glomerular basement membrane (GBM), an extracellular cell matrix; and the podocytes, visceral epithelial cells that tightly wrap around the exterior of glomerular capillaries.[2]

In the glomerulus are also present the mesangial cells that support the capillary loops, regulate blood flow and glomerular filtration rate, and have endocrine functions. Another cell population are the parietal epithelial cells at the Bowman's capsule side.[2]

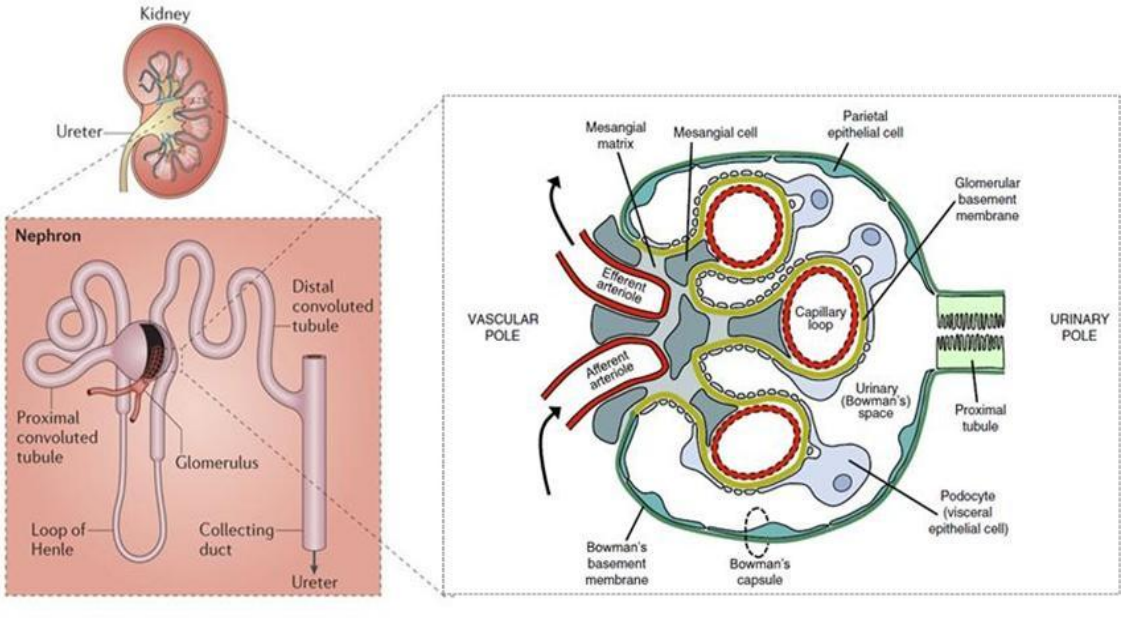


Figure 1.1. Graphical representation of the kidney, the nephron and the renal glomerulus (modified from Kurts, Nat Rev Immunol 2013, and Leeuwis, Adv Drug Deliv Rev 2010).[3, 4]

1.1.2 The glomerular endothelial cells

The endothelium constitutes the innermost layer of the GFB. Glomerular endothelial cell have two peculiarities: they are unusually flatten (height 50-150 nm) and they present the fenestrae, which are 60-100 nm wide transcellular pores comprising approximately 20-50% of the endothelium surface. These pores allow a more efficient and rapid passage of high volumes of fluid.[1, 2, 5]

Endothelial cells and the fenestrae are covered by the glycocalyx, a network of negatively charged glycoproteins, glycosaminoglycans and membrane-bound proteoglycans that cover the luminal face of glomerular capillaries constituting charge selective barrier.[6] Plasma proteins are absorbed within the glycocalyx through binding to glycosaminoglycans. This layer is the first size and charge selective barrier with a thickness of 500 nm that restricts the passage of large molecules (more than 200 kDa) and/or negatively charged proteins (such as albumin). Evidence on the glycocalyx function comes from many proteinuric animal models in which the reduction of glomerular endothelial cell glycocalyx is associated with proteinuria.[2]

In addition to the permeability barrier function, the endothelium contributes to the regulation of vasomotor tone inducing vascular smooth muscle contraction through the secretion of endothelin and platelet derived growth factor, and relaxation through the secretion of nitric oxide and prostacyclin.[7] Another important function is to maintain the fluidity of the blood and to control coagulation through multiple mechanisms including surface expression of thrombomodulin and anticoagulant heparan sulfate, secretion of protein S and prostacyclin, and the ability to catalyze the activation of plasminogen to plasmin.[7]

Moreover, endothelial cells are active participants in immune response both through vasodilator effects and by leukocyte recruitment. Cytokines or other inflammatory mediators activate the endothelial cells inducing the expression of adhesion molecules, such as E-selectin, Intercellular Adhesion Molecule 1 (ICAM-1), Vascular Cell Adhesion Molecule 1 (VCAM-1) and P-selectin, that promote contacts between the endothelium and circulating leukocytes, thus leading to leukocyte recruitment at sites of inflammation and migration into the tissues. These processes explain the presence of leukocytes within the glomerulus in diseases with an inflammatory or immune basis.[7, 8]

The endothelium has a major role in the regulation of the complement system during immune response by expressing on its surface a set of regulators including membrane cofactor protein (MCP), thrombomodulin, decay accelerating factor (DAF) and CD59.[9] Since complement system dysregulation causes membranoproliferative glomerulonephritis and C3 glomerulopathy, which are widely studied in this thesis, and is also involved in the pathogenesis of other glomerulopathies that present nephrotic syndrome, a more detailed description of the complement system is provided in section 2 of this chapter.

1.1.3 The glomerular basement membrane

The GBM is a 250 to 400 nm thin grid of extracellular matrix proteins located between the glomerular endothelial cells and the podocytes. The GBM supplies structural support for the glomerular capillaries and presents ligands for receptors of the adjacent endothelial cells, podocytes, and mesangial cells.[10, 11] In addition, it acts as a size and charge selective filtration barrier.[12]

The GBM is composed by three layers: lamina rara externa (adjacent to podocyte processes and mostly composed of heparan sulfate proteoglycans), lamina densa (constituting the large central zone and prevalently composed of type IV collagens, laminins and nidogens) and lamina rara interna (adjacent to endothelial cells, containing heparan sulfate proteoglycan). These molecules are synthesized and secreted by the podocytes and the glomerular endothelial cells.[13]

Heparan sulfate proteoglycans (HSPGs) are composed of a core protein to which negatively charged glycosaminoglycan side chains are bound. The HSPG found in GBM is prevalently agrin, but also perlecan, and collagen XVIII are present.[14, 15]

HSPGs contribute to the charge-selectivity of the filtration barrier. Indeed, the elimination of their anionic charges after intravenous heparanases administration in animal models results in increased ferritin permeability and collagen XVIII deficient mice show mesangial expansion and mild renal insufficiency suggesting that this HSPG may play a role in the filter barrier structure and function.[15, 16] In addition, glycosaminoglycans sequester and regulate availability of growth factors, such as vascular endothelial growth factor (VEGF), during their passage from podocytes to the fenestrated endothelium.[17]

Laminins are large heterotrimeric glycoproteins composed by three chains: α, β, and γ. The different chains are stabilized together through interchain disulfide bonds. Typically, laminin heterotrimer has a cross-like shape structure in which the laminin coiled-coil domains of each chain wrap around to form the longer arm of the cross. At the end of this arm there is the C-terminal laminin globular (LG) domain of the α chain through which laminin interacts with cell surface receptors such as integrins and dystroglycan. The three short arms are formed by alternating globular and rod-like domains and contain the globular laminin N-terminal (LN) domain. The LN domain mediates the laminin trimer-trimer interactions that lead to the formation of the laminin polymer in the GBM.[10] There are five different α chains, four β and three γ. The major laminin found in the mature GBM is the laminin α5β2γ1 also called laminin-521; however, during the GBM development there are transitions of the laminin trimers from laminin-111 to laminin-511 to laminin-521.[10] The importance of laminins is underlined by the massive proteinuria observed in laminin β2 chain deficient mice and in patients with Pierson syndrome that carry mutations in the laminin β2 gene. Interestingly, proteinuria preceded alterations in podocyte foot process morphology and slit-diaphragms in laminin β2 knockout mice, supporting the active role of the GBM as a size-selective filter.[10, 18]

Type IV collagens represent 50% of the all mature GBM proteins. Type IV collagen is a trimeric protein consisting of α chains rich in Gly-X-Y amino acid triplet repeats that wrap around one another to form the collagen triple helix. The Glycine at every third residue is the only amino acid small enough to fit at the center of the helix. Unlike most other collagen types, type IV collagens have interruption of the Gly-X-Y repeats that confer more flexibility to the molecules and to the GBM. Each type IV collagen has also a globular domain at the C-terminal (non collagen 1, NC1) and a 7S domain at the N-terminal.[10, 19] There are 6 collagen IV genes that encode α1-6 chains assembled to form three different trimers (α1α1α2, α3α4α5 and α5α5α6) called protomers. Maturation of the GBM involves the transition of the α1α1α2 collagen to the α3α4α5. Protomers are secreted in the extracellular matrix and then self-polymerize through NC1-NC1 and 7S-7S interactions to form a network distinct from that of the laminins.[10] Of note, mutations in type IV collagen encoding genes cause the Alport syndrome, a disease characterized by hematuria and progressive renal disease.[20] These patients usually present only mild proteinuria suggesting that type IV collagens probably have a minor role in the filtration selectivity of the glomerular barrier.[20, 21]

Nidogen-1 and nidogen-2 are two homologous glycoproteins that make also part of the GBM. They contain three globular-like domains with two rod-like domains that separate the globular-like domains providing to nidogen a dumbbell-shape. Nidogen-1 binds both laminin and type IV collagen and is supposed to link the corresponding distinct network and to provide extra stability to GBM under situations of unusual stress. However, double deficient mice for nidogen-1 and nidogen-2 show normal GBM and almost no renal abnormality.[10]

Recently, a new model to explain the charge-selective nature of the GFB was proposed. Micropuncture studies on animal models showed that there is an electrical field across

the GFB, which is generated by filtration and is directly proportional to filtration hydrostatic pressures. During the passage of plasma through the negatively charged gel-like GBM, due to a phenomenon called streaming potential, small anions such as Cl- and HCO3- pass the filter slightly faster than small cations such as Na+ and K+. This results in a potential difference that is negative in Bowman's space. It is estimated that it is created a field with strength of 1600 V/m. The electrical field across the GFB would drive albumin and the other plasma proteins, which are negatively charged, toward the capillary lumen by electrophoresis.[22]

1.1.4 The podocytes

The podocytes are highly specialized epithelial cells that cover the urinary face of the glomerular capillaries. Podocytes have multiple functions including size and charge barrier to plasma proteins, counteraction of the intraglomerular pressure and maintenance of the capillary loop shape, synthesis and maintenance of the GBM, and production and signal transduction.[23, 24]

The podocyte can be divided in cell body, major processes, secondary processes and foot processes. The foot processes interdigitate with neighboring foot processes, and adjacent foot processes are directly linked by specialized cell-cell junctions called slit-diaphragms. In proteinuric diseases, the retractions of foot processes are often reported, a phenomenon named effacement, but how this effacement is involved in the development of proteinuria is not fully understood.[25]

The slit-diaphragm divides the plasma membrane of foot processes into three different compartments: the basal, the apical, and the lateral (slit-diaphragm) surfaces. All these compartments are interconnected with the actin cytoskeleton of the foot processes; they

participate in signaling pathways and regulate podocyte shape and motility.[23, 26, 27]

The **basal** compartment anchors the podocyte to the GBM though several types of integrins, tetraspanins and dystroglycans.[28, 29] Integrins are heterodimeric α/β transmembrane proteins. The most common podocyte integrin is α3β1, which is an important receptor for the GBM component laminin-521. On the cytoplasmic surface of the podocyte membrane, α3β1 integrin associates with paxillin, talin, and vinculin, which mediate its connection to the actin cytoskeleton, and with integrin-linked kinase (ILK).[30, 31] ILK forms a complex with nephrin (NPHS1), a slit-diaphragm protein, and actinin-4 (ACTN4), a protein of the cytoskeleton. Mouse and humans α3-mutants present glomerular abnormalities and early-onset massive proteinuria.[28, 32, 33] Podocyte-specific deletion of ILK in mice results in glomerular abnormalities and proteinuria.[31, 34] Moreover, absence of ILK results in the redistribution of NPHS1 and ACTN4.[31]

Another integrin of the foot process basal surface, αvβ3, interacts with the urokinase plasminogen activator receptor and with vitronectin, a component of the extracellular matrix. Genetically modified mice with constitutively active beta3 integrin show proteinuria. Conversely, mice lacking b3 integrin or the avb3 integrin ligand vitronectin are protected from LPS-induced proteinuria.[35]

Among tetraspanins, CD151 is located to the base of the foot processes and interacts with cell-matrix adhesion complexes such as α3β1 integrin. CD151-deficient mice develop proteinuria with GBM alterations that seem to precede podocyte abnormalities suggesting that CD151 is involved in the maturation and/or maintenance of the GBM structure.[32]

Finally, αβ dystroglycan, a component of the basal membrane, binds laminin and agrin

on extracellular side, and is connected to the actin cytoskeleton on cytoplasm side via utrophin.[36]

The **apical** side is negatively charged due to the presence of the anionic proteins podocalyxin, podoplanin, and podoendin, which form the podocyte glycocalyx and thereby contribute as a charge filter of the GFB. Podocalyxin is a type 1 transmembrane sialoprotein with an N-terminal mucin-like domain. It has anti-adhesive function that maintains an open filtration pathway between neighboring foot processes in the podocyte by charge repulsion.[37] On the cytoplasmic side, podocalyxin is connected to the actin cytoskeleton via ezrin and Na+/H+ exchanger regulatory factor 2 (NHERF2). Podocalyxin-deficient mice lack foot processes and slit-diaphragms and are unable to filter primary urine.[38] Disruption in animal models of the podocalyxin/ezrin/NHERF complex leads to foot process effacement and proteinuria.[39]

Glomerular epithelial protein 1 (GLEPP1) is a transmembrane tyrosine phosphatase receptor type O exclusively present on the apical podocyte surface.[40] GLEPP1 seems to be involved in the regulation of podocyte structure and function by phosphorylation of podocyte proteins.[40] Glepp1 mutation cause proteinuria both in mice and in humans.[41]

The podocyte **slit-diaphragm** is a specialized cell-cell junction of the GFB with zipper-like structure and constant width of 40 nm.[42] The main function of the slit-diaphragm is to form a size and probably charge selective barrier for plasma proteins. The slit-diaphragm was first considered as a modified adherens junction, due to common morphologic features and proteins (P-cadherin, Fatty Acid Transporter tumor suppressor homolog-1 and -2, FAT1 and FAT2, and α-, β- and γ- catenins), together

with podocyte-specific proteins.[43] Later the presence of tight junction proteins, such as junction adhesion molecule A, occludin and cingulin, were observed in the slit diaphragm complexes.[44] Altogether these make the slit-diaphragm a specialized multi-protein junction for carrying out the renal ultrafiltration function. These specific proteins interact with the actin cytoskeleton, triggering signalling pathways and regulating podocyte motility.[26, 45-47]

Many proteins constitute the slit-diaphragm such as NPHS1, podocin (NPHS2), Nephrin-Like proteins 1, 2 and 3 (NEPH1, NEPH2 and NEPH3), FAT1, FAT2 and Calcium channel Transient Receptor Potential cation Channel 6 (TRPC6). CD2-Associated Protein (CD2AP) and Non-Catalytic Region of Tyrosine Kinase) Adaptor Proteins (NCKs) connect the slit-diaphragm to the actin cytoskeleton. Furthermore, other proteins that are not included in the slit-diaphragm, are also involved in cytoskeleton regulation including synaptopodin, dynamin and cytoplasmic cathepsin L (CATL).[27]

NPHS1 is a transmembrane adhesion protein that belong to the immunoglobulins superfamily and the first transmembrane protein identified in the slit-diaphragm.[48] Its extracellular domain forms homodimers and heterodimers with NEPH1 contributing to the formation of the zipper-like structure.[46] NPHS1 is recruited at the slit diaphragm by NPHS2. This is a U-shaped membrane protein of the slit-diaphragm that accumulates in oligomeric form in lipid rafts. NPHS2 is also associated with CD2AP and NEPH1 suggesting that it may have a crucial role in the assembly of the slit-diaphragm complex, as a scaffolding protein.[49-51] The transmembrane proteins NEPH1, NEPH2 and NEPH3 are also located to the slit-diaphragm and share certain homology with NPHS1.[52] NEPH1 seems to be involved with NPHS1 in actin recruitment via NCK proteins.[53] The importance of NEPH1 is underlined by the fact that Neph1 deficient

mice die perinatally due to massive proteinuria.[54]

TRPC6 is a transient non-selective cation channel that belongs to the Transient Receptor Potential Cation Channel (TRPC) family and interacts with the slit-diaphragm. It is involved in mechanosensation, regulates intracellular calcium concentration in response to the G-protein-coupled receptor activation and to tyrosine kinase receptors, and regulation of gene transcription.[55-57] TRPC6 interacts both with NPHS1 and NPHS2. TRPC6 is phosphorylated by Fyn, a member of the Src protein kinase family, thus enhancing TRPC6 channel conductivity, and it is dephosphorylated by calcineurin.[58, 59] Similar to TRPC6, NPHS1 is also phosphorylated by Fyn. Of note, Fyn-deficient mice develop proteinuria and foot process effacement.[60]

P-cadherin, vascular endothelial cadherin (VE-cadherin) and the two protocadherins FAT1 and FAT2, are also located to the slit-diaphragm.[42] FAT1 is involved in the formation of new intercellular junctions both during podocyte development and as a response to injury.[61] Fat1 deficient mice lack slit-diaphragms and show proteinuria and mutations in humans cause mixed glomerulo-tubular proteinuria.[62, 63] P-cadherin is associated with signaling proteins α-, β-, γ- catenins and colocalizes with the scaffold protein, zona occludens associated protein (ZO-1, a member of the membrane associated guanylate kinase family), but does not seem to be necessary for the renal filtration barrier, while the role of VE-cadherin in the podocyte is still unknown.[52]

NPHS1 is linked to the actin cytoskeleton through CD2AP and NCK intracellular proteins.[64-66] The NPHS1/CD2AP connection is critical for the functional renal filtration demonstrated by the observation that Cd2ap deficient mice die due to massive proteinuria and show foot process effacement.[66] Phosphorylation of NPHS1 also promotes NCK-dependent actin rearrangements, suggesting a critical role for NPHS1 in preventing foot process effacement.[64, 67, 68] NCK1 and NCK2 interact with tyrosine

phosphorylated NPHS1. Mice lacking both Nck proteins in podocytes develop massive proteinuria.[64, 65]

The cell polarity complex is localized at the cytoplasmic side of the slit-diaphragm and include partitioning defective 3 (Par3), partitioning defective 6 (Par6) and atypical protein kinase C (aPKC).[69] Par3 can bind directly to NPHS1 and NEPH1, and recruit Par6/aPKC to the slit-diaphragm area.[70] The deficiency of an aPKC isoform in podocytes results in altered distribution of the slit-diaphragm molecules and proteinuria suggesting that this cell polarity complex is essential for the GFB.[69]

Dendrin, the junctional adhesion molecule A (JAMA), densin, MAGI-1/2, Calcium/Calmodulin dependent Serine protein Kinase (CASK), αII spectrin, and βII spectrin have also been reported to be associated with the slit-diaphragm despite the role of these molecules in podocytes still needs to be clarified.[42]

1.1.5 The mesangial cells

The mesangial cells are specialized pericytes important for the keeping of the glomerular tuft integrity and forming the stalk that supports the capillary loops.[71] Mesangial cells are in direct contact with glomerular endothelial cells and are separated from podocytes by the GBM. They show contractile properties and they can regulate the distensibility of the glomerulus and the capillary blood flow in response to systemic arterial pressure.[72, 73] In this way, mesangial cells regulate the glomerular filtration rate of single nephrons.

Mesangial cells produce extracellular matrix, which provides support for glomerular capillaries. The mesangial matrix is composed of collagens (type I, III, IV, and V), laminin, fibronectin, heparan sulfate proteoglycans and chondroitin sulfate

proteoglycans.[74, 75] In addition, mesangial cells produce many growth factors including insulin-like growth factor (IGF), platelet-derived growth factor (PDGF), transforming growth factor β (TGF-β), interleukin (IL-1) and other factors that are essential for the homeostasis of podocytes and glomerular endothelial cells.[76-78]

Mesangial cell proliferation and matrix expansion have been reported in several diseases, such as membranoproliferative glomerulonephritis and IgA nephropathy. As the mesangial matrix expands, it affects the glomerular capillaries by reducing the area available for filtration and may eventually even occlude the capillary lumen.[79-81]

1.1.6 The Bowman's Capsule and the Parietal Epithelial Cells

Bowman's capsule surrounds the glomerular capillary tuft and is formed by a basement membrane associated to the parietal epithelial cells (PECs) that are flat cells linked by tight junctions.[82] PECs are involved in the reduction of permeability, contractility and mechanosensation of Bowman's capsule.[82] Recently, studies in both mice and humans showed that PECs migrate to replace lost podocytes, and perhaps tubular cells.[83, 84]

1.2 The complement system

1.2.1 The three pathways of complement activation

Complement system is a network of proteins that make part of the innate immune system, it acts as a bridge between innate and adaptive antibody-mediated immunity, and plays a fundamental role in the clearance of immune complexes and cell debris.[85] The complement system is activated through three different main pathways, namely the alternative, the classical and the lectin pathway, that converge to the formation of the C3 convertases, enzymatic complexes that cleave Complement component 3 (C3) into C3a and C3b. C3b may opsonize foreign and non-self surfaces such as bacteria or bind to C3 convertases to form the C5 convertases and initiate the terminal complement pathway (Figure 1.2).[85]

C3 is constantly spontaneously hydrolyzed at low level ("tick over"). In the alternative pathway (AP), factor B binds to hydrolyzed C3 [C3(H_2O)] or to C3b followed by Factor D that cleaves Factor B to form the AP C3 convertases [C3(H_2O)Bb and C3bBb]. This C3 convertase cleaves further C3 molecules to form C3b. A local increase of the C3b concentration leads to a rapid amplification of the AP positive loop.[85, 86]

The classical pathway (CP) is triggered by the binding of the Complement complex 1 (C1) to immunoglobulins G (IgGs) or immunoglobulins M (IgMs) bound to antigen. The C1 complex is composed of C1q, C1r and C1s molecules. C1 recognizes and binds to the Fc portion of the IgG or IgM in the immune complex. C1s then cleaves Complement components 4 (C4) and 2 (C2) to form the CP C3 convertase, C4bC2a. In addition pentraxin can recognize pathogens and eliminate them by direct binding to C1q.[85, 86]

The lectin pathway is activated when either mannose-binding lectin or ficolin bind to

carbohydrate moieties on surfaces of pathogens including yeast, bacteria, parasites and viruses. Both mannose-binding lectin and ficolin circulate in the serum as complexes with other associated proteins (Mannan-binding lectin Serine Proteases, MASPs). Binding to the pathogens induces conformational changes resulting in autoactivation of MASP2 (also known as MAP19) which cleaves C4 to form C4a and C4b. C4b attaches to the surface of the pathogens inducing C2 to bind, which is in turn cleaved by MASP2 to form C2b and C2a. C4b together with the attached C2a has enzymatic activity and forms the lectin pathway C3 convertase C4bC2a.[85, 86]

C3 convertases C3bBb of the alternative pathway, C4bC2a of the classical and lectin pathways, cleave the C3 to C3a and C3b. The newly generated C3b acts as an opsonin promoting phagocytosis and helps to amplify complement activation. The addition of a C3b molecule to the C3 convertases leads to the formation of the C5 convertases: C3bBbC3b and C4bC2aC3b. The C5 convertases cleave Complement component 5 (C5) to form C5a and C5b. The membrane attack complex (C5b-9, MAC), also called terminal complement complex (TCC), is then initiated by Complement components 6 (C6) and 7 (C7) binding to C5b followed by Complement component 8 (C8) and multiple molecules of Complement component 9 (C9) that bind to the C5bC6C7 complex. The C5b-9 complex forms a pore by inserting itself into cell membranes, resulting in cell lysis.[85, 86]

In addition to the three pathways, proteases released by neutrophils and macrophages, such as kallikrein, plasmin and factor XIIa (Hageman) can generate complement activation products. Studies in C3 deficient mice, which are unable to generate the conventional C5 convertase, showed that thrombin, a member of the coagulation pathway, can generate C5a in vivo.[87]

The anaphylatoxins C3a and C5a are small octapeptides that act as potent

chemoattractants for neutrophils and monocytes to sites of inflammation, act as vasodilators, induce smooth muscle contraction, histamine release from mast cells and oxidative bursts from neutrophils. They are implicated in the cytokine production and exert pleiotropic effects binding their receptors, C3aR and C5aR (CD88).[88, 89]

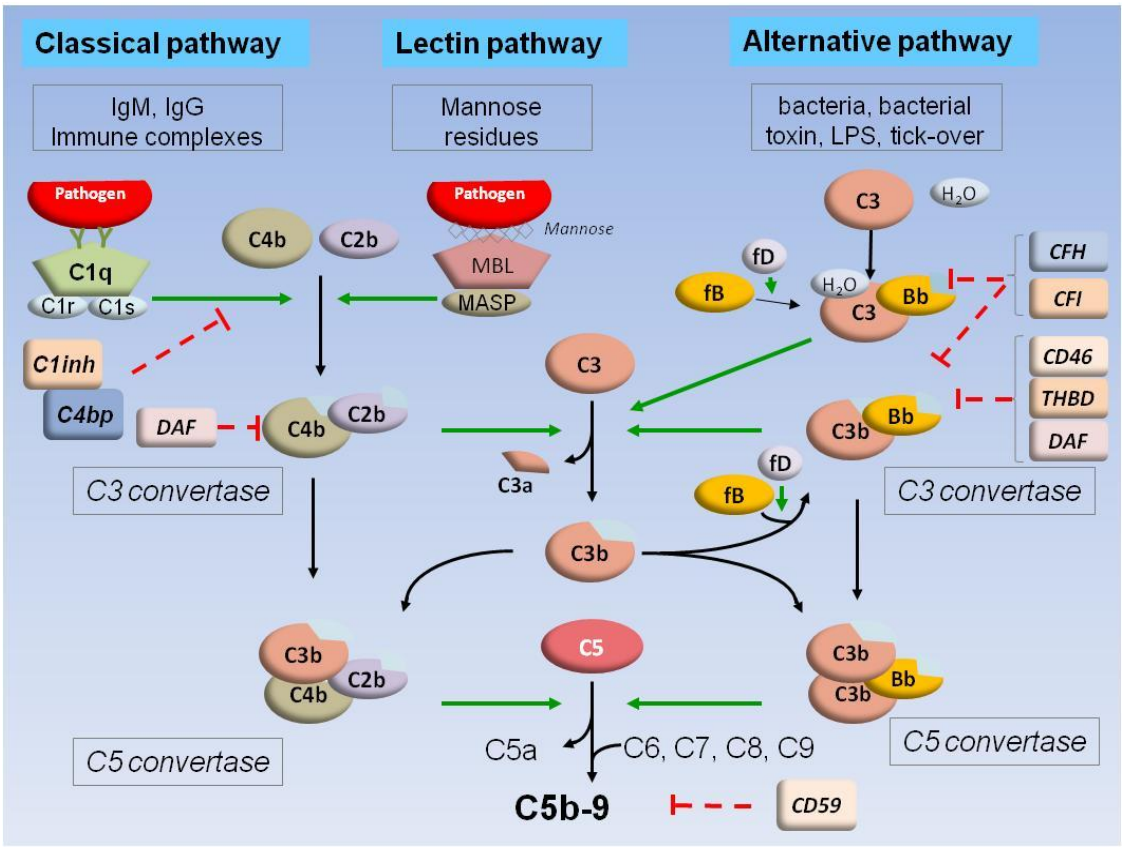


Figure 1.2. Graphical representation of the three pathways of activation of the complement system (based on Noris, N Engl J Med. 2009, and Noris, Semin Nephrol. 2013).[88, 95]

1.2.2 Regulation of the complement system

The inflammation generated by C3a and C5a is quickly reduced by plasma carboxypeptidases that cleave the C-terminal Arginine, resulting in C3a des-Arg and

C5a des-Arg, each of which has less than 10% of their original biological activity.[86]

C3b and C4b are also quickly inactivated by proteolytic cleavage into fragments iC3b, C3dg, C3c, and into C4c, C4d, respectively. This occurs by the serine protease Factor I (FI) in the presence of cofactors, which are present at the cell membrane (membrane cofactor protein, MCP/CD46, and complement receptor 1, CR1/CD35) or in plasma (Factor H, FH). The thrombomodulin (THBD) transmembrane protein also binds to C3b accelerating its FI-mediated inactivation in the presence of FH. Complement activation is also regulated by either preventing the assembly of the C3 convertase or, once it is formed, by inhibiting its activity through the decay acceleration factor (DAF/CD55), and C4 binding protein (C4BP). Finally, the C1 inhibitor (C1-INH) inactivates C1r, C1s and MASP2.[9]

The five FH-related proteins (FHR1 to 5), together with FH, comprise a family of structurally related proteins. Due to the high homology of FHR to FH, it is believed that the genes encoding for *CFHRs* derive from *CFH* duplication and subsequent chromosomal rearrangments.[90, 91] Similar to FH, the FHRs are involved in the alternative pathway of complement system. FHR1, FHR2 and FHR5 proteins form homodimers or heterodimers that compete with FH for binding to surface-fixed C3b, thus reducing the negative control of FH and enhancing complement activity.[91]

Properdin is a positive regulator of the complement system. It is released by activated neutrophils, stabilizes the AP C3 convertase by binding to C3b and preventing its cleavage by FI and FH.[86]

Finally, the TCC formation is controlled both before the integration into the membrane and the pore formation, by a set of different membrane and plasma regulators. The abundant S-protein (also called vitronectin), clusterin (also called SP-40), lipoproteins, anti-thrombin III, proteoglycans, such as heparin, and protamine, are all able to bind to

nascent C5b-7, preventing its membrane insertion. Vitronectin and clusterin bind to the nascent amphiphilic C5b-9 complex rendering it water soluble and lytically inactive. These water soluble complexes dissociate from the membrane and accumulate in solution (SC5b-9).[92, 93] In the final steps of TCC assembly, subsequent to C5b-7 insertion, host cells are protected from lysis due to autologous complement attack by CD59, a glycolipid-anchored membrane molecule, which interferes with the particular C9 interaction site on the C8α-chain that is needed for membrane insertion and subsequent polymerization of C9.[94]

1.3 Nephrotic syndrome

1.3.1 Epidemiology and clinical features

Nephrotic syndrome (NS) is a disorder clinically defined by massive proteinuria, and hypoalbuminemia. Edema and hyperlipidemia are also frequently present. The 70% of cases occur in children under 6 years of age. The annual incidence is around 2-7 cases per 100,000 children, with a prevalence of 16 cases per 100,000 children. The incidence in adults is 3 new cases per 100,000 individuals/year.[96]

The proteinuria is considered in the nephrotic-range when the protein excretion is >3.5 g/24h in adults and >40 mg/h/m^2 in children (despite the definition is more difficult in the young patients because of the large body size range). Hypoalbuminemia, a low level of albumin in blood, is defined as albuminemia <3.0 g/dl, although some authors use a 2.5 or a 3.5 cut-offs.[96, 97]

Edema is observed in the majority of patients with NS. Edema is defined as an abnormal accumulation of fluid in the interstitial compartment. Edemas are mainly localized in slope regions like foot or sacrum, and in regions with low tissue pressure such as periorbital region. In more severe cases, edema is generalized and may involve lungs leading to life-threatening conditions.[96]

The NS patients are also characterized by hyperlipidemia, high concentrations of lipids in blood, including total cholesterol, low density lipoprotein (LDL) and triglycerides. The hypercholesterolemia is the most frequent abnormality present in the 85% of patients. The less frequent is the hypertriglyceridemia showed only when the serum albumin is below 1-2 g/dl. Lipiduria, the excretion of fatty casts in urine, is also reported.[96, 98]

Some complications are frequent in patients with NS. In these patient thromboembolic events such as renal-vein thrombosis may occur. Development of cardiovascular disease is another important complication of NS in patients with prolonged clinical courses. There is a 5-fold increase of the risk for myocardial infarction in NS patients compared controls.[99] The risk of infections is also increased, especially in children. Primary bacterial peritonitis is the most common of infection with an incidence of 25-5% in NS children.[100] Due to loss of proteins other than albumin, other disorders may manifest including endocrine abnormalities (hypothyroidism), microcytic anemia and vitamin D deficiency.[98]

1.3.2 Pathophysiology of Nephrotic Syndrome

Independently of the underlying cause of NS, the abnormally increased permeability of the GFB leads to loss of proteins mainly with a molecular weight between 40-150 kDa, such as albumin, IgG, transferrin, ceruloplasmin and α1-acid glycoprotein. Small amounts of higher molecular weight proteins (200 kDa), such as the small forms of high density lipoprotein (HDL), can also be lost. However, proteins higher than 200 kDa such as IgM, macroglobulins, fibrinogen, XIII factor, fibronectin and large lipoproteins are not lost even in case of severe proteinuria.[101]

Albumin is the most abundant plasma protein and represents 70-90% of the proteins detected in the urine of NS patients. A fraction of the filtrated albumin is catabolized by the renal tubule increasing its catabolic rate. To compensate the albumin loss in urine, albumin synthesis in liver increases up to 300%. Hypoalbuminemia appears when the proteinuria and the renal albumin catabolism exceed its hepatic synthesis. The severity of the hypoalbuminemia is well correlated with the amount of proteinuria although other factors such as the age, the nutritional state and the type of renal lesion also

contribute.[98]

To explain edema in NS, there are two different hypotheses: the *underfilling hypothesis* and the *overfilling hypothesis*. In the *underfilling hypothesis*, the hypoalbuminemia determines a decrease in the intravascular oncotic pressure leading to fluid extravasation. The response to the hypovolemia is the activation of the renin angiotensin aldosterone system, the sympathetic nervous system and the release of anti-diuretic hormone. Finally, the plasma volume normalizes as a result of the increase of the extracellular space and the edema formation. On the other hand, according to the *overfilling hypothesis*, patients with nephrotic syndrome have a primary defect that enhances distal tubular sodium and water reabsorption, independently of the hemodynamic situation. The consequences of this defect are intravascular volume expansion that causes increased intravascular hydrostatic pressure and subsequent fluid extravasation leading to edema formation. This hypothesis is based on the observation that NS patients show resistance to Atrial Natriuretic Peptide (ANP), a peptide that is able to reduce the blood pressure by decreasing the amount of water, sodium and fat in the circulation. Another explanation is that the intrarenal sodium retention is due to activation of the epithelial sodium channel in the collecting duct. The precise cause of the edema and its persistence is still uncertain. The finding that most NS patients have normal or elevated intravascular volume suggests that the overfilling hypothesis is probably the most reliable. Nonetheless, the two hypotheses are not mutually exclusive and the volume status may depend on the stage of disease.[98]

Several factors could determine the hypercholesterolemia in NS. Early studies suggested that the reduction of the intravascular oncotic pressure may stimulate the synthesis of lipids and apolipoproteins by the liver, however the high cholesterol and LDL levels are rather caused by a decrease in lipoprotein due to the low level of tissue

low density lipoprotein receptors.[98] The predominant mechanism responsible for hypertriglyceridemia is a delayed triglyceride catabolism. The reduced catabolic rate is likely secondary to diminished lipoprotein lipase activity due to downregulation of the enzyme gene expression and to the presence of lipase inhibitors.[102]

Hyperlipidemia and hypertension observed in patients with NS are factor that promote atheromatosis. In addition, NS may lead to proteins involved in the coagulation and the immune response favoring a pro-inflammation and pro-coagulant state. This may explain the high risk of cardiovascular events in NS patients.[98, 103]

The susceptibility to infections is associated to the low serum levels of IgG due to urinary IgG loss. Other factors that may be involved are a dysregulation of T lymphocytes and decreased levels of complement factors B and D, resulting in a decreased ability to opsonize encapsulated bacteria. In addition, NS treatment widely use immunosuppressive drugs including steroids, cyclosporine, mycophenolate and others that further increases the risk of infection.[98]

1.3.3 Classification of disorders associated with NS based on pathogenesis

NS may be caused by a variety of glomerular or systemic diseases that determine the alteration of the size or the charge selectivity of the GFB. The forms that are associated with glomerular diseases intrinsic to the kidney and not related to systemic causes are defined as *primary*, also called *idiopathic*, because the disease cause is usually unknown. The NS is instead defined *secondary*, if the etiology is extrinsic to the kidney. In all cases, injury to glomeruli is an essential feature.

The patients can be classified based on the pattern of glomerular injury. There are several glomerular lesions including Focal Segmental Glomerulosclerosis (FSGS),

Minimal Change disease (MC), Congenital Nephrotic syndrome (CNS) of Finnish Type, Diffuse Mesangial Sclerosis (DMS), Membranoproliferative Glomerulonephritis (MPGN), Membranous Nephropathy (MN), Immunoglobulin A nephropathy (IgAN), and others.

MCD, DMS, FSGS are also defined as *podocytopathies*[104] because patient podocytes show foot process effacement, with a structural disorganization of the glomerular filtration barrier leading to proteinuria. The most common primary glomerulonephritides that can be associated with NS are described in section 1.4 of this chapter.

Secondary NS is due to systemic diseases that affect other organs (such as diabetes mellitus, hypertension and amyloidosis), malignancies (including colon or lung cancer), autoimmune diseases (including systemic lupus erythematosus, monoclonal gammopathy, Henoch-Schönlein Purpura), infections (including Human Immunodeficiency Virus, hepatitis B and C virus, Treponema pallidum), drug exposure (including heroin, mercury), and many others.[105]

Noteworthy, complex syndromes with extra-renal involvement have been associated with NS, such as Denys-Drash syndrome, Pierson syndrome, Nail-Patella syndrome, Alport syndrome and Schimke immuno-osseous dysplasia. These diseases are usually associated to genetic defects that may also be found in not syndromic NS patients (see section 4 of this chapter).

1.3.4 Classification of primary forms of NS based on therapy

The therapy of idiopathic NS aims to counteract the causes leading to glomerular injury and also the derived systemic complications.

Glucocorticoids (prednisone or prednisolone) are first-line treatment in NS. They show anti-inflammatory and immunosuppressive properties, acting on lymphocytes and inflammatory cells through the inhibition of the NF-κB transcription factor pathway.[106] The rationale for immunosuppressive therapy relies on the evidence that immune alterations have been identified among affected patients.[107] Specifically, it appears that in some patients T cells promote the production of a circulating factor that alters the glomerular permeability of the filtration barrier.[108] It is estimated that about 80% of patients with idiopathic NS will respond to corticosteroids with complete resolution of proteinuria and edema. Among this steroid-responsive group, the clinical course is variable, with up to 60% having frequent relapses or becoming dependent on steroid therapy to maintain them in remission. Based on these findings, it became important to establish some clinically relevant definitions for the diagnosis of NS and to clarify various patient responses to treatment. Patients who enter remission in response to corticosteroid treatment alone are referred to as having steroid-sensitive NS (SSNS), while patients who fail to enter remission after corticosteroid treatment are referred to as having steroid-resistant NS (SRNS). Some patients respond to initial corticosteroid treatment by entering complete remission but develop a relapse either while still receiving steroids or within 2 weeks of discontinuation of treatment following a steroid taper. Such patients typically require continued low-dose treatment with steroids to prevent development of relapse, and are therefore referred to as having steroid-dependent NS (SDNS). Patients who enter complete remission in response to steroids, remain in remission for several weeks following discontinuation of treatment but develop frequent relapses are defined as frequently relapsing NS (FRNS) if relapses occur 4 or more times in any 12-month period.[109]

Some patients with SRNS, FRNS or SDNS might benefit from other immunosuppressive treatments, such as alkylating agents, calcineurin inhibitors or

Rituximab, which however have significant toxicity profiles.[110] *Cyclosporin A* (CsA) and *Tacrolimus* are calcineurin inhibitors typically used in NS patients and both determine a proteinuria reduction in NS patients.[111, 112] Calcineurin inhibitors inhibit cytokine production from T-helper cells (Th1 and Th2) and have an inhibitory effect on antigen-presenting cells.[113] CsA may also reduce proteinuria through stabilization of the actin cytoskeleton in kidney podocytes through the inhibition of the calcineurin mediated dephosphorylation of synaptopodin.[59] Alkylating agents, such as *Cyclophosphamide*, cause a depletion of immune competent cells but the exact mechanisms of action are still unknown.[114, 115] Among the antiproliferative agents, *Mycophenolate Mofetil* has the minor nephrotoxic potential.[116]

Rituximab is an anti-CD20 chimeric murine/human antibody. CD20 is largely expressed on B-lymphocytes thus treatment with Rituximab cause a depletion of circulating B-lymphocytes, but had no effect on B-cell precursors or differentiated B-cell (plasma cell) populations [117]. Rituximab showed an unexpected efficacy in SDNS, inducing rapid and long-term remission even in multi-relapsing patients who did not respond to treatments primarily active on T-cells (including CsA, Cyclophosphamide and Tacrolimus).

Proteinuria is treated with renin-angiotensin-aldosterone system (RAAS) blockers including both, angiotensin-converting-enzyme inhibitors (ACEi) and angiotensin II receptor antagonists (AII-RA) which lower blood pressure, prevent worsening of kidney disease, and reduce the amount of protein excreted in the urine. The antiproteinuric effects of ACEis are determined to their ability to reduce glomerular capillary plasma flow rate, decrease transcapillary hydraulic pressure, and alter the permselectivity of the GFB.[118-120]

General measures to control edema include salt restriction, lying in supine position, and

use of diuretics drugs. Moderate exercise, a low cholesterol diet, hydroxymethylglutaryl-CoA reductase inhibitors, and weight lost in obese patients are recommended to reduce hyperlipidemia. Hygienic, dietetic and pharmacological measures are needed to counteract the high morbidity and mortality associated with NS and its complications.

1.3.5 Prognosis

The introduction of immunosuppressive therapies has contributed to decrease the NS mortality to 9%.[121] Nowadays, the single most important prognostic indicator for maintenance of long-term normal renal function in idiopathic NS is the patient's initial response to corticosteroids. Overall, close to 80% of newly diagnosed children treated with corticosteroids will achieve complete remission.[122] Among patients with SSNS, relapse is common. Risk factors for frequent relapses or a steroid-dependent course may be age of less than 5 years at onset and a prolonged time to initial remission.[123, 124]

It is estimated that 40% to 50% of patients with SRNS will progress to End Stage Renal Disease (ESRD) within 5-10 years of diagnosis, despite aggressive immunosuppression. For this subgroup, the ultimate treatment goal is renal transplantation, which can also pose serious challenges. NS recurs in the allograft in up to 30% of children with FSGS and leads to graft loss in about 50% of such patients.[125, 126] Patients who have familial forms of SRNS usually do not respond to immunosuppressive treatment and often progress to ESRD requiring renal transplantation. In most of these patients the genetic disease does not recur in the renal allograft, although the overall risk for disease recurrence is controversial: NS has been noted in a subset of patients because of immunological attack on a new antigen encountered for the first time in the transplanted kidney.[127-129]

1.4 Common primary glomerulonephritides associated to nephrotic syndrome

Idiopathic nephrotic syndrome can be associated with a variety of histological lesions and their frequency depends on the age at disease onset (Table 1.1). These histological types reflect distinct underlying mechanisms, although overlapping of both clinical features and pathogenesis may occur.

Table 1.1. Frequency of Primary Glomerulopathies in patients with Nephrotic Syndrome

Type of Primary Glomerulopathy	Frequency (%)	
	Childhood	**Adults**
Minimal change nephropathy	55 to 76%	15 to 20%
Focal segmental glomerulosclerosis	8 to 31%	23 to 35%
Membranous nephropathy	<1 to 7%	33 to 36%
IgA nephropathy	1 to <5%	8 to 9%
Membranoproliferative glomerulonephritis	4 to 5%	2 to 6%
Mesangial proliferative (other than IgAN)	<5 to 22%	na
Other	1 to 5%	<7%

Data obtained from Orth (1998)[130] and Bonilla-Felix (1999)[131] for children, and from Korbet (1996)[132] and Haas (1997)[133] for adults.

1.4.1 Podocytopathies

1.4.1.1 Minimal change disease (MC)

Histopathological features

On light microscopy (LM), MC is characterized by minimal or no glomerular alterations. The "minimal" findings include podocyte swelling and mild mesangial expansion. MC was also known in the past as lipoid nephrosis due to the lipids found in

the renal tubular cells, as well as to the presence of lipid-laden proximal tubular cells and macrophages in the urine.

Commonly no glomerular immune deposits are detected on immunofluorescence (IF). In a minority of patients, there may be weak mesangial positivity for IgM, IgA and/or C3.

Electron microscopy (EM) shows extensive effacement of foot processes in the absence of other abnormalities of the capillary walls. Slit-diaphragms are almost always obliterated.

Clinical features

Annual incidence of primary MC is 2.0 per 100,000 individuals in children and 0.6 per 100,000 individuals in adults.[134, 135] MC is the most common histopathological pattern in children with idiopathic NS accounting for 55-76% of cases (Table 1.1). Patients with MC usually manifest the diseases at age 2-6 years. MC also accounts for 15-20% of idiopathic NS in adults. The incidence of MC is a higher in elderly than middle-aged adults. More than 90% of children and 77-96% of adults go to remission after appropriate steroid treatment. The long term outcome is excellent for steroid sensitive patients. The risk of ESRD is <3% in patients with initial response to corticosteroid, while up to 50% steroid non responders develop ESRD within 10 year.[133, 136]

1.4.1.2 Focal Segmental Glomerulosclerosis (FSGS)

Histopathological features

The term FSGS describes a histological pattern of injury characterized on LM by

scarring (sclerosis) in some portions of the glomerular capillary tuft (segmental) and only in some glomeruli (focal). As the disease progresses, a more diffuse and global pattern of sclerosis evolves and tubular atrophy and interstitial fibrosis develop.

FSGS can be further divided in the following subtypes: i) collapsing; ii) tip-lesion; iii) cellular variant; iv) perihilar variant; v) FSGS not otherwise specified (NOS).[137] The order from *i* to *v* is important because the assignment to one subtype presumes that all the previous have been excluded. The collapsing variant is characterized by implosive glomerulus tuft collapse. The tip variant is characterized by the adhesion of the tuft's most peripheral portion with the Bowman's capsule at the urinary pole. This occurs by extracellular matrix adhesion and/or confluence of podocytes with parietal or tubular epithelial cells.[137] In the cellular variant there is segmental endocapillary hypercellularity associated with variable glomerular epithelial cell proliferation. The perihilar variant is defined as perihilar hyalinosis and sclerosis involving the majority of glomeruli with segmental lesions.[137] FSGS is defined as not otherwise specified when it cannot be assigned to any other variant.[137] It is the most common histological subtype[138]. Repeated biopsies have shown that other variants may evolve into FSGS-NOS over time.

On IF, FSGS typically shows IgM and C3 coarse segmental deposits entrapped in areas of sclerosis.

On EM, an extensive effacement of the foot processes is observed without other abnormalities in the GBM.

Clinical features

Annual incidence of primary FSGS is about 0.1 per 100,000 individuals in children and

0.8 per 100,000 individuals in adults.[134, 135] Approximately 75-90% of children and 50-60% of adults with focal segmental glomerulosclerosis have the nephrotic syndrome at presentation.[139] Hematuria may also be present. The renal outcome of FSGS is similar in children and in adults, although children may sometimes show a better response to therapy. Forty to 69% of nephrotic patients treated for appropriately long period with steroids attain complete remission. More than 90% of the responders show stable renal function after 10 years after the clinical onset, while more than 60% of patients who do not enter remission progress to ESRD.[140-142]

1.4.1.3 Diffuse mesangial sclerosis (DMS)

Histopathological features

On LM, DMS is characterized by mesangial matrix expansion, with no mesangial hypercellularity, and sclerosis that conduces to the obliteration of the capillary lumen and to the contraction of the glomerular tuft.

No glomerular immune deposits are found on IF, although nonspecific staining for IgM, C3, and C1q may be identified in the mesangium of some glomeruli.

On EM, effacement of foot processes is present.

Clinical features

DMS is characterized by early onset of NS (in the first year of life) and rapid progression to end stage renal disease. This entity has been described as an isolated disorder or also a symptom of syndromes with renal and extrarenal abnormalities such as Denys-Drash or Pierson Syndrome.

1.4.1.4 Congenital Nephrotic syndrome of Finnish Type

Histopathological features

On LM, the first signs of the disease are increased mesangial cellularity and matrix, and the characteristic focal cystic dilatations of proximal tubules, followed by radial dilatations of the proximal tubules, mesangial proliferation and focal segmental sclerosis.

Immunofluorescence does not detect immune deposits, although, IgM and C3 staining can be observed as the disease progresses.

EM shows effacement of podocyte foot processes and disappearance of the slit diaphragm.

Clinical features

Nephrotic syndrome is defined as congenital (CNS) when it is manifested in the first 3 months of life. CNS of Finnish type (CNF) represents the most severe form of NS.

The disease is manifested in the fetus and can be suspected around 16-18 weeks of gestation by elevated α-fetoprotein levels in the maternal serum. The placenta is abnormally large, while the fetus is small for gestational age in 25% of cases.

Proteinuria at birth and it is very high ranging from 1 to 6 g/day. Microhematuria and signs of tubular dysfunction, such as aminoaciduria and glycosuria, may also be present. Renal function is usually normal at birth but it progress to ESRD within the first three years of life. The mortality is high due to the complications of NS, especially infections and sepsis.

1.4.1.5 Genetic causes of Podocytopathies

The first gene identified as a cause of SRNS was NPHS1 that underline the importance of the slit-diaphragm and the podocyte in the pathogenesis of the disease.[143]

To date more than 30 genes have been described to carry disease-causing mutations for SRNS that encode for proteins involved in different aspects of the podocyte and the glomerular basement membrane. These proteins are involved in the slit diaphragm, the podocyte cytoskeleton, the podocyte development and regulation of homeostasis, mitochondrial energy metabolism, the nuclear pore complex, and the glomerular basement membrane and the cell-matrix interactions.[144] Interestingly, defects in these proteins can cause isolated glomerulopathy, complex syndromes, where the nephropathy is associated with extra-renal manifestations, or both.

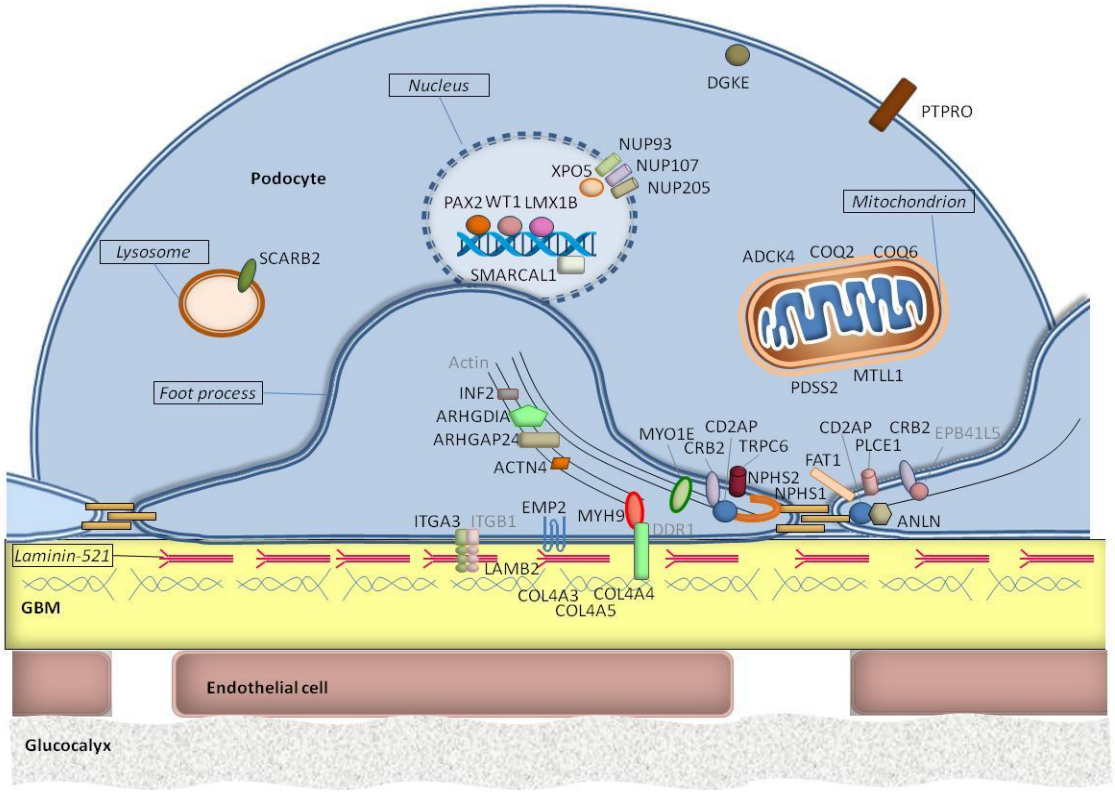


Figure 1.3. Graphical representation of the glomerular filtration barrier and of some genes associated with podocytopathies. In grey genes that are not associated with nephropathy in humans.

Slit diaphragm

There is evidence that mutations affecting proteins that make part of the podocyte slit diaphragm, such as nephrin (*NPHS1*), podocin (*NPHS2*), calcium channel Transient Receptor Potential cation Channel 6 (*TRPC6*) and CD2-Associated Protein (*CD2AP*), cause SRNS.[145, 146]

The *NPHS1* encodes for nephrin, a major component of the slit diaphragm (see chapter 1.1.4). Mutations in *NPHS1* are characterized by an autosomal recessive inheritance and are the major cause of CNS in Finland (causing the CNF) due to a founder effect of two mutations (L41fs*91 and R1109*) that account for more than 94% of all mutations.[143] Both mutations lead to truncated protein products causing the absence of nephrin from the podocyte slit diaphragm.[143] These mutations are rare in non-Finnish populations. The frequency of *NPHS1* mutations in CNS patients is lower worldwide than in Finnish patients, accounting for 39% to 80% of cases.[147-150] Less frequently, *NPHS1* mutations are also found in SRNS patients with age of disease onset between 6 months and 12 years that show a pattern of MC or FSGS on renal biopsy.[151, 152] In addition, a case with adult-onset SRNS carrying two compound heterozygous *NPHS1* mutations has been described.[153] In at least one patient with later-onset NPHS1-associated SRNS a milder effect of the mutation on the nephrin function has been demonstrated.[151]

NPHS2 encodes for podocin that physically interacts with nephrin at the slit diaphragm. Mutations in *NPHS2* are characterized by autosomal recessive inheritance and are identified more frequent in familial cases (30-40% of cases) compared to sporadic cases (10-30% of cases).[128, 147, 154-157] *NPHS2* mutations are the main cause of SRNS in patients with age of onset between 4 months and 12 years, and are also found in congenital (15-39%), adolescent (3-11%) and adult forms (5-16%).[150, 152, 158, 159] Interestingly, in adults the presence of two mutated alleles is rare, while it is frequently

observed the combination of one mutation with the R229Q variant in *trans*. R229Q is a common polymorphism with an allele frequency of 2.7% in the general population and leads to a disease phenotype only when it is associated specifically with certain mutations affecting the C-terminal portion of the podocin because of an altered heterodimerization and entrapment of the 229Q-podocin in the endoplasmic reticulum.[160] *NPHS2* mutations are associated with a FSGS diagnosis in kidney biopsy in approximately 80-96% of cases and the remaining patients show a pattern of MC.[128, 150, 156] NPHS2 pathogenic mutations include missense, nonsense, frameshift, and splicing variants. The most prevalent mutation in European patients is the *NPHS2* R138Q that accounts for 32-44% of all affected *NPHS2* alleles and has an allelic frequency of 0.16% in the European non-Finnish general population of the Exome Aggregation Consortium (ExAC) database (v0.3).[128, 156] In congenital and infantile SRNS, the 94% of patients with NPHS2 mutations carry at least one truncating mutation or homozygous R138Q mutations.[147] Missense mutations such as the p.R138Q or the p.R168H characterized by retention of the mutant protein in the endoplasmic reticulum result in an earlier onset and a more severe phenotype compared to mutations that do not disrupt the mutant protein trafficking to the plasma membrane, such as G92C, V180M and R238S.[128, 161]

TRPC6 encodes for the short transient receptor potential canonical 6. TRPC6 co-localizes and physically interacts with podocin and nephrin. It is critical in adjusting the calcium influx close to the slit diaphragm and is functionally connected to the actin cytoskeleton.[162, 163] Mutations in *TRPC6* are characterized by autosomal dominant inheritance and are found in 0.5-3% of patients both in children and adults.[150, 152, 162] Mutation in this gene are characterized by high phenotypic variability, even within members of the same family, and incomplete penetrance with healthy mutation carriers.[162, 164, 165] Histopathological findings typically show FSGS and occasionally

MC.[164-167] The majority of *TRPC6* mutations are gain-of-function mutations leading to increased channel activity by increasing calcium current amplitudes or by delaying channel inactivation, whereas some loss-of-function mutations have been described.[162, 168]

PLCE1 encodes for the phospholipase C epsilon 1 protein, which catalyzes the hydrolysis of phosphatidylinositol 4,5-bisphosphate to generate inositol 1,4,5-trisphosphate (IP3) and diacylglycerol (DAG), thus opening DAG-sensitive classical transient receptor channels including TRPC6.[168] Mutations in *PLCE1* show an autosomal recessive pattern of inheritance.[169] On renal biopsy, DMS is the most frequently finding but FSGS can also be observed. Most patients carrying *PLCE1* mutations show severe phenotype with early-onset of the disease and rapid progression to ESRD.[169] In four different families with patients carrying *PLCE1* mutation, four unaffected individuals carrying homozygous truncating mutations have been reported suggesting that modifiers modifier genes or environmental factors are required to manifest the disease.[170, 171]

CD2AP encodes for the CD2-Associated Protein, a cell adhesion molecule found on the surface of T lymphocytes and natural killer cells that is implicated in T-cell adhesion to antigen-presenting cells. In podocytes, CD2AP interacts with nephrin, podocin and F-actin suggesting its role in anchoring the slit diaphragm to the actin cytoskeleton.[50, 172, 173] *CD2AP* mutations cause FSGS with both autosomal recessive and autosomal dominant mode of inheritance. Accordingly to a recessive mode, two compound heterozygous *CD2AP* mutations in a congenital form and one homozygous mutation in an infantile form have described.[174, 175] Both parents of the homozygous patient carried a heterozygous mutation and were asymptomatic.[175] Accordingly to a dominant mode, 5 patients carrying heterozygous mutations have been reported with prevalently adult-

onset.[176-178] Similar to humans, Cd2ap deficient mice present congenital NS with rapid progression to ESRD, while Cd2ap haploinsufficient mice show only mild glomerular abnormalities in more advanced age.[176] *CD2AP* mutations were not identified in a large international cohort, suggesting a low contribution of this gene to SRNS.[152]

Podocyte Cytoskeleton

SRNS can also be caused by mutations in proteins involved in the organization of the podocyte cytoskeleton such as inverted formin 2 (*INF2*), α-actinin-4 (*ACTN4*), myosin 1E (*MYO1E*), Rho GDP dissociation inhibitor-α (*ARHGDIA*), RhoA-activated Rac1 GTPase-activating protein (*ARHGAP24*), and anillin (*ANLN*).[145, 146, 179, 180]

The INF2 protein has the unique ability to accelerate both polymerization and depolymerization of actin.[181] The α-actinin-4 protein belonging to the spectrin gene superfamily that binds actin and interconnects actin filaments.[182] MYO1E is localized at the plasma membrane of foot processes in podocytes, and is involved in podocyte cytoskeleton organization and motility.[179, 183] ARHGAP24 is upregulated in podocytes during their differentiation and inhibits RAC1and CDC42 activity thus suppressing lamellipodia formation and cell spreading.[180]. ARHGDIA interacts with RHO GTPases, including Ras Homolog Family Member A (RHOA), Ras-Related C3 Botulinum Toxin Substrate 1(RAC1), and Cell Division Cycle 42 (CDC42), and sequester them in an inactive state in the cytosol.[184] Finally, ANLN is an F-actin binding protein that is enriched in the cytoskeleton.[185]

The inheritance is autosomal dominant for *INF2*, *ACTN4*, *ARHGAP24* and ANLN, while it is autosomal recessive for *MYO1E* and *ARHGDIA*.

Typically, patients with *INF2* mutations present FSGS with moderate proteinuria and

adolescent or early-adult onset. However, there is high phenotypic variability. Nephrotic range proteinuria and NS at onset have been described. The age of clinical presentation ranges from 5 years to 72 years, while the age of progression to ESRD ranges from 13 to 70 years.[186-188] Intra-familial variability is also widely present in *INF2*-assocaited FSFS and incomplete penetrance of the disease is also common.[187, 189, 190] Interestingly, INF2 mutations are also frequently detected in patients with Charcot-Marie-Tooth disease associated with FSGS.[191] Mutations *INF2* account for about 9-17% of familial cases and less than 1% of sporadic cases.[152, 187-189] To date dozen of mutations have been reported and all of them are located within the diaphanous inhibitory domain (DID) of INF2. Mutations in exons 2 to 4, which encode for the DID domain, account for more than 90% of all mutations.[187-189, 192] Some reported mutations are recurrent with certain codons appearing to be hotspots. The R218Q, R218W, R214H and R214C mutations were described in 6, 2, 4 and 4 independent families, respectively.[187-189] A possible founder effect of these mutations has been excluded, since no common ancestral disease-associated haplotype was found.[187]

Patients carrying *ACTN4* mutations present FSGS with adolescent or adult onset, although childhood onset cases have been described. Like other dominant FSGS-causing genes the disease presents incomplete penetrance.[193-196] Mutations in the ACTN4 gene were reported to account for approximately 2% of autosomal dominant FSGS.[188] However, no mutations were identified in a large cohort, suggesting a low contribution of *ACTN4* in the pathogenesis of FSGS.[152] Animal models and *in vitro* studies suggest both gain-of-function and loss-of-function mechanisms for *ACTN4* mutations.[197] Most of the described ACTN4 mutations are missense involving residues within or next to the actin binding site of α-actinin-4.[182, 188, 193, 194, 198] Some of these mutations increase the affinity of α-actinin-4 to F-actin and/or lead to protein mislocalization followed by formation of α-actinin-4/F-actin aggregates around the

nucleus with subsequent impairment of podocyte spreading and motility.[193, 194, 197, 199, 200]

MYO1E mutations cause a rare autosomal recessive form of SRNS that is commonly manifested in the first decade of life. On renal biopsy FSGS is mainly observed, although a case with MC has been described.[174, 179] Mutations in the *MYO1E* gene account for the 0.3% of SRNS cases, but may involved in the pathogenesis of a significant number of families that have some degree of consanguinity.[152, 174, 179] The available functional studies show that at least in some mutations lead to impaired function or loss of important MYO1E domains thus causing abnormal subcellular localization of the protein and reduction of podocyte motility.[179]

ARHGAP24 mutations are associated with severe, early-onset FSGS.[180] In a large cohort, mutations in this gene accounted for 0.06% of mutations.[152] Functional studies showed that mutated *ARHGAP24* results in a marked increase in the level of active Rac1causing dysregulation of the RhoA-Rac1 signaling.[180]

All reported patients with homozygous or compound heterozygous *ARHGDIA* mutations showed an early onset of the disease (3 months to 2.4 years), rapid progression to ESRD and DMS on renal biopsy.[201, 202] In some patients, extra-renal features such as intellectual disability, sensorineural hearing loss, seizures or cortical blindness were present. *ARHGDIA* mutations account for 0.06% of mutations in SRNS patients.[152] Functional studies showed that all three described mutant GDIα proteins present impaired actin polymerization, smaller cell size, increased cellular projections and slow podocyte motility.[201-203]

Heterozygous missense mutations in ANLN have recently been described in two families with FSGS. The age at proteinuria onset ranged from 9 to 69 years and ESRD occurred between 35 and 75 years.[185] The mutant anillin displays reduced binding to the

slit diaphragm-associated CD2AP. Knockdown of *ANLN* in zebrafish causes a loss of glomerular filtration barrier integrity, podocyte foot process effacement, and an edematous phenotype.[185]

Glomerular Basement Membrane and GBM-Podocyte Interactions

Another group of genes associated with SRNS are those encoding proteins involved in focal adhesions that tether basal surface of the podocytes to the underlying glomerular basement membrane including integrins *ITGA3* and *ITGB4*, *EMP2*, *LAMB2*, and collagens *COL4A3*, *COL4A4* and *COL4A5*.

The diseases associated with mutations in *LAMB2* gene, encoding for the β2 chain of laminin-521, show an autosomal recessive inheritance. Mutations in LAMB2 are associated with Pierson syndrome, which is characterized by congenital NS with histological lesions of DMS and ocular malformations (microcoria, abnormal lens with cataracts, and retinal abnormalities).[204] Childhood onset of Pierson syndrome has also been reported with progressive visual impairment leading to blindness at approximately 2 years of age, and NS and ESRD between 5 and 10 years of age.[205] Interestingly, missense *LAMB2* mutations are also found in patients with congenital NS with no or only minor ocular changes.[206-208] Complete loss-of-function mutations, such as truncating mutations, seem be associated with the complete Pierson syndrome, whereas missense mutations may display variable phenotypes ranging from a milder variant of Pierson syndrome to isolated congenital NS.[147, 206-208]

Mutations in the genes encoding for the components of the collagen IV α3α4α5 protomer are also associated with NS. *COL4A5* is characterized by X-linked recessive inheritance, although affected females can also be observed usually with a milder

phenotype. In the *COL4A3* and *COL4A4* genes, both an autosomal dominant and recessive mode of inheritance have been described. Traditionally, mutations in these genes are associated with Alport syndrome, a glomerulopathy associated with variable ocular anomalies and sensorineural hearing loss. Typical renal features in Alport syndrome are microscopic hematuria with progressive proteinuria and, eventually, renal failure.[21] The allelic disease thin basement membrane nephropathy (TBMN) is characterized by persistent microscopic hematuria, but progressive proteinuria and ESRD are rarely reported. Finally, mutations in *COL4A3–5* genes have been found in patients who had a primary diagnosis of FSGS with or without NS, even though on in-depth clinical evaluation at least some of them presented features of Alport syndrome, such as characteristic findings on electron microscopy and hearing loss.[21, 209] Remarkably, also Alport syndrome and TBMN show FSGS on LM. Inheritance of Alport syndrome is X-linked in 65% of cases (in patients with *COL4A5* mutations) and dominant (20%) or recessive (15%) in other cases (that is, in patients with *COL4A3* and *COL4A4* mutations).23[21] In adult-onset FSGS mutations in *COL4A3–5* may be found in up to 20% patients, mainly in the *COL4A4* gene.[210]

The *EMP2* gene encodes for Epithelial Membrane Protein 2, a transmembrane protein that regulates cell-membrane localization of integrins, caveolins, and glycosylphosphatidyl-inositol-linked proteins. In podocytes and endothelial cells, EMP2 deficiency results in an increased caveolin-1 levels and decreased cell proliferation. It is characterized by autosomal recessive inheritance. Patients with EMP2 mutations may present SRNS or SSNS.[211] Patients describe to date presented a disease onset in the first 3 years of life.[211]

The *ITGA3* encodes for the integrin alpha-3 subunit. Integrin alpha-3 is present in different tissues. In podocytes, it makes part of the integrin α3β1 heterodimer that it is

localized at the basal membrane interacts with the GBM component laminin-521.[10] Mutations in this gene are characterized by autosomal recessive inheritance and are found in syndromic cases with congenital nephrotic syndrome, interstitial lung disease and epidermolysis bullosa.[212] In a large SRNS cohort, *ITGA3* mutations were found in 0.3% of patients.[152]

Mitochondrial dysfunction

NS is also observed in patients carrying mutations in genes encoding for proteins involved in the biosynthesis of Coenzyme Q10 (CoQ10), such as Prenyl Decaprenyl Diphosphate Synthase Subunit 2 (*PDSS2*), Coenzyme Q6 FAD-dependent Monooxygenase (*COQ6*), Coenzyme Q2 Polyprenyltransferase (*COQ2*) and AarF Domain Containing Kinase-4 (*ADCK4*).[213-216] CoQ10 is a lipophilic molecule composed of a benzoquinone and a decaprenyl side chain. It is present in all cell membranes and acts as an electron carrier in the mitochondrial respiratory chain transporting electrons from complexes I and II to complex III. In addition, CoQ10 is an antioxidant, a membrane stabilizer, a regulator of mitochondrial permeability transition pores and a cofactor for uncoupling proteins in brown adipose tissue.[213] The PDSS2 protein is the first enzyme of the CoQ10 biosynthetic pathway and synthesizes the prenyl side-chain of coenzyme Q10. Coenzyme Q6 FAD-dependent Monooxygenase is required for the C5-ring hydroxylation. Coenzyme Q2 Polyprenyltransferase acts in the final steps in the CoQ10biosynthesis. The ADCK4 protein resides both to podocyte foot processes and to the mitochondria interacts COQ6 and COQ7 of the CoQ10 biosynthesis pathway.[214]

PDSS2, COQ6, COQ2 and *ADCK4* abnormalities are all characterized by autosomal recessive inheritance and account for approximately 0.1%, 0.5%, 0.2% and 0.2-2% of SRNS patients, respectively.[152, 214]

Mutations in *PDSS2* cause Leigh syndrome, a congenital disease characterized by encephalomyopathy and nephrotic syndrome; however, isolated nephropathy may also be present.[213]

Mutations in *COQ6* cause a syndrome characterized by nephrotic syndrome with sensorineural deafness.[216] In the reported patients, the age of onset ranges from 0.2 to 6 years. The main histopathologic feature is FSGS but also DMS may be observed.[216]

Patients carrying the CoQ2 mutations manifest the disease within the first two years of life. Inherited COQ2 mutations cause a primary glomerular disease with renal lesions that vary in severity and are not necessarily associated with neurological signs.[215] On kidney biopsy focal segmental glomerular sclerosis, crescentic glomerulonephritis is observed.[215, 217]

The *ADCK4* gene is associated with a nephropathy that is typically manifested in adolescence with mild to moderate proteinuria and occasionally hematuria. On kidney biopsy, patients show a FSGS pattern of injury.[214, 218] In most patients the disease is limited in the kidney, however some patients show features compatible with neurologic dysfunction.[214, 218]

Podocyte Development

Mutations are also found in genes involved in the regulation podocyte development and homeostasis including transcription factors *WT1*, *PAX2* and *LMX1B*, and the Crumbs complex component *CRB2*.[110]

PAX2 regulates the levels of GDNF and c-Ret that mediate the ureteric bud induction in early kidney development.[219] PAX2 represses *WT1* expression in the presence of groucho and transducin-like enhancer proteins, while it induces *WT1* expression in the

absence of these proteins. Following the expression of *WT1*, PAX2 is down-regulated in the podocyte.[220] LMX1B in podocytes binds to the promoters *NPHS2*, *CD2AP*, *COL4A3* and *COL4A4* inducing their expression.[220] Mutations in *PAX2*, *WT1* and *LMX1B* show an autosomal dominant inheritance. *WT1* and *LMX1B* mutations account for 5% and 0.2% of SRNS patients, respectively.[152] Mutations in PAX2 account for approximately 4% of families with FSGS.[221]

The *WT1* gene encodes the Wilms Tumor protein 1. Abnormalities in this gene are associated with a wide phenotypic spectrum of diseases ranging from Wilms tumor to syndromic forms of glomerular disease and genitourinary abnormalities including Denys-Drash, Frasier Syndromes and WAGR syndrome (Wilms tumor, Aniridia, Genitourinary abnormalities, and mental Retardation).[222-225] Patients with *WT1* mutations may also manifest isolated SRNS.[226-230] The type and localization of the mutations correlate with the renal and extra-renal phenotypes.[230] Nephropathy secondary to WT1 mutations is usually manifested in the first decade of life but it may also occur in adolescents and adults. Histological findings show DMS in most patients and FSGS in a minority.[230]

PAX2 mutations have been first associated with congenital abnormalities of the kidney and urinary tract (CAKUT) as part of a syndrome known as renal coloboma syndrome (also called papillorenal syndrome).[231] Renal coloboma syndrome is highly variable even in patients carrying the same mutation within the same family and some affected individuals may show only renal features.[232] Moreover, PAX2 mutations have been found in a significant portion of patients with primary FSGS. The disease is commonly manifested from the second to the fourth decade but it may be observed from children to elderly individuals.[221]

Mutations in *LMX1B* cause Nail-Patella syndrome, a disorder characterized by dysplasia

of the patellae, nails and elbows associated to iliac horns, glaucoma and in 50% of cases glomerulopathy. On renal biopsy a pattern of FSGS is observed. *LMX1B* mutations are also identified in patients with primary FSGS and no extra-renal manifestations.[233, 234] The age of onset is highly variable and ranges from 5 to 70 years of age, and some of them reached ESRD between 26 and 70 years of age.[233, 234]

The *CRB2* gene encodes the transmembrane Crumbs homolog 2 protein that is required for podocyte foot process arborization, slit diaphragm formation, and proper nephrin trafficking.[235] Following an autosomal recessive inheritance, homozygous or compound heterozygous *CRB2* mutations are identified in patients with childhood-onset isolated SRNS with age of onset ranging from 9 months to 6 years. FSGS is observed on kidney biopsy.[235] Of note, CRB2 mutations may also be detected in fetuses and infants with cerebral ventriculomegaly and echogenic kidneys, in which histopathological features of congenital nephrosis are found.[236] *CRB2* mutations very rare in SRNS patients.[152]

Nucleoporins

Recently, mutations in nucleoporin genes, such as nucleoporin 93 (*NUP93*), 107 (*NUP107*), 205 (*NUP205*), and exportin 5 gene (*XPO5*), have been described as a cause of autosomal recessive SRNS. Nucleoporins make part of the nuclear pore complex (NPC), a huge protein complex embedded in the nuclear envelope, which mediates the transport of proteins, RNAs and ribonucleoprotein particles between the nuclear interior and the cytoplasm, Nucleoporins exert their function by interacting with transport receptors such as exportins and importins that shuttle their cargo through the NPC. *NUP93*, *NUP107*, *NUP205* encode for nucleoporins 93, 107 and 205, respectively, while *XPO5* encodes for exportin 5. NUP93 physically interacts with NUP205 within the inner ring of NPC. NUP93, NUP205 and exportin 5 are expressed in developing

podocytes. In addition, exportin 5 colocalizes with synaptopodin in foot processes of rat podocytes.[237, 238] Mutations in *NUP93*, *NUP107*, *NUP205* and *XPO5* have been described in SRNS patients with age of onset between 1 and 12 years. On kidney biopsy most of the patients showed a pattern of FSGS, while in a minority MC and MDS was observed. Noteworthy, in addition to the glomerular phenotype, tubular abnormalities were commonly observed.[237, 238]

Apical membrane

The *PTPRO* gene encodes for the transmembrane podocyte protein Glepp1, which is localized in the apical membrane of podocytes and is involved in the regulation of the glomerular pressure and filtration rate.[239, 240] PTPRO has also been associated to SRNS in two Turkish consanguineous families (five individuals) with autosomal recessive inheritance.[239] Patients with *PTPRO* mutations had a disease onset between 5 and 14 years of age. On kidney biopsy FSGS or MCD with severe foot process effacement was observed. Studies on animal models showed that Ptpro-deficient mice present modification of podocyte structure, hypertension and glomerular filtration rate but no proteinuria.[240] Screening in a large cohort of 1783 unrelated, international families found no mutations in *PTPRO*.[152]

Primary Cilium

The *TTC21B* gene encodes the retrograde intraflagellar transport protein 139 (IFT139). IFT139 is a component of the intraflagellar transport-A complex that associates with the motor protein dynein2 to regulate retrograde trafficking in the primary cilium.[241] The primary cilium is a microtubule-based organelle with an antenna-like structure that is

present on the surface of most cells and plays a crucial role during development in sensing flow changes and mediating signaling pathways involved in the establishment of cell polarity.[242] The *TTC21B* gene is characterized by autosomal recessive inheritance. Mutations in *TTC21B* gene have been identified in nine individuals with different ciliopathies including isolated nephronophthisis, nephronophthisis with extra-renal manifestations and Jeune Asphyxiating Thoracic Dystrophy.[243-245] Ttc21b-null mice are characterized by embryonic lethality and ciliopathy-like features.[246] Recently, TTC21B mutations have been identified in patients from seven unrelated families with FSGS and tubulointerstitial lesions. Clinically these patients were characterized by late-onset proteinuria (9-23 years), high blood pressure and ESRD at ages 15-32 years. Renal biopsies show FSGS together with typical alterations of nephronophthisis including tubular basement membrane thickening.[247]

1.4.2 Immune-complex-mediated Membranoproliferative glomerulonephritis and C3 glomerulopathy

Histopathological features

Membranoproliferative glomerulonephritis (MPGN) is a pattern of glomerular injury observed in light microscopy characterized by mesangial hypercellularity, endocapillary proliferation, and capillary-wall remodeling with double contour formation.[248] On immunofluorescence (IF), C3 staining is observed. Immunoglobulin deposits are also frequently detected. These are usually IgG and/or IgM deposits but also IgA and C1q can be found.[249]

Traditionally, through electron microscopy (EM), MPGN was divided into type I, with subendothelial deposits, type II or Dense-Deposit Disease (DDD), with intramembranous electron-dense deposits, and type III, with subendothelial and subepithelial deposits.[248] A better understanding of MPGN's pathogenesis has led to reclassification into immune-complex-mediated MPGN (IC-MPGN, also called immunoglobulin-associated MPGN, Ig-MPGN) and complement-mediated MPGN.[248, 250, 251] IC-MPGN shows immunoglobulin and C3 positive staining on IF, supporting glomerular immune-complex deposition as the trigger of classical complement pathway. Complement-mediated MPGN, named C3 glomerulopathy (C3G), presents predominant C3 staining resulting from alternative complement pathway dysregulation.[250-253] Predominant C3 glomerular staining is defined as C3 intensity ≥2 orders of magnitude more than any other immune-reactant on a 0 to 3 scale.[253] C3G is further divided into DDD and C3 glomerulonephritis (C3GN), the latter lacking the characteristic linear or ribbon-like intramembranous highly electron-dense deposits. Notably, the term C3G is also used to define other proliferative patterns (mesangial and endocapillary) or even nonspecific alterations sharing C3-dominant glomerular staining.[253] In all subtypes,

mesangial, subendothelial and subepithelial deposits may be found on EM.

Clinical features

MPGN commonly presents with the nephritic syndrome (microscopic hematuria, non-nephrotic proteinuria, and renal insufficiency). However, up to one third of the patients show relatively preserved renal function and nephrotic syndrome at onset.[254] Following the historical classification, types I and III MPGN are more common in children. The majority of patients have microscopic hematuria, and some have macroscopic hematuria. Hypertension may be present at the diagnosis and is less common in children. MPGN type II/Dense deposit disease commonly presents in children between the ages of 5 and 15 years. Clinically is similar to type I and III, although nephrotic range proteinuria is observed in approximately 50% of patients.[254]

According to the current IC-MPGN/C3G classification, microhematuria rate is similar between the subtypes and is present in 51-76% of patients. On the other hand, nephrotic syndrome is more frequently observed in IC-MPGN (65%) compared to DDD (38%) and C3GN (27%). Age of onset is higher in C3GN (median age 26 years) than in DDD (median 12 years).[255, 256] Finally, the prevalence of patients with low serum C3 levels is higher in DDD compared to C3GN.[255, 256] The risk to progress to ESRD within ten years from onset is 25-50%.[255, 256]

Pathogenesis

The presence of glomerular C3 deposits in IC-MPGN and C3G underline the importance of the complement system activation in these disorders. The association of significant immunoglobulin and C1q deposits suggested an activation of the

complement system through the classical pathway. Indeed, immunoglobulin deposits are frequently observed in the secondary forms of MPGN. Thereby, the presence of idiopathic IC-MPGN was questioned, and it is believed that only few IC-MPGN patients remain idiopathic after extensive evaluation.[248, 253, 257, 258] On the other hand, the presence of glomerular C3 deposits with no or scanty immunoglobulins implicate complement activation through the alternative (or the lectin) pathway.[248, 252]

Both acquired and genetic abnormalities are found in IC-MPGN and C3G. The most frequently involved is C3NeF. It is an IgG auto-antibody that binds and stabilizes the alternative pathway C3 convertase C3bBb leading to over-production of activated C3b. C3NeF is found in approximately 80% of DDD patients but also in about 40% of C3G and 40% of IC-MPGN patients.[256] Other less frequently observed acquired abnormalities include autoantibodies against complement factors H (FH) or B (FB).[259, 260]

Genetic abnormalities also account for a smaller portion of IC-MPGN and C3G patients. Mutations are found in the two components of the alternative pathway C3 convertase, C3 and Factor B.[261-263] Mutations are also identified in complement regulator proteins such as FH, factor H-related 5, factor I (FI) and CD46/MCP.[256, 262] In a study where both IC-MPGN and C3G patients were screened for the *CFH, CFI* and *CD46/MCP* genes, genetic abnormalities were identified in 17% of IC-MPGN patients, in 17% of DDD and in 20% of C3GN patients (19% in all C3G patients).[256] In a successive study in which the *C3*, *CFB*, *CFH*, *CFI*, *CFHR5*, and *CD46/MCP* were analyzed a positive genetic diagnosis was provided in 43% of patients carrying the clinical diagnosis of C3G.[262]

Together with rare pathogenic genetic variants, common genetic variants increase the risk to present DDD. Indeed, DDD patients segregate a complotype given by the

combination of the risk alleles 102G and 314L in C3, and 402H and V62 in FH. This complotype is associated with higher AP activity in healthy control subjects, which is consistent with the interaction of these two proteins.[264] These data are supported by *in vitro* studies showing that C3(102G) activates complement alternative pathway more efficiently than C3(R102). The C3 convertase negative regulator FH binds more strongly to C3(R102) than the C3(102G). FH cofactor activity in CFI-mediated C3 inactivation is reduced for C3b(102G) compared to C3b(R102), favoring alternative pathway amplification. Finally, C3 102G and FH V62 combined with the CFB R32 common variant have a 6-fold higher hemolytic activity compared to the combination of the variants C3 R102, FH 62I and CFB 32Q.[265] Altogether, these data show that DDD is a complex genetic disease.

1.4.3 Membranous glomerulonephritis

Pathological features

The hallmark in MN is the uniform thickening of the glomerular basement membrane caused by granular subepithelial IgG deposits.[266] The pathological evolution of the glomerular lesions has been divided in four stages.

In stage 1, light microscopy may show a normal or slightly thickened GBM. In stage 2, the GBM is prominent, GBM projections toward urinary space, called spikes, may be seen with appropriate staining. In stage 3, the deposits are incorporated within the GBM. In stage 4, there is an irregular thickening of the GBM due to reabsorbed deposits. In patients showing complete remission, the GBM deposits may disappear; the GBM turns normal but some lucent areas may remain. In patients with progressive disease, segmental sclerosis and tubule-interstitial fibrosis may arise.[266, 267]

On immunofluorescence microscopy, granular polyclonal IgG deposits (often IgG4) are observed along the subendothelial portion of glomerular basement membrane. Complement C3 can also be associated with the IgG deposits.[266, 267]

Electron microscopy shows subepithelial deposits with increasing matrix spike reaction during disease progression. Foot process effacement is also extensively present. In stage 3, the GBM can envelop the deposits though matrix reaction, thus resulting in a laddering appearance. The deposits may become rarefied as the disease becomes chronic (stage 4). Mesangial deposits are absent or rare in primary MN.[266, 267]

Clinical features

Membranous nephropathy primarily affects middle-aged and elder adults. The disease

may be idiopathic or secondary to infections (e.g., hepatitis B and C), tumors (e.g. lung and colon cancer), systemic autoimmune diseases (lupus erythematosus) or drugs (e.g., gold and penicillamine).

Annual incidence of primary membranous nephropathy has been estimated as 1.2 per 100,000.[134, 135] It is clinically characterized by proteinuria, which is usually nonselective. Microscopic hematuria is also commonly present. About 60%-80% of MN patients show nephrotic syndrome at disease presentation and 60% of the remaining patients develop nephrotic syndrome during disease course. Arterial hypertension and renal impairment at clinical onset is observed in approximately 20% of patients.[266, 268]

Clinical outcome is variable and difficult to predict individually. Spontaneous remission occurs in about one third of the patients, usually within the first 2 years after onset. The remaining patients can be divided equally into those with persistent proteinuria that maintain renal function long term and those patients who will progress to ESRD despite immunosuppressive therapy. Predictors of spontaneous remission are baseline proteinuria <8 g/day, female sex, age younger than 50 years, and preserved renal function at presentation.[268]

Detection of circulating anti-PLA2R antibodies is useful for diagnosis, since they seem to be specific for MN, even though they can also be detected in secondary forms. The finding of PLA2R on podocytes on renal biopsy can also aid in the diagnosis of membranous glomerulonephritis. Circulating anti-PLA2R1 titer correlates with disease activity, prognosis and response to treatment, as well as recurrences in kidney graft. During treatment, a decrease in antibody titers often precedes proteinuria response and can identify responders from patients refractory to treatment.

Pathogenesis

MN is an autoimmune disease caused by antibody formation against podocyte antigens. So far, the role of three different antibodies has been established that recognize human endogenous antigens such as neutral endopeptidase (NEP), M-type phospholipase A2 receptor 1 (PLA2R1) and thrombospondin type-1 domain-containing 7A (THSD7A).

Antibodies against an exogenous antigen, the cationic bovine serum albumin (BSA), have been described in children younger than 5 years affected by MN, however the source of the cationic BSA remains obscure.[268]

The anti-NEP autoantibodies were the first to be discovered and explain some cases with congenital onset. Anti-NEPs are produced in mothers with neutral endopeptidase deficiency that become alloimmunised to the paternally inherited NEP antigen expressed in the placenta. Maternal anti-NEP antibodies cross the placenta and bind to NEP expressed on fetal glomerular podocytes. Therefore the disease is manifested in the newborn.

Beck et al. showed that in white people 70% of primary MN cases represent autoimmune responses against the M-type PLA2R, a type I transmembrane glycoprotein expressed on glomerular podocytes.[269] These anti-PLA2R1 can be detected in glomerular subepithelial immune deposits on kidney biopsy and on podocytes. These findings have been widely confirmed and circulating anti-PLA2R autoantibodies are found in 52–82% of MN patients.[270]

Auto-antibodies against THSD7A are another cause of MN. THSD7A is a large 250kDa protein that is located at the basal surface of the podocyte. Anti-THSD7As are found in 5-10% of MN patients that are negative for anti-PLA2R1.[271]

IgG4 autoantibodies against aldose reductase (AR), SOD2, and α-enolase (αENO) have

been described in 34%, 26% and 43% of MN patients. Such levels were significantly higher in MN compared to controls and patients with other glomerulopathies. In this study, 60% of patients carried anti-PLA2R1, while no patient had anti-NEP antibodies. Coexistence of more types of auto-antibodies was observed with 10% of patients being positive for all three anti-AR, anti-SOD2 and anti-αENO.[272] However, these findings are reported only by one group, and anti-AR and anti-SOD2 may be found in other diseases; therefore further validation for these autoantibodies is required.

Subsequently, a genome-wide association study in white patients of European ancestry showed that the rs4664308 polymorphisms of the PLA2R gene combined with the rs2187668 polymorphisms in the HLA-DQA1 locus increase the risk for primary MN. The risk alleles of these two genes have an additive effect; individuals carrying all four risk alleles have an odds ratio of 78 for developing MN compared to those carrying all four protective alleles.[273] These findings were confirmed in an independent Spanish cohort (Table 1.2).[274]

Table 1.2. Summary of GWAS results from membranous nephropathy cohorts.

Study	Cohort	dbSNP locus	Locus			*P* value	Odds ratio (95% CI) per allele
			Chr	Minor allele	Nearest gene		
Stanescu *et al.* (2011)	556 white European MN patients and 2,338 controls	rs4664308	2	A	*PLA2R1*	8.6×10^{-29}	2.28 (1.96–2.64)
		rs2187668	6	G	*HLA-DQA1*	8.0×10^{-93}	4.32 (3.73–5.01)
Bullich *et al.* (2014)	89 white European MN patients and 286 controls	rs4664308	2	A	*PLA2R1*	0.03	1.53 (1.05–2.23)
		rs2187668	6	G	*HLA-DQA1*	2.5×10^{-6}	2.58 (1.72–3.87)

Modified from Jiang (2013).[275] MN: membranous nephropathy; Chr: chromosome.

1.4.4 IgA nephropathy

Pathological features

The diagnosis of IgA nephropathy (IgAN) is based on a histopathologic evaluation of renal biopsy, which shows mesangial hypercellularity and predominant immunoglobulin A (IgA) deposition in the glomeruli.[276] Nonetheless, histopathologic findings in IgAN are highly heterogeneous.[276]

On light microscopy, the features may vary greatly among patients and within the individual biopsy sample. Increased mesangial matrix and mesangial hypercellularity are commonly observed. Other glomerular lesions may include focal necrosis, segmental scarring, and crescents in Bowman's space.[276]

Immunofluorescence microscopy shows a predominant IgA deposition in the glomerular mesangium, which can be either alone or associated with IgG and/or IgM. The frequency of IgA deposits alone varies across centers from 0 to >85%. In addition to immunoglobulin deposits, complement C3 and properdin are almost always observed. C4 or C4d, mannose-binding lectin, and terminal complement complex (C5b–9) are frequently present, whereas C1q is usually absent.[276]

On electron microscopy, it is usually observed electron-dense material corresponding to immune deposits on immunofluorescence. These deposits are generally located in mesangial and paramesangial areas but they can be occasionally also present in subepithelial and subendothelial portions of glomerular basement membranes.[276]

An international consensus working group developed the Oxford classification of IgA nephropathy based on four pathological features: mesangial hypercellularity, segmental glomerulosclerosis, endocapillary hypercellularity, and tubular atrophy/interstitial fibrosis. These features have prognostic significance and the classification predicts more

accurately the risk of progression of renal disease.[277]

Clinical features

IgAN is a major cause of kidney failure worldwide. The incidence of IgAN is estimated as 2.5 cases per 100,000 adults.[135] The prevalence of this disease is highest in Asia, intermediate in Europe and USA, and lowest in Africa.[134]

IgAN is a highly heterogeneous disorder, in which the clinical presentation ranges from asymptomatic hematuria to rapidly progressive glomerulonephritis evolving in ESRD. The typical presentation is hypertension or mild edema in young oligosymptomatic adults.[134] Microscopic hematuria (persistent or intermittent), episodic gross hematuria and chronic kidney disease are the main characteristic, while NS is observed in 10% of patients.[278] Among patients with NS, 48% show complete remission, 32% partial remission and 20% no response to treatment.[278] Up to 50% of IgAN patients show a progressive loss of glomerular filtration rate, while most of the remaining patients present persistent hematuria and/or proteinuria. Only in a minority of patients sustained clinical remission occurs (spontaneously or after treatment).[134, 279]

The prognosis in IgAN patients is highly variable. At least half of the patients have a benign disease course or remit spontaneously, while about 30%–50% progress to ESRD within 20 years of the first biopsy. Kidney transplantation is the treatment of choice for ESRD, but the disease recurs in up to 50% of patients.[280]

To predict individual outcome, a three-point score was established in a French cohort, based on proteinuria >1 g/day, presence of hypertension, and histological changes in the biopsy sample. This score efficiently predicted the risk of death or ESRD (ranging from 4% to 64% after 20 years in those with a score of 0 or 3, respectively). Moreover, the

current Oxford-MEST classification of IgAN has been developed to better reflect outcome. It is based on four histological features associated with a progressive course: glomerular mesangial hypercellularity; endocapillary hypercellularity; segmental glomerulosclerosis and/or tuft adhesions; and tubular atrophy and/or interstitial fibrosis in more than 25% of the section.

Pathogenesis

IgA nephropathy is likely a systemic disease with extrarenal causes in which the kidneys are damaged as innocent bystanders. This is suggested by the fact that IgA nephropathy frequently recurs in renal allografts. Conversely, IgA deposits in a kidney from a donor with subclinical IgA nephropathy disappeared within several weeks after transplantation in a patient with a different kidney disease.[281]

IgAN is an immune-mediated disease caused by abnormal deposition of deficient O-linked glycan IgA1 antibodies. Type 1 IgA heavy chains are composed of one variable (V_H) and 3 constants regions (Cα1-3). Between Cα1 and Cα2 there is a hinge region with Serine and Threonine residues that can be subject to O-glycosylation (Figure 1.1A). Three to six of the 9 potential glycosylation sites are usually glycosylated. The carbohydrate composition of the O-linked glycans on normal human circulatory IgA1 is variable, with the prevailing glycan being a tetrasaccharide consisting of one N-acetylgalactosamine (GalNAc), one galactose (Gal), and two sialic acid molecules (one linked to GalNAc and one to Gal) (Figure 1.1B). In patients affected by IgAN, a greater fraction of circulatory IgA1 with hinge-region O-linked glycans without galactose is observed. These O-linked glycans consist of only a terminal GalNAc or a sialylated GalNAc (Figure 1.1B) and the corresponding IgA1s are as defined galactose-deficient O-linked glycan (Gd-IgA1).[279, 282]

The IgAN patients have elevated serum Gd-IgA1 levels compared to controls and 50-78% of patients have serum Gd-IgA1 levels above the 95th percentile of healthy controls. Family studies have shown that elevated Gd-IgA1 levels are heritable, with 25-33% of asymptomatic family members displaying levels that are just as elevated as those of patients. The heritability of Gd-IgA1 has been estimated at >50%.

The presence of high Gd-IgA1levels is not sufficient to cause the disease. A multi-hit pathogenetic model has been proposed in which genetic and environmental factors are involved. According to this model, individuals with high Gd-IgA1 levels are prone to auto-sensitization and production of anti–Gd-IgA1 IgG or IgA antibodies results in the formation of immune complexes. These pathogenic immune complexes deposit in the glomerular mesangium activating mesangial cells that produce cytokines, proliferate and excessively produce extracellular matrix leading to renal injury.[279, 283] Interestingly, both the circulating concentrations of Gd-IgA1 and of anti-Gd-IgA1 autoantibodies are associated with IgAN progression.[134]

Four different GWAS studies are available in IgAN that have identified fifteen independent risk loci with genome-wide significance. These loci cumulatively explain about 6-8% of the genetic risk.[279] The susceptibility loci for IgAN implicate defects in adaptive immunity (MHC region), mucosal barrier innate immunity (*VAV3*, *DEFA* gene cluster, *TNFSF* , *CARD9*, *ITGAM*, *ITGAX* and *HORMAD2*) and alternative pathway of complement (*CFHR1*).[279, 282] A IgAN genetic risk score was created as the weighted sum of the number of risk alleles. Subjects carrying all the risk alleles have a 3.71-fold higher risk to develop IgAN than the worldwide average, while subjects carrying all the protective alleles have a 5.03-fold less risk to develop the disease. Therefore, the two extremes show an 18.66-fold difference in IgAN risk.[279]

Six independent susceptibility loci have been identified in the MHC region that involve

prevalently *HLA-DR*/*HLA-DQ* but also *HLA-DP* and *TAP1*/*TAP2*/*PSMB8*/*PSMB9*. Imputation of classical HLA alleles using the data from SNP arrays identified a role for HLA-DQA1*0101, HLA-DQA1*0102, HLA-DQB1*0201 and HLA-DQB1*0301.[279] However, it must be considered that imputation of classic HLA alleles by SNP data may miss or impute with low confidence rare alleles due to the limited size of the reference database given the highly polymorphic HLA locus.[284, 285] The HLA genes are responsible for antigen processing and presentation, and it is hypothesized that they are involved in the genesis of the anti-Gd-IgA1 autoantibodies. One of the MHC region susceptibility loci contains the *TAP1*, *TAP2*, *PSMB8*, and *PSMB9* genes.[279] These are interferon-regulated genes that are involved in antigen processing for presentation by MHC-I molecules and in modulation of cytokine production and cytotoxic T cell response.[280] Interestingly, IgAN patients present increased mRNA levels of PSMB8.[286]

Another group of genes involved in the IgAN pathogenesis includes genes that participate in the innate immune defense of the mucosal barrier. The *TNFSF13* gene, encoding APRIL, is a powerful B cell–stimulating cytokine that promotes IgA1 to IgA2 class switching. The fact that APRIL levels are elevated in the serum of some IgAN patients further supports the role of *TNFSF13* in IgAN.[287] Moreover, *TNFSF13* mutations cause partial IgA deficiency and reduced IgA antibody responses to mucosal immunization both in mice and humans.[288] Another susceptibility locus contains the DEFA cluster that includes genes encoding the α-defensin antimicrobial peptides. The α-defensin peptides are released at mucosal surfaces and have microbiocidal and chemoattractant properties.[279] Mammalian α-defensins exert their bactericidal effects through the conserved triple-stranded β-sheet that creates pores in the target membranes and permeabilizes the target cell envelope.[289] DEFA1, DEFA3 and DEFA4 peptides are synthesized in neutrophils, while (DEFA5 and DEFA6 are constitutively secreted by the intestinal Paneth cells into the gut lumen.[280] GWAS suggest the involvement in the

IgAN pathogenesis also for *ITGAM* and *ITGAX*, which encode for integrins αM and αX, respectively. These integrins mark intestinal dendritic cells that maintain the balance between inflammation and tolerance, and combine with the integrin β2 chain to form leukocyte-specific complement receptors 3 and 4 (CR3 and CR4, respectively). Studies in mice show that ITGAM regulates the IgA-producing plasma cells in the bowel.[290] *CARD9*, encoding for Caspase Recruitment Domain Family Member 9, and *VAV3*, encoding for Vav Guanine Nucleotide Exchange Factor 3, are involved in NF-κB activation and are essential for maintenance of the intestinal epithelial barrier and control of the local inflammatory response to infection. CARD9 mediates intestinal repair and control of bacterial infection after intestinal epithelial injury in mice. Moreover, CARD9 deficiency results in susceptibility to invasive fungal infections. VAV proteins are required for the proper enterocyte differentiation in colon and the prevention of spontaneous ulcerations of intestinal mucosa.[279] The role of *HORMAD2* is not clear, however the same gene was associated with inflammatory bowel disease.[291]

A common deletion involving the two CFH-related genes *CFHR3* and *CFHR1* is protective against IgAN. The *CFH* gene and *CFHR1-R5*, encoding factor H (FH) and five FH-related (FHR) proteins, are located in tandem in the cluster of complement activation regulators at chromosome 1q32 and present a high level of sequence identity. The presence of repeated sequences favors genomic rearrangements through non-allelic homologous recombination and the most frequently observed recombination is the deletion of *CFHR3-CFHR1* (allele frequency of 20% in Europeans), which is strongly associated with anti-FH autoantibodies development in atypical hemolytic uremic syndrome and has a protective role in age-related macular degeneration.[90, 292] FH and FHRs are involved in the alternative pathway of complement system. FHR1 forms homodimers or heterodimers with FHR2 and FHR5 that compete with FH for binding to surface-fixed C3b, thus reducing the negative control of FH and enhancing complement

activity.[91] FHR1 deficiency should lead to reduced complement activation at GBM. The presence of complement C3 and properdin deposits in the glomeruli of IgAN patients, which are frequently associated with the presence of the terminal complement complex (C5b–9) and the absence of C1q,[276] underlines the role of the complement alternative pathway in IgAN and supports a possible involvement of FHR1. The role of FHR3 is unknown.

Environmental factors may also play a role in IgAN pathogenesis. The increased incidence of IgAN in some geographical areas might represent an untoward consequence of protective adaptation to mucosal invasion by local pathogens. Bioinformatic analyses showed a very strong positive association of the IgAN genetic risk score with local pathogen diversity from different geographical regions and in particular with the helminth diversity. Furthermore, some secondary IgAN forms are caused by schistosomiasis, which is a helminth infection. In addition, the ability of ITGAM-positive dendritic cells to stimulate CD4+ T cells is specifically impaired during schistosome infection. The enhanced immune response conferred by risk alleles in IgAN would be multi-locus adaptation response to mucosal infections.[279]

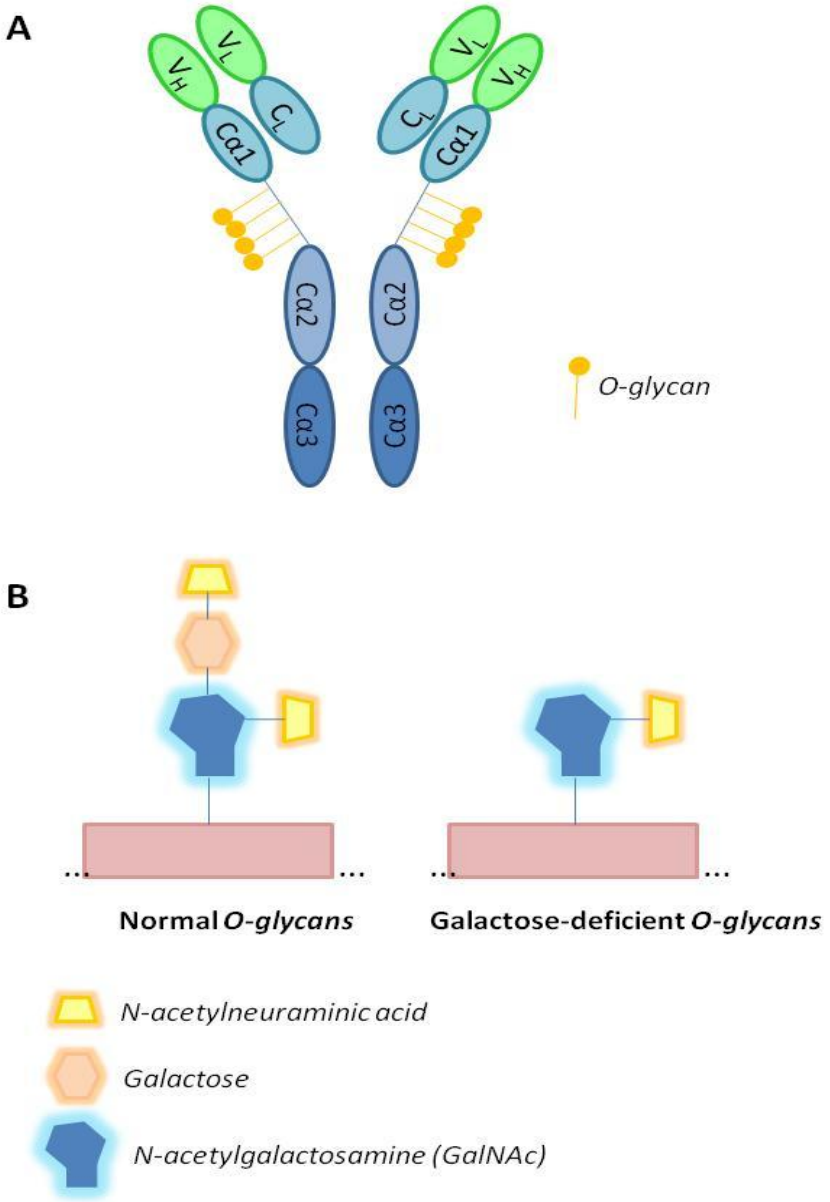


Figure 1.4

Figure 1.4. A) Graphical representation of an IgA1 antibody with the presence of O-glycans in the hinge region. B) Representation of normal and galactose-deficient O-glycans.

Table 1.3. Summary of GWAS results from IgA nephropathy cohorts.

Study	Cohort	dbSNP locus	Locus			*P* value	Odds ratio (95% CI) per allele
			Chr	Minor allele	Nearest gene		
Yu *et al.* (2012)[27]	4,137 Han Chinese IgAN patients and 7,734 controls	rs2738048	8	A	*DEFA*s	3.18×10^{-14}	0.79 (0.74–0.84)
		rs3803800	17	G	*TNFSF13*	9.40×10^{-11}	1.21 (1.14–1.28)
		rs4227	17	T	*MPDU1*	4.31×10^{-10}	1.23 (1.16–1.32)
		rs12537	22	C	*MTMR3*	1.17×10^{-11}	0.78 (0.72–0.84)
		rs2523946	6	T	*HLA-A*	1.74×10^{-11}	1.21 (1.15–1.28)
		rs660895	6	A	*HLA-DRB1*	4.13×10^{-20}	1.34 (1.26–1.42)
		rs1794275	6	G	*HLA-DQA/B*	3.43×10^{-13}	1.30 (1.21–1.39)
Feehally *et al.* (2010)[25]	244 European IgAN patients and 4,980 controls	rs3115573	6	A	*NLNHP4*	2.00×10^{-6}	1.55 (1.29–1.85)
		rs3130315	6	G	*NLNHP4*	2.00×10^{-6}	1.55 (1.29–1.85)
Gharavi *et al.* (2011)[26]	3,144 Han Chinese IgAN patients and 2,822 controls*	rs3766404	1	C	*CFHR1,3*	4.24×10^{-5}	0.77
		rs6677604	1	A	*CFHR1,3*	2.96×10^{-10}	0.68
		rs2856717	6	T	*HLA-DQB1*	8.44×10^{-16}	0.73
		rs9275596	6	C	*HLA-DQB1*	1.59×10^{-26}	0.63
		rs9357155	6	A	*PSMB8*	2.11×10^{-12}	0.71
		rs2071543	6	A	*PSMB8*	5.77×10^{-12}	0.73
		rs1883414	6	T	*HLA-DPB2*	4.84×10^{-9}	0.78
		rs3129269	6	T	*HLA-DPB2*	8.54×10^{-9}	0.79
		rs2412971	22	A	*HORMAD2*	1.86×10^{-9}	0.80
		rs2412973	22	A	*HORMAD2*	4.46×10^{-9}	0.80
Kiryluk *et al.* (2014)	4,418 European and 3,240 Han Chinese IgAN patients and 10,229 European and 2,287 Han Chinese controls	rs17019602	1	G	*VAV3*	6.8×10^{-9}	1.17
		rs6677604	1	A	*CFHR1,3*	4.8×10^{-14}	0.74
		rs7763262	6	T	*HLA-DR/DQ*	1.8×10^{-38}	0.71
		rs9275224	6	A	*HLA-DR/DQ*	5.9×10^{-30}	0.74
		rs2856717	6	T	*HLA-DR/DQ*	1.1×10^{-15}	0.79
		rs9275596	6	C	*HLA-DR/DQ*	2.5×10^{-31}	0.69
		rs2071543	6	A	*TAP2-PSMB9*	1.5×10^{-4}	0.87
		rs1883414	6	T	*HLA-DP*	1.5×10^{-11}	0.82
		rs2738048	8	C	*DEFA*	1.6×10^{-4}	0.91
		rs10086568	8	A	*DEFA*	1.0×10^{-9}	1.16
		rs4077515	9	T	*CARD9*	1.2×10^{-9}	1.16
		rs11150612	16	A	*ITGAM-ITGAX*	1.3×10^{-11}	1.18
		rs11574637	16	C	*ITGAM-ITGAX*	8.1×10^{-13}	0.76
		rs3803800	17	A	*TNFSF13*	9.3×10^{-6}	1.12
		rs2412971	22	A	*HORMAD2*	4.8×10^{-12}	0.83

Modified from Jiang (2013).[275] *All values are given for all cohorts combined (the Beijing discovery cohort, plus both the Shanghai and European replication cohorts).

dbSNP: database of single nucleotide polymorphisms; GWAS: genome-wide association study; IgAN, IgA nephropathy.

2. AIMS

The global aim of the present thesis was to evaluate the contribution of genetics to the pathogenesis and clinical outcome of two kidney disorders characterized by the presence of nephrotic syndrome: the immune complex-mediated membranoproliferative glomerulonephritis (IC-MPGN) and C3 glomerulopathy (C3G), and the steroid-resistant nephrotic syndrome (SRNS).

The first part of this thesis was focused on IC-MPGN/C3G. Specific aims of this part were:

To evaluate the prevalence of complement genetic and acquired abnormalities in a large cohort of patients with idiopathic IC-MPGN/C3G.

To investigate whether specific patterns of complement abnormalities and genetic variants correlate with disease manifestations and whether they can predict the long-term evolution of these complex diseases.

To evaluate whether the current IF-based C3G/IC-MPGN classification captures all patients with complement-mediated pathogenesis.

To explore whether cluster analysis based on histologic findings, clinical features, serum complement profile and genetic data could be a tool for better understanding the disease etiology.

The second part of this thesis aimed to analyze the genetic causes of SRNS and to develop a genetic test based on NGS able to screen the coding regions and the polymorphisms associated with the most common primary proteinuric glomerular disorders. Specific aims of this part were:

To create a NGS based genetic test containing genes associated with the most common primary proteinuric glomerular disorders.

To create a strategy and a score that allows to distinguish true from false variant during NGS in order to increase the sensitivity and reduce the false positive error detection rate of the Ion Torrent Personal Genome Machine (PGM) platform.

To search for mutations in known SRNS-associated genes in patients with the sporadic and the familial form of SRNS, in order to evaluate the prevalence of mutations in patients.

To evaluate whether SRNS patients carry mutations in genes associated with other primary proteinuric glomerular disorders.

To search for new gene/s involved in the pathogenesis of SRNS by screening selected candidate genes in the entire cohort.

To search for new gene/s involved in the pathogenesis of SRNS in familial cases without mutations in the known genes using autozygosity and linkage studies all along the genome and whole exome studies.

3. COMPLEMENT GENE VARIANTS DETERMINE THE RISK OF IMMUNOGLOBULIN-ASSOCIATED MPGN AND C3 GLOMERULOPATHY AND PREDICT LONG-TERM RENAL OUTCOME

3.1. Introduction

Membranoproliferative glomerulonephritis (MPGN) is an uncommon cause of chronic proteinuric nephropathy diagnosed based on a glomerular injury pattern characterized by mesangial hypercellularity, endocapillary proliferation, and capillary-wall remodeling with double contour formation.[248] The onset varies widely with patients presenting with asymptomatic hematuria and proteinuria, nephritic or nephrotic syndrome, or rapidly progressive glomerulonephritis. MPGN may be secondary to infections, autoimmune diseases and malignancies, or idiopathic, when a clear underlying etiology cannot be identified. Traditionally, through electron microscopy (EM), MPGN was divided into type I, with subendothelial deposits, type II or Dense-Deposit Disease (DDD), with intramembranous electron-dense deposits, and type III, with subendothelial and subepithelial deposits.[248] A better understanding of MPGN's pathogenesis has led to reclassification into immunoglobulin-associated (Ig-MPGN; also called immune complex-mediated MPGN, IC-MPGN) and complement-mediated MPGN.[248, 250, 251] Ig-MPGN shows immunoglobulin- and complement-positive staining by immunofluorescence (IF), supporting glomerular immune-complex deposition as the trigger of classical complement pathway. Complement-mediated MPGN, named C3 glomerulopathy (C3G), presents predominant C3 staining resulting from alternative

complement pathway dysregulation.[250-253] C3G is further divided into DDD and C3 glomerulonephritis (C3GN), the latter lacking the intramembranous electron-dense deposits. The term C3G is also used to define other proliferative patterns (mesangial and endocapillary) or even nonspecific alterations sharing C3-dominant glomerular staining.[253]

Although Ig-MPGN is considered an immune-complex mediated disease with complement activation through the classical pathway, the presence of mutations affecting alternative pathway regulatory proteins in a few cases, and C3 nephritic factors (C3NeFs) - autoantibodies that stabilize the alternative pathway C3 convertase - in about 50% of patients, suggests the alternative pathway plays a role.[256]

Alternative pathway dysregulation is well documented in C3G: low serum C3 levels associated with C3NeFs are observed in most patients and few carry autoantibodies against complement factors H (FH) or B (FB), or mutations in genes encoding C3 and the regulatory proteins FH, factor H-related 5, factor I (FI) and CD46.[256, 259, 260, 293-295]

In this study we evaluated the complement genetic and biochemical profile in a large cohort of patients with idiopathic Ig-MPGN/C3G. We also investigated whether specific patterns of complement abnormalities and genetic variants correlate with disease manifestations and whether they can predict the long-term evolution of these complex diseases.

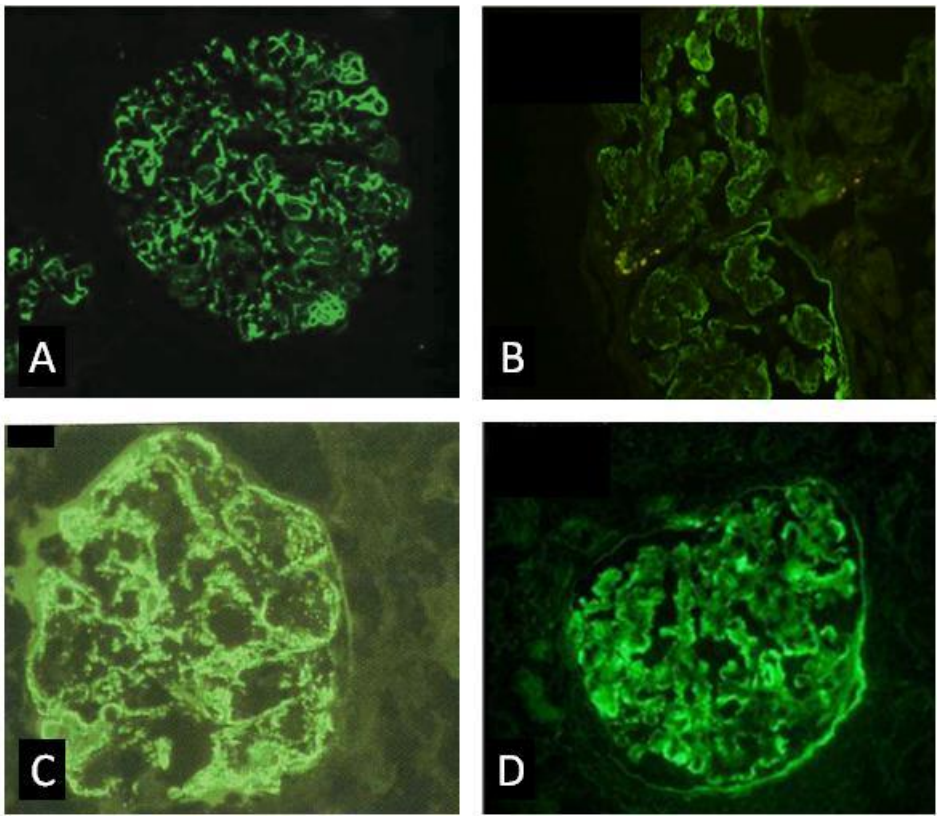


Figure 3.1. Examples of immunofluorescence studies for C3 (A), C4d (B), IgG (C) and C1q (D) in renal biopsy specimens from patients with IC-MPGN. Modified from Sethi, J Am Soc Nephrol 2015,[296], Sethi, Semin Nephrol 2011[250] and Yabuuchi, Case Rep Nephrol Dial 2016.[297]

3.2 Specific aims

1. To evaluate the prevalence of complement genetic and acquired abnormalities in a large cohort of patients with idiopathic Ig-MPGN/C3G.

2. To investigate whether specific patterns of complement abnormalities and genetic variants correlate with disease manifestations and whether they can predict the long-term evolution of these complex diseases.

3.3. Patients and Methods

Patients

Patients (n=227) were consecutively recruited through the Registry of MPGN established at the Mario Negri Institute (Figure 3.2). All kidney biopsy reports were centrally reviewed, adopting the recent classification.[250, 253] MPGN diagnosis was based on the typical light microscopy pattern.[250] Patients with glomerular immunoglobulin and C3 deposits at IF were considered Ig-MPGN. C3G was diagnosed based on the presence of MPGN or mesangial proliferative patterns under light microscopy with "dominant C3" glomerular staining (intensity≥2 orders of magnitude more than any other immune-reactant on a 0 to 3scale).[253] Based on EM, C3G was further classified as DDD or C3GN.[253] Even though only patients with idiopathic MPGN are recruited through our Registry, after a deeper evaluation we excluded 87 patients: 11 did not meet diagnostic criteria, 18 were affected by secondary MPGN following a comprehensive diagnostic work-up, 7 received an MPGN diagnosis years after atypical hemolytic uremic syndrome (aHUS), 4 received an MPGN diagnosis on allograft but the native kidney biopsy was not available, 23 had no IF studies, 15 had no EM, and for 9 DNA samples were not available.

All participants provided informed written consent. The study adheres to the Declaration of Helsinki and was approved by the Ethics Committee of the Azienda Sanitaria Locale of Bergamo (Italy).

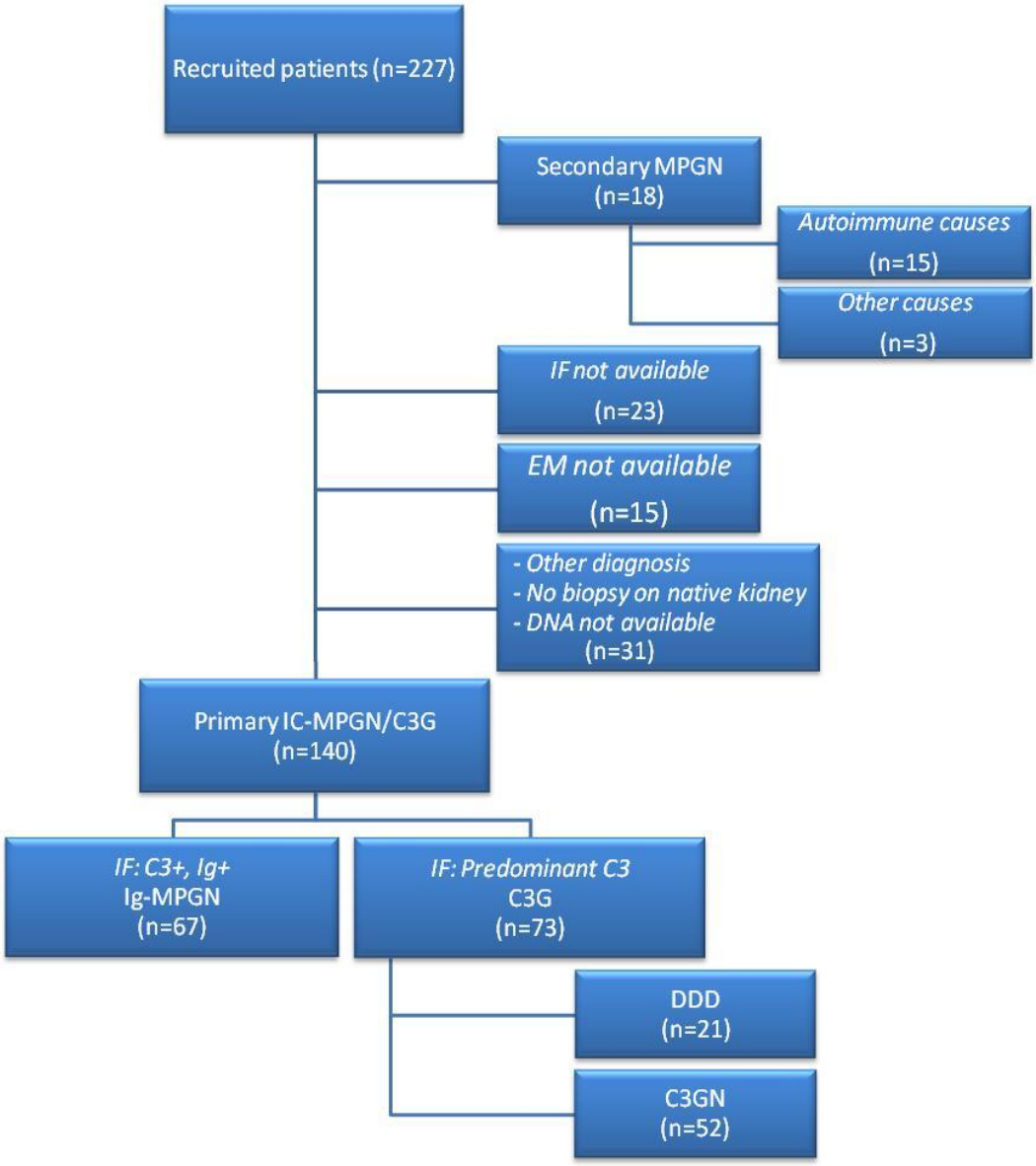


Figure 3.2. Histopathologic classification of the patients recruited through the Registry of Idiopathic Membranoproliferative Glomerulonephritis (MPGN). Ig-MPGN: immunoglobulin-associated MPGN; C3G: C3 glomerulopathy; DDD: Dense Deposit Disease; C3GN: C3 glomerulonephritis; IF immunofluorescence microscopy; EM: electron microscopy.

Complement component assays

Serum C3 and C4 levels were assessed by kinetic nephelometry.[294] To assess plasma SC5b-9 levels and C3NeF activity, blood was collected in ice-cold EDTA tubes, immediately centrifuged at 4° C to avoid ex-vivo complement activation, and frozen at -80°C until assay. SC5b-9 levels were measured using the MicroVue SC5b-9 Plus EIA commercial kit (SC5b-9 Plus, Quidel). Serum C3, C4 and plasma SC5b-9 were assessed at Mario Negri Institute. C3NeF activity was determined at the Cordelier Research Center by purifying IgG from plasma and assessing their ability to stabilize cell-bound C3bBb convertase as previously described.[298]

DNA Sequencing

Genomic DNA was extracted from peripheral blood using the Nucleon™ BACC2 Genomic DNA extraction kit (GE Healthcare, Little Chalfont, UK).

Through next-generation sequencing, genetic screening of all exons and flanking regions of *CFH* (NG_007259.1), *CD46* (NG_009296.1), *CFI* (NG_007569.1), *CFB* (NG_008191.1), *C3* (NG_009557.1) and *THBD* (NG_012027.1) genes was performed by highly multiplex PCR using the Ion AmpliSeq™ Library Kit 2.0 and 311 primer pairs (subdivided into two pools), designed using the Ion AmpliSeq Designer tool. After clonal amplification using the Ion PGM™ Template OT2 200 Kit, sequencing was performed on Ion PGM Sequencer. Variants were identified using TorrentSuite Software 3.6, and putative mutations were validated by Sanger sequencing. Regions with a depth of amplicon coverage ≥40x reads were considered appropriately analyzed. Regions with low coverage or regions not covered by the Ion AmpliSeq Designer were sequenced by Sanger method. In 9 patients, screening of CFH, CD46 and CFI coding

exons was performed by Sanger sequencing as previously described.[294]

Genetic variant annotation

Genetic variants were functionally annotated using the ANNOVAR software. Missense variants, insertions/deletions in the coding regions, splicing variants affecting the first two nucleotides flanking exons, or variants lying within the Kozak consensus sequence were considered functional. Variants that did not exceed an allelic frequency of 0.01 in none of the ESP or the 1000 Genomes Project subpopulations were considered rare. Likely pathogenic variants were defined as the variants with minor allele frequency in the ExAC database <0.001 and with Combined Annotation Dependent Depletion (CADD)[299] phred score ≥10.

The Combined Annotation Dependent Depletion is a method that integrates different annotation in a general framework to estimate the relative pathogenicity of human genetic variants.[299] It was implemented as a support vector machine trained to differentiate 14.7 million fixed or nearly fixed derived alleles in humans from 14.7 million simulated variants. The fixed or nearly fixed human-derived alleles are those alleles that differ from the inferred human-chimpanzee ancestral genome and have an allele frequency of at least 95% in the human population (according to the 1000 Genomes Project variant catalog). Deleterious variants that reduce organism fitness are depleted by natural selection in fixed but not simulated variation. The support vector machine uses 63 distinct annotations that span a range of data types including species conservation metrics like GERP, phastCons, and phyloP; regulatory information like genomic regions of DNase hypersensitivity and transcription factor binding; transcript information like distance to exon-intron boundaries or expression levels in commonly studied cell lines; and protein-level scores like Grantham, SIFT, and PolyPhen. CADD

objectively integrates the diverse annotations into a single, quantitative score (C-score). "C-scores" were pre-computed for all 8.6 billion possible human single nucleotide variants and enable scoring of short insertions/deletions. C-scores were also scaled to a Phred-like scale ranging from 1 to 99 on the basis of their ranking relative to all possible substitutions in the human reference genome [−10log10 (rank/total number of substitutions)]. A Phred-scaled C-score of greater or equal 10 indicates that these are predicted to be the 10% most deleterious substitutions that may occur to the human genome, a score of greater or equal 20 indicates the 1% most deleterious and so on. The authors state that the choice of a cutoff on deleteriousness to identify potentially pathogenic variants is always arbitrary and there is no natural choice. They suggest a cutoff of Phred-scaled C-score somewhere between 10 and 20. Since C3G and Ig-MPGN are considered complex diseases that may derive by adding up the small effect of multiple hits, the less stringent cutoff of 10 was adopted for these diseases. For steroid-resistant nephrotic syndrome, which is believed a monogenic disease, we adopted a more stringent cutoff of 15 (http://cadd.gs.washington.edu/).[299]

The ExAC database, containing 60,706 subjects, was adopted as reference to assess allelic frequencies. To perform association studies, we used as a control group the 286 unphenotyped subjects of European origin from the 1000 Genomes Project (CEU, TSI, GBR and IBS subpopulations), for whom individual genotypes were available.

Statistical analyses

Categorical data were analyzed with χ^2 test or Fisher exact test, where appropriate. One-way ANOVA, Kruskal-Wallis and multivariate regression tests were used for continuous variables.

Degree of mesangial proliferation, endocapillary proliferation, interstitial inflammation, interstitial fibrosis, and arteriolar sclerosis, as well as IF findings were graded using a scale of 0 to 3, including 0, trace, 1+, 2+ and 3+. These findings were considered as quantitative variables after substituting 'trace' with 0.5, and analyzed with one-way ANOVA. EM findings were considered as dichotomous variables and were analyzed with χ^2 test or Fisher exact test where appropriate.

Survival analyses considered as cumulative fractions of patients free of events were estimated using Kaplan-Meier survival curves and with Cox proportional-hazard regression. Multivariate analyses included all the covariates that reached statistical significance at the univariate analysis. To perform multivariate analyses with genetic variants, homozygous genotypes for the reference allele were codified as 0, heterozygous as 1 and homozygous for the alternative allele as 2. In the dominant model the group with at least one alternative allele was compared with the group with no alternative alleles. In the recessive model the group homozygous for the alternative allele was compared with the group with no or one alternative allele. P-values <0.05 were considered statistically significant. Analyses were performed with the MedCalc software and with the R software (https://cran.r-project.org/).

3.4. Results

3.4.1 Pathologic Features

Among 140 patients from 135 families studied, 52% had C3G (DDD: 21, C3GN: 52) and 48% Ig-MPGN (Figure 3.2), a distribution comparable with that of a previously reported large cohort of idiopathic MPGN.[300] Kidney biopsy was performed a median of 0.3 and 0.4 years after disease onset for Ig-MPGN and C3G, respectively (DDD: 0.6 years, C3GN: 0.4 years), without significant differences between histology groups (Table 3.1). Glomerular C3 staining was intense in all three histology groups. As expected, immunohistology showed stronger immunoglobulin and C1q staining in Ig-MPGN, while EM showed the presence of intramembranous highly electron-dense ribbon-like deposits in DDD and higher prevalence of subendothelial and subepithelial deposits in Ig-MPGN and C3GN. Interestingly, two patients with intramembranous highly electron-dense ribbon-like deposits by EM were classified as Ig-MPGN (C3 3+, IgA 0, IgG 3+, IgM 0, C1q 2+ and fibrinogen 0 in one patients, and C3 3+, IgA 1+, IgG 2+, IgM traces, C1q 1+ and fibrinogen 1+ in the other patients) confirming that some DDD cases would not be classified as C3G even with the criterion of dominant C3 (C3 of ≥2 orders of magnitude of intensity by IF greater than any other immune reactant).[253, 300] Additionally, we observed a higher degree of mesangial proliferation, endocapillary proliferation, and interstitial inflammation in Ig-MPGN compared with C3G (Table 3.1).

Table 3.1. Light microscopy, immunofluorescence and electron microscopy features of kidney biopsies in patients classified according to the recent Ig-MPGN and C3G classification.

	Ig-MPGN	C3G		
			DDD	C3GN
N	67	73	21	52
Time Onset to Biopsy (yr) - median (IQR)	0.3 (0.1-2.3)	0.4 (0.1-3.6)	0.6 (0.1-3.9)	0.4 (0.1-3.1)
Light microscopy				
N. of glomeruli	17.7 (±9.1)	17.3 (±9.8)	16.6 (±9.1)	17.6 (±10.2)
Sclerotic glomeruli	9% (±15%)	5% (±15%)	2% (±7%)	7% (±16%)
Crescents	6% (±18%)	4% (±14%)	8% (±22%)	2% (±1%)
Degree of mesangial proliferation*	1.98 (±0.95)[a]	1.61 (±0.95)[a]	1.82 (±0.78)	1.53 (±1.01)
Degree of endocapillary proliferation*	1.33 (±1.03)[b]	0.93 (±1.13)[b]	0.97 (±1.18)	0.92 (±1.12)
Degree of interstitial inflammation*	0.9 (±0.78)[c]	0.6 (±0.86)[c]	0.69 (±0.84)	0.57 (±0.87)
Degree of interstitial fibrosis*	0.6 (±0.85)	0.37 (±0.72)	0.17 (±0.38)	0.44 (±0.79)
Degree of arteriolar sclerosis*	0.27 (±0.63)	0.28 (±0.75)	0.03 (±0.12)	0.37 (±0.85)
Immunofluorescence				
C3*	2.61 (±0.59)	2.71 (±0.46)	2.87 (±0.23)	2.65 (±0.51)
IgA*	0.55 (±0.85)[d,e1,e2]	0.04 (±0.18)[d]	0.11 (±0.32)[e1]	0.02 (±0.1)[e2]
IgG*	1.7 (±1.16)[f,g1,g2]	0.18 (±0.4)[f]	0.13 (±0.28)[g1]	0.2 (±0.43)[g2]
IgM*	1.25 (±0.94)[h,i1,i2]	0.34 (±0.55)[h]	0.58 (±0.55)[i1]	0.22 (±0.42)[i2]
C1q*	1.16 (±1.04)[k,l1,l2]	0.15 (±0.33)[k]	0.16 (±0.34)[l1]	0.15 (±0.34)[l2]
Fibrinogen*	0.34 (±0.71)	0.38 (±0.81)	0.35 (±0.84)	0.39 (±0.81)
Electron microscopy				
Mesangial deposits	51%	63%	53%	68%
Subepithelial deposits	37%[m,n]	48%	10%[m,o]	63%[n,o]
Subendothelial deposits	81%[p]	53%	19%[p,q]	67%[q]
Intramembranous granular deposits	37%[r]	37%	0%[r,s]	52%[s]
Intramembranous highly electron-dense ribbon-like deposits	3%[t]	29%	100%[t,u]	0%[u]

[a]p=0.029; [b]p=0.041; [c]p=0.043; [d]p=2x10^{-6}; [e1]p=0.026; [e2]p=1x10^{-5}; [f]p=1x10^{-18}; [g1]p=2x10^{-7}; [g2] p=1x10^{-17}; [h]p=1x10^{-10}; [i1]p=0.013; [i2]p=4x10^{-11}; [k]p=3x10^{-11}; [l1]p=1x10^{-4}; [l2]p=1x10^{-9}; [m]p=0.016; [n]p=0.005; [o]p=3x10^{-5}; [p]p=2x10^{-7}; [q]p=0.0002; [r]p=0.001; [s]p=3x10^{-5}; [t]p<2x10^{-16}; [u]p<2x10^{-16}

*Degree of mesangial proliferation, endocapillary proliferation, interstitial inflammation, interstitial fibrosis, and arteriolar sclerosis, as well as IF findings were graded using a scale of 0 to 3, including 0, trace, 1+, 2+ and 3+.

Table 3.2. Clinical and laboratory findings in different histology groups**.**

	Ig-MPGN	C3G		
			DDD	C3GN
N	67	73	21	52
Gender (% males)	51%	41%	38%	42%
Data at onset				
Age (y) - Mean (SD)	20.1 (±15.2)[a]	15.0 (±11.9)[a]	15.9 (±11.4)	14.6 (±12.1)
Microhematuria	86%	86%	90%	84%
Gross hematuria	35%	38%	38%	38%
Proteinuria	88%	90%	86%	92%
Nephrotic syndrome	43%	29%	29%	29%
Renal impairment	27%[b]	14%	5%[b]	17%
Trigger event	31%	42%	41%	42%
C3NeFs positive	44%[e]	54%	78%[e,f]	44%[f]
Serum C3 (mg/dl) - Median (IQR)	45 (13-77)[c]	38 (15-73)	20 (9-46)[c,d]	44 (17-87)[d]
Serum C4 (mg/dl) - Median (IQR)	21 (12-29)	21 (17-27)	26 (20-29)	21 (17-25)
Low Serum C3 & normal C4	67%	74%	86%	69%
Plasma SC5b-9 (ng/ml) - Median (IQR)	515 (286-1860)	417 (228-1033)	353 (252-623)	486 (215-1140)
Mutation carriers*	17%	18%	14%	20%
Mutation carriers and/or C3NeFs*	56%	65%	83%	58%
Familiarity for nephropathy*	11%	14%	10%	16%

[a]p=0.028; [b]p=0.035; [c]p=0.047; [d]p=0.028; [e]p=0.012; [f]p=0.016.

*calculated in unrelated patients (64 Ig-MPGN, 21 DDD and 50 C3GN).

Nephrotic syndrome was defined as: 24-hour proteinuria exceeding 3.5g in adults or 40mg/h/m2 in children together with albuminemia ≤3 g/L. Renal impairment was defined as abnormal serum creatinine levels. Familiarity for nephropathy was defined as the presence of at least one relative (up to 3rd degree) with biopsy-proven Ig-MPGN/C3G, or proteinuria and/or renal impairment without other apparent cause. Quantitative variables are expressed as mean (±S.D.) unless otherwise specified.

Serum C3: reference 90-180 mg/dl; serum C4: reference 10-40 mg/dl; plasma SC5b-9: reference ≤303 ng/ml.

3.4.2 Genetic screening

Likely pathogenic variants, defined as genetic variants with Minor Allelic Frequency (MAF) ≤0.001, and Combined Annotation Dependent Depletion (CADD) pathogenic score ≥10, were detected in 24/135 (18%) unrelated patients, and their prevalence was comparable between Ig-MPGN and C3G (17 and 18%) and did not differ between the three histology groups (Ig-MPGN, DDD and C3GN, Table 3.2).

We compared the cumulative allelic frequency of the likely pathogenic variants between our cohort and 286 ethnically matched controls from the 1000 Genome Project. The patient group is significantly enriched in likely pathogenic variants in all six genes (p=1.9x10^{-7}) compared to controls, and significant differences were also observed between *CFB* (p=0.032) and *C3* (p=1.8x10^{-4}) variants separately (Table 3.3). The screening of the available affected family members of three unrelated patients showed segregation of the variants with the disease, further supporting their involvement in disease pathogenesis.

Table 3.3. Cumulative frequency of patients and controls carrying variants with MAF <0.001 and CADD pathogenic score ≥10.

	MAF≤0.001 & CADD≥10		
	Pt	**Ctr**	***p***
All	0.18	0.03	**1.9x10^{-07}**
CFH	0.04	0.01	0.118*
CD46	0.01	0.00	0.321*
CFI	0.01	0.00	0.242*
CFB	0.02	0.00	**0.032***
C3	0.09	0.01	**1.8x10^{-04}**
THBD	0.01	0.00	0.539*

*Fisher exact test. For this analysis we considered as controls the 286 European subjects from the 1000 Genome Project for whom the individual genotypes were available. Ig-MPGN, n=64; DDD, n=21; C3GN, n=50.

Of the likely pathogenic variants (herein defined as mutations), 5 have already been described in patients with Ig-MPGN, C3G or aHUS, while 19 are novel (Figure 3.3 and Table 3.4).[294, 295, 301-306] We identified seven additional variants that have previously been reported in Ig-MPGN, C3G or aHUS patients, but we did not include these variants among the mutations, because they were relatively common in the control population or because they were not predicted as damaging by CADD (Table 3.4).

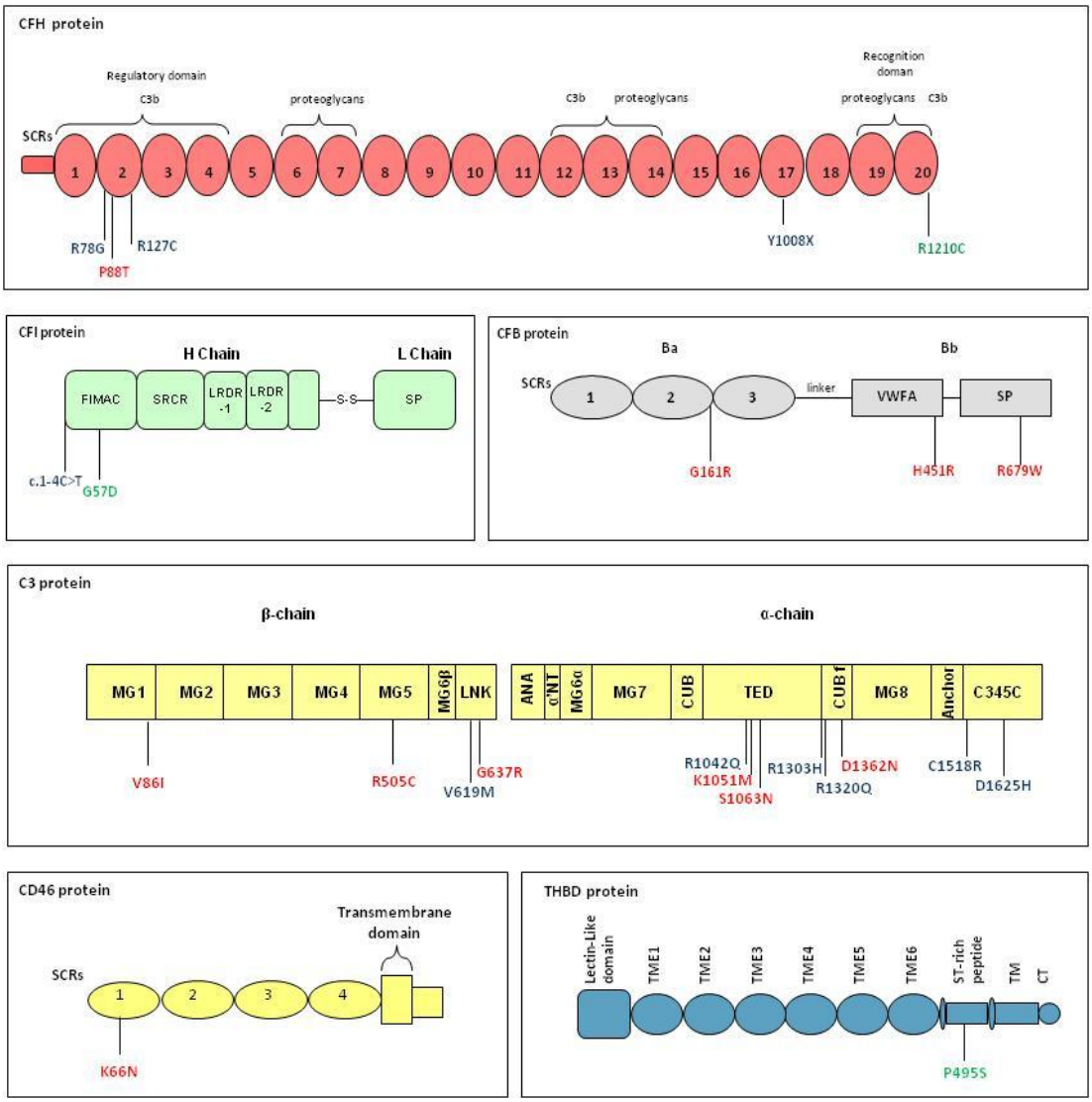


Figure 3.3 Graphical representation of CFH, CD46, CFI, CFB, C3, THBD genes and location of the mutations identified in Ig-MPGN (red), DDD (green) and C3GN (blue) patients.

Table 3.4. Likely pathogenic variants and other variants previously reported in Ig-MPGN, C3G or aHUS patients.

Mutation	Zygo-sity	Histology group	ExAC Global Frequency	ExAC NFE Frequency	CADD	New/ known
Likely pathogenic variants						
CFH R78G	Hom	C3GN	0	0	10	(306-308)
CFH P88T	Hom	Ig-MPGN	0	0	19	New
CFH R127C	Het	C3GN	0	0	18	New
CFH Y1008X	Hom	C3GN	0	0	37	New
CFH R1210C	Het	DDD	0.0002	0.0003	16	(256, 294, 301)
CD46 K66N	Het	Ig-MPGN	0.0005	0.0009	13	New
CFI c.1-4C>T	Het	C3GN	8.2E-06	1.5E-05	18	New
CFI G57D	Het	DDD	8.2E-06	0	16	New
CFB G161R	Het	Ig-MPGN	6.8E-05	7.9E-05	18	New
CFB H451R	Het	Ig-MPGN	0	0	24	New
CFB R679W	Hom	Ig-MPGN	0	0	17	New
C3 V86I	Het	Ig-MPGN	0	0	12	New
C3 R505C	Het	Ig-MPGN	8.2E-06	1.5E-05	16	New
C3 V619M	Het	C3GN	0.0003	0.0004	14	New
C3 G637R	Het	Ig-MPGN	0.0002	0.0004	11	New
C3 R1042Q*	Het	C3GN	0	0	21	(301)
C3 K1051M	Het	Ig-MPGN	1.6E-05	3.0E-05	15	(304)
C3 S1063N	Het	Ig-MPGN	9.1E-05	0	12	New
C3 R1303H	Het	C3GN	8.2E-06	8.2E-06	17	New
C3 R1320Q	Het	C3GN	0	0	19	New
C3 D1362N	Het	Ig-MPGN	4.9E-05	6.0E-05	10	New
C3 C1518R	Het	C3GN	0	0	15	New
C3 D1625H	Het	C3GN	0	0	10	New
THBD P495S	Het	DDD	0.0004	0.0010	12	(309)
Variants previously reported in Ig-MPGN, C3G or aHUS patients						
CFH I216T	Het	Ig-MPGN	8.2E-05	0	6	(293)
CFH Q950H	Het	Ig-MPGN	0.0036	0.0061	13	(307)
CFH T956M	Het	C3GN	0.0012	0.0013	17	(302, 303)
CFI R317W*	Het	C3GN	9.9E-05	9.0E-05	8	(310)
CFB E566A[§]	Het	5 patients[$]	0.011	0.010	0	(303, 311)
C3 K633R[§]	Het	C3GN	0.0004	0.0007	2	(306)
THBD D486Y	Het	C3GN	0.0065	0.0004	12	(309)

*One patient carried both *C3* R1042Q and *CFI* R317W; [$]found in 3 I-MPGN, 1 DDD and 1 C3GN patients, 5/135 (0.04) patients with *CFB* E566A *vs.* 5/286 (0.02) controls, p=0.302; [§]one patient carried both *C3* K633R and *CFB* E566A variant.

ExAC Global Frequency: variant frequency in all subjects of the ExAC database; ExAC NFE Frequency: variant frequency in non-Finnish European subjects of the ExAC database. In brackets the reference of the variants already described in MPGN, C3G or aHUS patients.

Mutations were homozygous in four patients, all with parental consanguinity, and heterozygous in twenty. Three out of the 4 homozygous mutations affected *CFH* and the fourth *CFB*. The highest frequency of genetic abnormalities was found in *C3* (12 patients, 50%), followed by *CFH* (5 patients, 21%), *CFB* (3 patients, 13%), *CFI* (2 patients, 8%), *CD46* (1 patient, 4%) and *THBD* (1 patient, 4%).

C3 mutations were found only in Ig-MPGN and C3GN but not in DDD (Figure 3.3 and Table 3.4). The mutations affected domains involved in C3 convertase formation, such as the C345C domain, which binds FB, and the MG5 domain, which participates in the C3bBb-C3 interaction, or in domains (MG1, TED and CUBf) involved in the binding of C3 with the complement regulators FH and FI.[312-314] Mutations of *CFB*, the counterpart of C3 in the formation of the alternative pathway C3 convertase, were found in three Ig-MPGN patients and occurred in functional domains that are involved in C3-FB interactions (Figure 3.3).[315]

Mutations in complement regulatory genes were found both in Ig-MPGN and C3G, regardless histologic pattern (Figure 3.3 and Table 3.4). All *CFH* mutations affected domains involved in the C3-FH interaction. Three of the 5 mutations occurred in the N-terminal short consensus repeat (SCR) 2, at variance with aHUS-associated mutations that mainly clustered in SCRs 18-20.[294] The *CFI* G57D is in the FIMAC domain, critical for FI-mediated C3b degradation,[316] while the c.-4C>T in the Kozak sequence should affect protein translation (Figure 3.3). The *CD46* K66N is located within SCR1, whose function is unclear.[317] The *THBD* P495S has previously been reported in patients with aHUS and results in less effective C3b inactivation.[309]

3.4.3 Susceptibility genetic variants

We then investigated whether common genetic variants causing amino-acid change, and non-coding *CFH* and *CD46* variants previously associated with MPGN, promote susceptibility to Ig-MPGN or C3G (Table 3.5).[256] We observed a higher frequency of the c.-366A (OR=1.7, 95%CI 1.1-2.6, *p*=0.015) and the c.*783T (OR=1.6, 95%CI 1.1-2.5, *p*=0.023) *CD46* alleles in Ig-MPGN patients compared with 286 ethnically matched controls from the 1000 Genomes Project. Fifty-nine out of 64 Ig-MPGN patients (92%) carried at least one CD46 c.-366A allele. We also observed a higher frequency of the CFB R32 allele in Ig-MPGN patients compared to controls (OR=1.8, 95%CI 1.1-2.8, *p*=0.020) and significance fitted with a recessive inheritance mode (OR=1.8, 95%CI 1.1-3.1, *p*=0.031). Considering a recessive mode of inheritance, there was a trend of association of *CD46* c.-366A with C3G, which was significant when we considered DDD alone (OR=3.3, 95%CI 1.3-8.4, *p*=0.009). C3G patients had a higher frequency of the *CFH* V62 (OR=1.7, 95%CI 1.04-2.7, *p*=0.032) and *THBD* A473 (OR=2.1, 95%CI 1.2-3.7, *p*=0.012) alleles compared with controls and significance fitted with a recessive inheritance mode. The *CFH* V62 also significantly conferred disease risk (recessive model) in DDD (OR=3.3, 95%CI 1.1-10.1, *p*=0.027). Among the 71 C3G patients, 41 (58%) carried both the *CFH* V62 and *THBD* A473 alleles in homozygosity; 24 (34%) were homozygous for one susceptibility allele (*CFH* V62 n=9, *THBD* A473 n=15); only 6 (8%) patients carried no homozygous susceptibility variants.

Complement gene mutations were associated with a risk of Ig-MPGN (OR=6.4, 95%CI 2.5-16.2, *p*=0.0001) and C3G (OR=6.9, 95%CI 2.8-16.9, $p=3x10^{-5}$). We then evaluated whether mutations and the susceptibility genetic variants, namely homozygous or heterozygous *CD46* c.-366A allele for Ig-MPGN and homozygous *CFH* V62 and *THBD* A473 for C3G, synergized to increase disease risk (Figure 3.4). For Ig-MPGN, mutations combined with the *CD46* c.-366A susceptibility variant increased disease risk

(OR=19.2, 95%CI 4.6-80.8, compared with the reference group with subjects lacking both mutations and *CD46* c.-366A). For C3G, disease risk was highest in subjects carrying mutations and both susceptibility variants (OR=23.9, 95%CI 4.0-143.2 compared with. subjects lacking both mutations and susceptibility variants). The disease risk was intermediate (OR=8.2, 95%CI 1.9-35.4) for subjects with mutations and only 1 susceptibility variant.

Table 3.5. Association study between common polymorphic variants and histology groups.

SNP-ID	Risk Allele	Group	Alleles				Dominant model (homozygous or heterozygous for the risk allele)				Recessive model (homozygous for the risk allele)			
			Freq	OR (95%CI)	*p* (*vs.* Ctr)	*p* (Ig-MPGN *vs.* C3G)	Freq	OR (95%CI)	*p* (*vs.* Ctr)	*p* (Ig-MPGN *vs.* C3G)	Freq	OR (95%CI)	*p* (*vs.* Ctr)	*p* (Ig-MPGN *vs.* C3G)
***CFH* c.-331C>T (rs3753394)**	T	**Ig-MPGN**	0.36	1.36 (0.91-2.04)	0.134	0.389	0.56	1.32 (0.77-2.28)	0.315	0.750	0.16	1.85 (0.84-4.06)	0.120	0.198
		C3G	0.31	1.09 (0.73-1.62)	0.676		0.54	1.18 (0.7-1.99)	0.52		0.08	0.92 (0.36-2.34)	0.87	
		DDD	0.26	0.86 (0.42-1.75)	0.679		0.43	0.77 (0.32-1.89)	0.57		0.10	1.05 (0.23-4.77)	1.00*	
		C3GN	0.33	1.19 (0.76-1.88)	0.443		0.58	1.42 (0.77-2.61)	0.26		0.08	0.87 (0.29-2.61)	1.00*	
		Ctr	0.29				0.49				0.09			
***CFH* V62I (rs800292)**	V	**Ig-MPGN**	0.80	1.35 (0.84-2.2)	0.218	0.474	0.95	1.20 (0.33-6.65)	1.00*	0.668*	0.66	1.48 (0.84-2.61)	0.172	0.550
		C3G	0.84	1.69 (1.04-2.7)	**0.032**		0.97	2.04 (0.46-18.7)	0.544*		0.70	1.85 (1.06-3.24)	**0.030**	
		DDD	0.88	2.42 (0.93-6.3)	0.061		0.95	1.18 (0.17-52.2)	1.00*		0.81	3.3 (1.08-10.1)	**0.027**	
		C3GN	0.82	1.49 (0.86-2.6)	0.149		0.98	2.90 (0.43-124.2)	0.485*		0.66	1.51 (0.8-2.83)	0.200	
		Ctr	0.75				0.94				0.56			
***CFH* Y402H (rs1061170)**	H	**Ig-MPGN**	0.35	0.93 (0.63-1.4)	0.741	0.337	0.55	0.89 (0.51-1.53)	0.661	0.490	0.16	0.99 (0.47-2.09)	0.983	0.411
		C3G	0.41	1.19 (0.82-1.73)	0.363		0.61	1.13 (0.66-1.91)	0.661		0.21	1.43 (0.75-2.76)	0.277	
		DDD	0.45	1.42 (0.76-2.68)	0.270		0.57	0.98 (0.40-2.39)	0.961		0.33	2.68 (1.02-7.00)	0.063*	
		C3GN	0.39	1.1 (0.71-1.7)	0.662		0.62	1.20 (0.65-2.22)	0.569		0.16	1.02 (0.45-2.32)	0.962	
		Ctr	0.37				0.58				0.16			
***CFH* E936D (rs1065489)**	D	**Ig-MPGN**	0.21	1.18 (0.73-1.89)	0.504	0.995	0.31	0.94 (0.53-1.69)	0.845	0.321	0.11	2.58 (0.99-6.75)	0.068*	0.084*
		C3G	0.21	1.18 (0.75-1.86)	0.481		0.39	1.35 (0.79-2.31)	0.270		0.03	0.61 (0.13-2.76)	0.744*	
		DDD	0.14	0.73 (0.30-1.78)	0.492		0.29	0.83 (0.31-2.21)	0.709		0.00	nc	1.00*	
		C3GN	0.24	1.39 (0.84-2.30)	0.202		0.44	1.63 (0.89-3.00)	0.114		0.04	0.88 (0.19-4.00)	1.00*	
		Ctr	0.19				0.33				0.05			

Ig-MPGN, n=64; DDD, n=21; C3GN, n=50, controls, n=286. *p*: *p*-value assessed by chi-square test or with Fisher Exact test (*) where appropriate (*p*-values in bold are statistically significant); Freq: frequency; OR: Odds-ratio with 95% confidence interval in brackets; nc: not calculable. SNP: Simple Nucleotide Polymorphism. *CD46* c.-366A>G and c.*783T>C are in strong linkage disequilibrium (r^2=0.98).

Table 3.5. *continue*

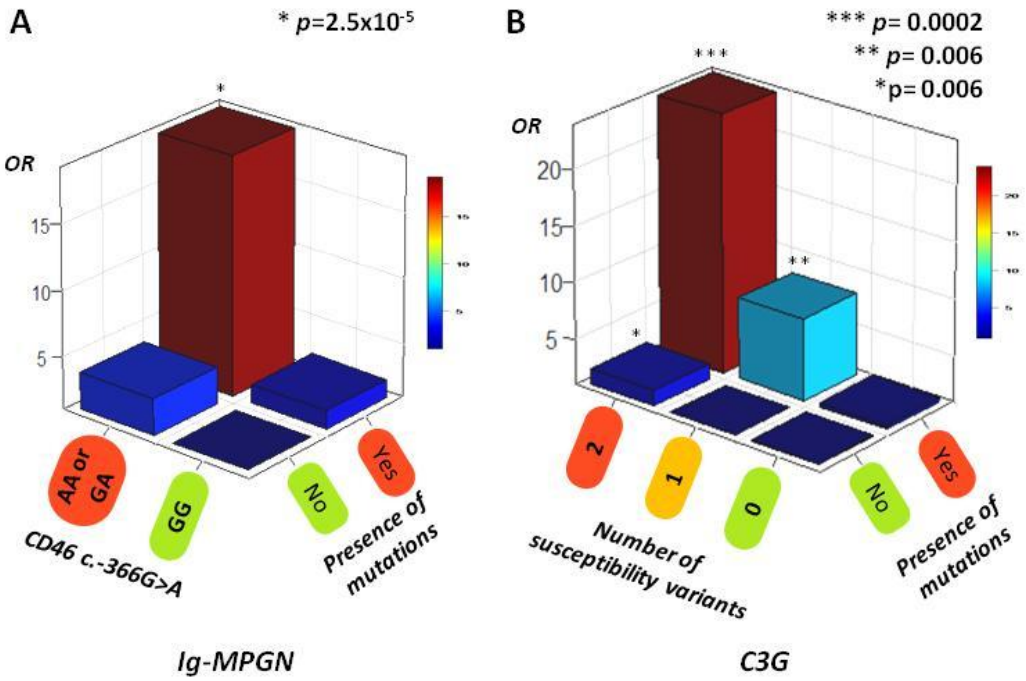
SNP-ID	Risk Allele	Group	Alleles				Dominant model (homozygous or heterozygous for the risk allele)				Recessive model (homozygous for the risk allele)			
			Freq	OR (95%CI)	*p* (*vs.* Ctr)	*p* (Ig-MPGN *vs.* C3G)	Freq	OR (95%CI)	*p* (*vs.* Ctr)	*p* (Ig-MPGN *vs.* C3G)	Freq	OR (95%CI)	*p* (*vs.* Ctr)	*p* (Ig-MPGN *vs.* C3G)
***CD46* c.-652G>A (rs2796267)**	A	**Ig-MPGN**	0.70	1.45 (0.95-2.21)	0.087	0.607	0.90	1.95 (0.80-4.79)	0.137	0.448	0.50	1.42 (0.81-2.49)	0.213	0.859
		C3G	0.67	1.26 (0.83-1.89)	0.276		0.85	1.28 (0.59-2.76)	0.531		0.48	1.33 (0.77-2.32)	0.303	
		DDD	0.74	1.72 (0.79-3.76)	0.167		0.88	1.63 (0.36-7.34)	0.746*		0.59	2.03 (0.75-5.50)	0.154	
		C3GN	0.64	1.12 (0.71-1.79)	0.620		0.84	1.18 (0.50-2.79)	0.709		0.44	1.14 (0.60-2.15)	0.687	
		Ctr	0.62				0.82				0.41			
***CD46* c.-366A>G (rs2796268)**	A	**Ig-MPGN**	0.72	1.68 (1.1-2.56)	**0.015**	0.555	0.92	2.44 (0.93-6.39)	0.062	0.235	0.52	1.75 (0.99-3.03)	0.056	0.988
		C3G	0.69	1.44 (0.97-2.13)	0.071		0.86	1.24 (0.59-2.59)	0.566		0.51	1.75 (1.00-2.95)	0.051	
		DDD	0.76	2.11 (1.02-4.37)	0.041		0.86	1.24 (0.35-4.37)	1.00*		0.67	3.3 (1.29-8.42)	**0.009**	
		C3GN	0.65	1.24 (0.79-1.94)	0.349		0.86	1.24 (0.53-2.92)	0.622		0.45	1.34 (0.73-2.48)	0.344	
		Ctr	0.60				0.83				0.38			
***CD46* A309V (rs35366573)**	V	**Ig-MPGN**	0.02	1.19 (0.12-6.09)	0.687*	0.229*	0.03	1.19 (0.12-6.09)	0.658*	0.228	0.00	nc		
		C3G	0.00	nc	0.363*		0.00	nc	0.357*		0.00	nc		
		DDD	0.00	nc	1.00*		0.00	nc	1.00*		0.00	nc		
		C3GN	0.00	nc	0.607*		0.00	nc	0.6*		0.00	nc		
		Ctr	0.01				0.02				0.00			
***CD46* c.*783T>C (rs7144)**	T	**Ig-MPGN**	0.71	1.62 (1.07-2.46)	**0.023**	0.653	0.92	2.44 (0.93-6.39)	0.062	0.235	0.50	1.65 (0.96-2.84)	0.071	0.869
		C3G	0.69	1.44 (0.97-2.13)	0.071		0.86	1.24 (0.59-2.59)	0.566		0.51	1.75 (0.99-2.95)	0.051	
		DDD	0.74	1.85 (0.91-3.76)	0.083		0.86	1.24 (0.34-6.82)	1.00*		0.62	2.68 (1.08-6.67)	**0.029**	
		C3GN	0.66	1.3 (0.83-2.03)	0.259		0.86	1.24 (0.53-2.92)	0.622		0.47	1.46 (0.79-2.68)	0.224	
		Ctr	0.60				0.83				0.38			

Table 3.5. *continue*

SNP-ID	Risk Allele	Group	Alleles				Dominant model (homozygous or heterozygous for the risk allele)				Recessive model (homozygous for the risk allele)			
			Freq	OR (95%CI)	*p* (*vs.* Ctr)	*p* (Ig-MPGN *vs.* C3G)	Freq	OR (95%CI)	*p* (*vs.* Ctr)	*p* (Ig-MPGN *vs.* C3G)	Freq	OR (95%CI)	*p* (*vs.* Ctr)	*p* (Ig-MPGN *vs.* C3G)
***CFB* L9H (rs4151667)**	H	**Ig-MPGN**	0.03	0.59 (0.20-1.71)	0.328	1.00*	0.06	0.62 (0.21-1.85)	0.392	1.00*	0.00	nc	1.00*	1.00*
		C3G	0.03	0.55 (0.19-1.58)	0.259		0.06	0.58 (0.20-1.70)	0.312		0.00	nc	1.00*	
		DDD	0.02	0.44 (0.01-2.8)	0.714*		0.05	0.46 (0.06-3.56)	0.705*		0.00	nc	1.00*	
		C3GN	0.03	0.60 (0.11-1.98)	0.607*		0.06	0.63 (0.18-2.16)	0.594*		0.00	nc	1.00*	
		Ctr	0.05				0.10				0.01			
***CFB* R32W/Q (rs12614 & rs641153)**	R	**Ig-MPGN**	0.86	1.75 (1.09-2.81)	**0.020**	0.104	0.99	4.10 (0.53-31.8)	0.203*	0.089*	0.74	1.79 (1.05-3.05)	**0.031**	0.286
		C3G	0.79	1.07 (0.68-1.69)	0.769		0.93	0.62(0.21-1.80)	0.366*		0.66	1.23 (0.71-2.13)	0.452	
		DDD	0.74	0.79 (0.36-1.61)	0.514		0.86	0.29 (0.07-1.09)	0.087*		0.62	1.05 (0.42-2.60)	0.923	
		C3GN	0.82	1.24 (0.72-2.15)	0.436		0.96	1.12 (0.25-5.12)	1.00*		0.67	1.33 (0.70-2.52)	0.386	
***CFB* G252S (rs4151651)**	S	**Ig-MPGN**	0.02	0.56 (0.06-2.43)	0.756*	0.609*	0.03	0.56 (0.06-2.43)	0.753*	0.607*	0.00	nc		
		C3G	0.01	0.26 (0.01-1.67)	0.221*		0.02	0.26 (0.01-1.67)	0.215*		0.00	nc		
		DDD	0.03	0.84 (0.02-5.63)	1.00*		0.05	0.84 (0.02-5.63)	1.00*		0.00	nc		
		C3GN	0.00	nc	0.148*		0.00	nc	0.143*		0.00	nc		
		Ctr	0.03				0.06				0.00	nc		
***CFB* K565E (rs4151659)**	E	**Ig-MPGN**	0.01	0.67 (0.01-5.27)	1.00*	1.00*	0.02	0.67 (0.01-5.27)	1.00*	1.00*	0.00	nc		
		C3G	0.01	0.63 (0.01-4.94)	1.00*		0.02	0.63 (0.01-4.94)	1.00*		0.00	nc		
		DDD	0.00	nc	1.00*		0.00	nc	1.00*		0.00	nc		
		C3GN	0.01	0.91 (0.02-7.2)	1.00*		0.02	0.91 (0.02-7.2)	1.00*		0.00	nc		
		Ctr	0.01				0.02				0.00	nc		

Table 3.5. *continue*

SNP-ID	Risk Allele	Group	Alleles				Dominant model (homozygous or heterozygous for the risk allele)				Recessive model (homozygous for the risk allele)			
			Freq	OR (95%CI)	p (*vs.* Ctr)	p (Ig-MPGN *vs.* C3G)	Freq	OR (95%CI)	p (*vs.* Ctr)	p (Ig-MPGN *vs.* C3G)	Freq	OR (95%CI)	p (*vs.* Ctr)	p (Ig-MPGN *vs.* C3G)
***CFB* E566A (rs45484591)**	A	**Ig-MPGN**	0.02	2.72 (0.42-14.17)	0.166*	0.670*	0.05	2.72 (0.42-14.17)	0.164*	0.668*	0.00	nc		
		C3G	0.01	1.62 (0.15-10.01)	0.631*		0.03	1.62 (0.15-10.01)	0.630*		0.00	nc		
		DDD	0.02	2.76 (0.06-25.48)	0.347*		0.05	2.76 (0.06-25.48)	0.349*		0.00	nc		
		C3GN	0.01	1.15 (0.02-10.39)	1.00*		0.02	1.15 (0.02-10.39)	1.00*		0.00	nc		
		Ctr					0.02				0.00			
***C3* P314L (rs1047286)**	L	**Ig-MPGN**	0.16	0.76 (0.45-1.26)	0.279	0.642	0.27	0.60 (0.33-1.09)	0.091	0.426	0.06	1.84 (0.41-6.64)	0.297*	0.709*
		C3G	0.19	0.88 (0.55-1.41)	0.587		0.33	0.81 (0.46-1.40)	0.446		0.04	1.24 (0.21-4.97)	0.725*	
		DDD	0.21	1.05 (0.49-2.25)	0.902		0.33	0.82 (0.32-2.11)	0.686		0.10	2.89 (0.29-15.07)	0.194*	
		C3GN	0.17	0.81 (0.46-1.41)	0.454		0.33	0.80 (0.42-1.52)	0.494		0.02	0.58 (0.01-4.22)	1.00*	
		Ctr	0.21				0.38				0.03			
***C3* R102G (rs2230199)**	G	**Ig-MPGN**	0.21	0.94 (0.59-1.5)	0.784	0.947	0.31	0.71 (0.4-1.26)	0.238	0.473	0.11	2.21 (0.73-6.09)	0.148*	0.271
		C3G	0.21	0.96 (0.61-1.5)	0.843		0.37	0.92 (0.54-1.58)	0.756		0.06	1.09 (0.26-3.59)	0.774*	
		DDD	0.26	1.24 (0.61-2.54)	0.550		0.38	0.96 (0.38-2.38)	0.923		0.14	2.99 (0.51-12.1)	0.115*	
		C3GN	0.19	0.84 (0.49-1.44)	0.533		0.37	0.90 (0.48-1.69)	0.747		0.02	0.38 (0.01-2.56)	0.484*	
		Ctr	0.22				0.39				0.05			
***THBD* A473V (rs1042579)**	A	**Ig-MPGN**	0.84	1.31 (0.78-2.21)	0.301	0.216	0.95	0.81 (0.21-4.68)	0.727*	0.104*	0.73	1.51 (0.82-2.77)	0.181	0.458
		C3G	0.89	2.06 (1.16-3.66)	**0.012**		1.00	nc	0.131*		0.79	2.04 (1.10-3.79)	**0.022**	
		DDD	0.93	3.17 (0.96-10.4)	0.066		1.00	nc	1.00*		0.86	3.28 (0.94-11.4)	0.079	
		C3GN	0.88	1.79 (0.94-3.38)	0.071		1.00	nc	0.380*		0.76	1.73 (0.86-3.46)	0.118	
		Ctr	0.80				0.96				0.65			



Ig-MPGN

Mutation/s	SVs	Pt	Ctr
Yes	1	15.6%	2.1%
Yes	0	1.6%	1.0%
No	1	76.6%	80.8%
No	0	6.3%	16.1%

C3G

Mutation/s	SVs	Pt	Ctr
Yes	2	9.9%	0.7%
Yes	1	8.5%	1.7%
Yes	0	0.0%	0.7%
No	2	47.9%	35.3%
No	1	25.4%	47.2%
No	0	8.5%	14.3%

Figure 3.4. Panel A. Graphical representation of the risk (expressed as odds-ratio, OR) of developing Ig-MPGN obtained through the combination of the mutations with the *CD46* c.-366A susceptibility variant. The group lacking both mutations and CD46 c.-366A was taken as reference. **Panel B.** Risk (expressed as OR) of developing C3G obtained through the combination of the mutations with the homozygous *CFH* V62 and homozygous *THBD* A473 (susceptibility variants). The group lacking both mutations and susceptibility variants was taken as reference. Histogram colors are scaled. Below each graph, the percentages of patients (Pt) or controls (Ctr) carrying mutations and/or susceptibility variants (SV) are reported.

3.4.4 Complement component assessment

Serum C3NeFs, C3 and C4 and plasma SC5b-9 at first investigation are reported in Table 3.2. There were no differences in the interval between onset and first investigation between histology groups. C3NeFs were positive in 57/116 tested patients (49%) and their prevalence was comparable between Ig-MPGN and C3G. However, when we considered the DDD and C3GN subgroups separately, the prevalence of C3NeFs in DDD (78%) was significantly higher compared with Ig-MPGN (44%, p=0.012) and C3GN (44%, p=0.016). Similarly, C3 levels were lower in DDD compared with Ig-MPGN (20 *vs.* 45mg/dl, p=0.047) and C3GN (20 *vs.* 44mg/dl, p=0.028). However, the prevalence of subjects with low C3 and normal C4, and plasma SC5b-9 levels did not differ between Ig-MPGN and C3G and among the 3 histology groups.

We then investigated whether genetic or acquired complement alternative pathway abnormalities were associated with a specific complement profile. Ig-MPGN and C3G patients carrying complement gene mutations, C3NeFs, or both, presented lower C3 and higher SC5b-9 levels (Figure 3.5) than patients without mutations or C3NeFs. However, in both groups intense C3 deposits in the kidney biopsy were observed (mutation and/or C3NeFs 2.8±0.4; no mutation or C3NeFs 2.5±0.7). Additionally, in C3G patients the homozygous *THBD* A473 allele was associated with lower C3 (p=0.026) and higher SC5b-9 levels (p=0.005) (Figure 3.6).

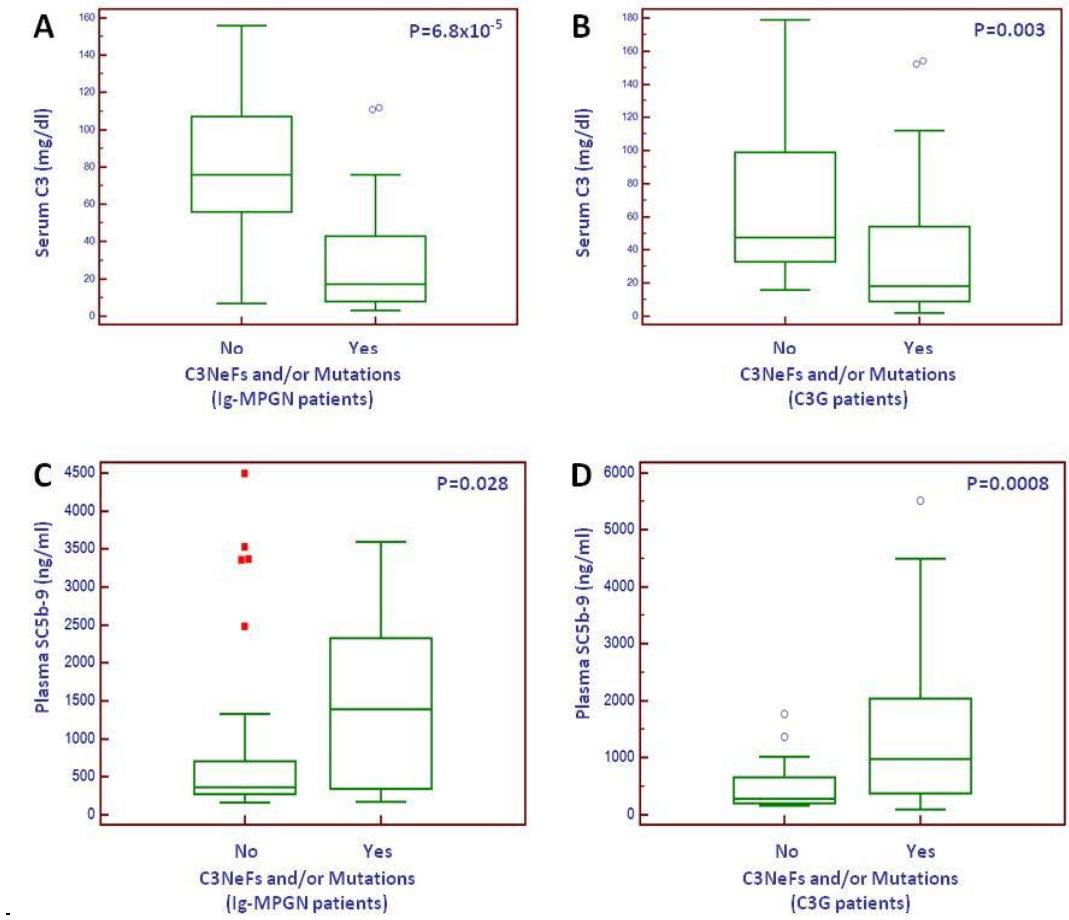


Figure 3.5. A) Box plot shows that Ig-MPGN patients with mutations and/or C3NeFs have lower serum C3 levels than Ig-MPGN patients without mutations or C3NeFs (median 15 *vs*. 66, respectively). **B)** Box plot shows that C3G patients with mutations and/or C3NeFs have lower serum C3 levels than C3G patients without (median 18 *vs*. 60, respectively). **C)** Box plot shows that Ig-MPGN patients with mutations and/or C3NeFs have higher plasma SCb-9 levels than Ig-MPGN patients without (median 1391 *vs*. 857, respectively). **D)** Box plot shows that C3G patients with mutations and/or C3NeFs have higher plasma SCb-9 levels than C3G patients without (median 983 *vs*. 463, respectively).

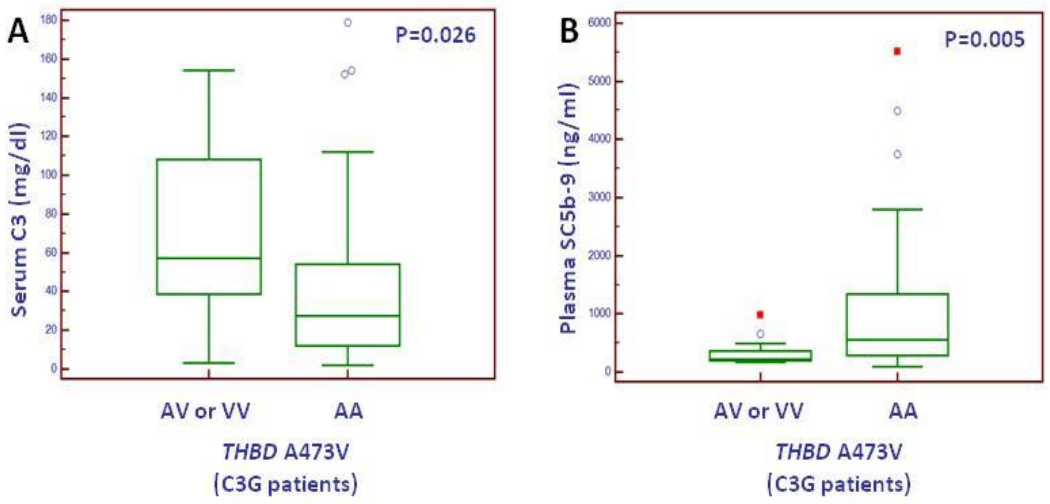


Figure 3.6. A) Box plot shows that C3G patients homozygous for the *THBD* A473 risk allele (AA) have lower serum C3 levels than those heterozygous (AV) or homozygous for the 473V allele (VV) (median 27 *vs.* 62, respectively, p=0.026). B) Box plot shows that C3G patients homozygous for the *THBD* A473 risk allele (AA) have higher plasma SC5b-9 levels than heterozygous (AV) or homozygous for the 473V allele (VV) patients (median 561 *vs.* 347, respectively, p=0.005).

3.4.5 Clinical parameters and predictors of long-term outcome

The mean age of onset was higher in Ig-MPGN than in C3G patients (20.1 *vs.* 15.0 years, *p*=0.028) (Table 3.2). The main clinical features at onset were proteinuria and microhematuria in all histology groups. Nephrotic syndrome at onset was present in 43%, 29% and 29% of patients with Ig-MPGN, DDD and C3GN, respectively. Renal impairment at onset was observed in 27%, 5% and 14% of Ig-MPGN, DDD and C3GN patients, respectively, and the difference was statistically significant between Ig-MPGN and DDD (Table 3.2).

The median follow-up was 4.8 years (interquartile range, IQR: 1.8-9.5). Sixteen patients developed ESRD (time from onset median, IQR: 5.8 years, 1.7-10.6). Kaplan-Meier survival analysis showed no differences in cumulative renal survival between

histologic groups (Figure 3.7). In order to evaluate whether the absence of correlation between renal survival and histology was influenced by clinical, therapeutic, biochemical or genetic confounders, we performed multivariate Cox proportional-hazards regression. The results, shown in Tables 3.6 and 3.7, confirmed the absence of any significant difference in renal survival between histology groups. However, other factors independently predicted progression to ESRD. Indeed, patients without mutations or C3NeFs had a higher risk of progression to ESRD, while a higher proportion of sclerotic glomeruli or of crescents in the kidney biopsy, and nephrotic syndrome at onset, predisposed to ESRD.

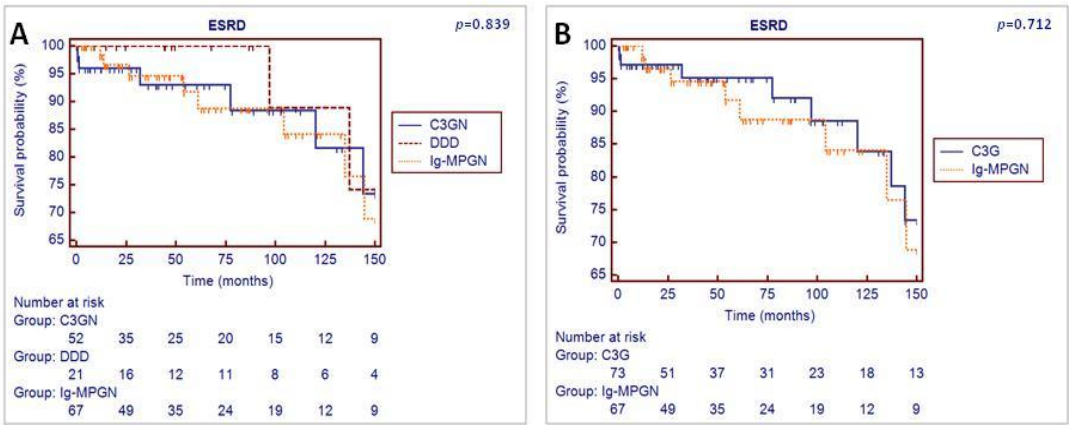


Figure 3.7. Kaplan-Meier renal survival analysis according to histology groups.

Table 3.6. Univariate analysis of the association of long-term renal outcome with clinical, laboratory and genetic features.

	HR	HR 95%CI	*p*
Gender (male=1, female=2)	1.02	0.38-2.72	0.974
Familiarity for nephropathy	0.38	0.09-1.71	0.209
Clinical data at onset			
Age of onset	1.01	0.98-1.04	0.588
Age of onset≥16	1.71	0.64-4.6	0.288
Gross hematuria	0.85	0.27-2.69	0.780
Nephrotic syndrome	4.80	1.73-13.34	**0.0027**
Renal impairment	4.74	1.76-12.77	**0.0021**
High Blood Pressure	2.17	0.76-6.22	0.148
Trigger	0.35	0.08-1.58	0.171
Serum C3	1.00	0.99-1.01	0.697
Serum C4	1.04	0.99-1.09	0.106
Plasma SC5b-9	1.00	1.00-1.00	0.221
Treatment			
Steroids	1.49	0.42-5.29	0.535
ACEi/ARB	0.51	0.17-1.53	0.231
Intensified immunosuppression	3.32	1.04-10.65	**0.043**
Light microscopy			
Sclerotic glomeruli	14.70	1.83-117.91	**0.011**
Crescents	22.03	3.27-148.42	**0.0015**
Degree of Mesangial proliferation	1.06	0.66-1.71	0.815
Degree of Endocapillary proliferation	1.15	0.73-1.8	0.546
Degree of Interstitial inflammation	1.51	0.84-2.72	0.169
Degree of Interstitial fibrosis	1.11	0.64-1.93	0.702
Degree of Arteriolar sclerosis	1.35	0.76-2.39	0.308
Immunofluorescence			
C3	0.98	0.38-2.52	0.969
IgA	0.54	0.16-1.79	0.312
IgG	1.20	0.81-1.79	0.369
IgM	1.26	0.76-2.1	0.370
C1q	1.35	0.85-2.14	0.199
FBG	1.05	0.60-1.86	0.859
Electron microscopy			
Mesangial deposits	0.66	0.23-1.88	0.440
Subepithelial deposits	0.76	0.25-2.32	0.635
Subendothelial deposits	1.28	0.41-3.95	0.672
Intramembranous non DDD-like deposits	0.50	0.15-1.75	0.281
Intramembranous DDD-like deposits	0.82	0.22-3.04	0.771
Diagnosis of C3GN*	0.93	0.33-2.68	0.900
Diagnosis of DDD9*	0.63	0.13-2.94	0.557
Absence of mutations or C3NeFs	3.20	1.2-9.1	**0.025**

*Ig-MPGN was used as reference category.

Table 3.6. *continue*

	HR	HR 95%CI	*p*
Common genetic variants			
***CFH* c.-331C>T (additive model)**	1.61	0.83-3.13	0.159
***CFH* c.-331C>T (dominant model)**	1.57	0.54-4.55	0.402
***CFH* c.-331C>T (recessive model)**	2.37	0.82-6.86	0.112
***CFH* V62I (additive model)**	0.35	0.08-1.46	0.147
***CFH* V62I (dominant model)**	0.34	0.08-1.49	0.152
***CFH* V62I (recessive model)**	nc		
***CFH* H402Y (additive model)**	0.98	0.5-1.92	0.959
***CFH* H402Y (dominant model)**	1.44	0.32-6.36	0.633
***CFH* H402Y (recessive model)**	0.80	0.29-2.2	0.661
***CFH* E936D (additive model)**	1.50	0.77-2.91	0.232
***CFH* E936D (dominant model)**	1.57	0.58-4.21	0.374
***CFH* E936D (recessive model)**	2.17	0.62-7.68	0.228
***CD46* c.-652G>A (additive model)**	0.82	0.38-1.75	0.600
***CD46* c.-652G>A (dominant model)**	0.37	0.1-1.33	0.129
***CD46* c.-652G>A (recessive model)**	1.13	0.39-3.26	0.821
***CD46* c.-366A>G (additive model)**	0.94	0.44-2	0.866
***CD46* c.-366A>G (dominant model)**	0.68	0.15-3.05	0.616
***CD46* c.-366A>G (recessive model)**	1.04	0.38-2.87	0.942
***CD46* A309V (additive model)**	3.98	0.5-31.35	0.190
***CD46* A309V (dominant model)**	3.98	0.5-31.35	0.190
***CD46* A309V (recessive model)**	nc		
***CD46* c.*783T>C (additive model)**	0.99	0.47-2.1	0.984
***CD46* c.*783T>C (dominant model)**	0.87	0.32-2.34	0.781
***CD46* c.*783T>C (recessive model)**	1.39	0.31-6.14	0.667
***CFB* L9H (additive model)**	nc		
***CFB* L9H (dominant model)**	nc		
***CFB* L9H (recessive model)**	nc		
***CFB* R32W (additive model)**	0.50	0.11-2.22	0.365
***CFB* R32W (dominant model)**	0.50	0.11-2.22	0.365
***CFB* R32W (recessive model)**	nc		
***CFB* R32Q (additive model)**	1.07	0.31-3.68	0.914
***CFB* R32Q (dominant model)**	1.09	0.31-3.85	0.890
***CFB* R32Q (recessive model)**	nc		
***CFB* E566A (additive model)**	nc		
***CFB* E566A (dominant model)**	nc		
***CFB* E566A (recessive model)**	nc		
***C3* P314L (additive model)**	0.93	0.36-2.44	0.885
***C3* P314L (dominant model)**	1.03	0.36-2.98	0.952
***C3* P314L (recessive model)**	nc		
***C3* R102G (additive model)**	1.17	0.51-2.68	0.713
***C3* R102G (dominant model)**	1.49	0.55-3.99	0.433
***C3* R102G (recessive model)**	nc		
***THBD* A473V (additive model)**	1.37	0.47-3.95	0.566
***THBD* A473V (dominant model)**	1.03	0.29-3.64	0.959
***THBD* A473V (recessive model)**	nc		

Table 3.7 Multivariate analysis of the association of long-term renal outcome with clinical, laboratory and genetic features.

	All patients		
	HR	**HR 95%CI**	***p***
Absence of mutations or C3NeFs	7.1	1.9-26.3	**0.004**
Sclerotic glomeruli (% of glomeruli)	69.3	3.1-1553	**0.008**
Crescents (% of glomeruli)	39.7	3.3-481	**0.004**
Nephrotic syndrome at onset	10.9	2.5-47	**0.002**

HR: Hazard ratio calculated by Multivariate Cox proportional-Hazards analysis. CI: Confidence Interval. nc: not calculable. Nephrotic syndrome was defined as: 24-hour proteinuria exceeding 3.5g in adults or 40mg/h/m2 in children together with albuminemia ≤3 g/dL. Intensified immunosuppression was also included in multivariate Cox Regression analysis but was not significantly associated with progress to ESRD (HR=3.9, 95%CI 0.65-23.9, p= 0.138).

3.4.6 Treatment

Seventy-two percent of patients received steroids, 74% angiotensin-converting enzyme inhibitors (ACEi) and/or angiotensin receptor blockers (ARB), and 40% intensified immunosuppression (cyclosporine A, mycophenolate and/or cyclophosphamide). Steroids, ACEi/ARB, or intensified immunosuppression did not correlate with renal outcome (Tables 3.6, 3.7 and 3.8). Eculizumab was administered for a mean (±S.D.) time of 14.9 (±10.1) months to 14 patients (4 Ig-MPGN, 4 DDD and 6 C3GN patients) who did not respond to other treatments. None of them have so far developed ESRD. However, with the exception of a previously published case,[318] the follow-up is too short to draw any conclusions.

Table 3.8. Treatment received by Ig-MPGN and C3G patients.

	Ig-MPGN	**C3G**		
			DDD	**C3GN**
Steroids	85%	64%	55%	68%
Intensified immunosuppression	45%	35%	39%	34%
Eculizumab	6%	15%	22%	12%
ACEi/ARB	80%	68%	83%	62%

Ig-MPGN, n=67; DDD, n=21; C3GN, n=52.
Intensified immunosuppression: cyclosporine A, mycophenolate and/or cyclophosphamide; ACEi: angiotensin-converting enzyme inhibitors; ARB: angiotensin receptor blockers.

3.5. Discussion

In this study, we classified a large cohort of idiopathic Ig-MPGN and C3G patients following current criteria,[250, 253] performed complement gene screening to disclose the genetic basis of these diseases, and correlated genetic, and biochemical data with clinical features to identify outcome predictors. We find genetic or acquired complement abnormalities in the majority of Ig-MPGN and C3G patients. We identify mutations in *C3*, *CFB*, *CFH*, *CD46* and *CFI*, in accordance with previous studies,[253, 263] and we report a *THBD* mutation in C3G. Finally, we report that patients without complement gene mutations or C3NeFs are characterized by worse renal survival.

A mutation in *C3*, leading to hyperfunctional C3 convertase, was reported in a family with DDD.[263] Here, we observe that *C3* mutations are also associated with Ig-MPGN and C3GN, and that they are the most frequent genetic abnormalities in these groups (Tables 3.3 and 3.4, and Figure 3.2). A mutation in the gene encoding FB was reported in a patient with C3GN.[311] Another study described a C3GN patient carrying the E566A *CFB* variant,[261] which however is present in about 2% of the control population, and in our cohort is not associated with either Ig-MPGN or C3G. In our cohort *CFB* mutations clustered in the Ig-MPGN histology group, supporting the association between alternative pathway dysregulation and Ig-MPGN. Altogether, mutations in the two components of the alternative pathway C3 convertase account for 63% of the identified genetic abnormalities. These findings are consistent with recent data showing a higher frequency of rare and novel *C3* and *CFB* variants in C3G patients compared with controls.[262]

Twenty-seven percent of mutations occur in genes encoding complement regulators and most of them affect *CFH*.[256] Notably, we report genetic abnormalities in *THBD*, an anticoagulant endothelial glycoprotein that also regulates the complement alternative

pathway (Table 3.4). We describe a Proline to Serine substitution at position 495 of thrombomodulin in a DDD male patient with disease onset at age 12, showing non-nephrotic range proteinuria and microhematuria at onset, low serum C3 levels, and C3NeF positivity. The *THBD* P495S mutation was previously reported in a patient affected by aHUS and leads to less efficient inactivation of C3b.[309] Interestingly, another patient carries the *THBD* D486Y rare variant, which was previously found in two aHUS patients and is characterized by defective suppression of the complement alternative pathway through FI-mediated C3b inactivation.[309] The D486Y carrier is a female with C3GN, disease onset at age 8, non-nephrotic range proteinuria at onset, low serum C3 levels and C3NeF positivity. Neither patient showed signs of TMA and/or aHUS. The allelic frequency of the D486Y variant is 0.7% in the general population, thus it cannot be considered causative of a rare disease such as C3G, but may instead be a disease risk factor. Indeed, C3G is considered a complex multifactorial disease.[264] It is possible that in the two patients carrying P495S and D486Y, the presence of C3NeFs acted as a second hit, leading to the development of C3G.

Thrombomodulin binds C3b and FH and enhances FI-mediated C3b inactivation, protecting the kidney from alternative pathway activation.[309] Thrombomodulin expression on glomerular endothelial cells is increased in MPGN and lupus nephritis, but not in other glomerulonephritides, and this upregulation may be a protective response to counterbalance complement hyperactivation occurring in MPGN glomeruli.[319, 320] Thrombomodulin also exists in soluble form in plasma that increases during inflammation and vascular injury, and regulates coagulation and inflammation.[321]

Further evidence supporting the involvement of *THBD* in the pathogenesis of C3G is provided by our finding that the *THBD* A473 common variant is significantly associated

with the risk of developing the disease (Table 3.5). Functional studies on *THBD* A473V are lacking. Based on A473V proximity to D486Y and P495S, we hypothesize that the 473 residue is relevant for C3b inactivation. Further evidence supporting the involvement of *THBD* in the pathogenesis of C3G is provided by our finding that the *THBD* A473 common variant is significantly associated with the risk of developing the disease, and that A473 inactivates C3b less efficiently than 473V. This hypothesis is supported by our observation that A473 homozygous C3G patients have lower serum C3 and higher SC5b-9 levels than 473V carriers. To identify the functional effects of A473V, additional studies are needed.

We also identified an association between C3G risk and the homozygous *CFH* V62 allele (Table 3.5). *CFH* V62 is less efficient than 62I at binding and inactivating C3b, and confers susceptibility to DDD, aHUS, and age-related macular degeneration.[264, 322, 323] Association studies highlighted the genetic complexity of both C3G and Ig-MPGN. Indeed, we show that the risk of developing C3G or Ig-MPGN strongly increases only when mutations are combined with common susceptibility variants (Figure 3.4). The combination of mutations with the homozygous *CFH* V62 and *THBD* A473 risk alleles strongly increases the probability of developing C3G. Conversely, in subjects carrying mutations but no susceptibility variants, the risk of C3G is comparable to subjects lacking both mutations and susceptibility variants.

In addition, we observed an association between Ig-MPGN and the *CD46* c.-366A and the *CFB* R32 alleles. Functional studies on the CD46 c.-366A>G polymorphism are not available, however its presence in the promoter of the gene suggests that it may impact gene expression. On the other hand, functional data show that the CFB R32 is characterized by an increased capability to form convertase and amplify complement activation compared to the 32Q and 32W alleles.[265, 324] Similarly to C3G, the risk of

Ig-MPGN is higher when the mutations are combined with a susceptibility variant, *CD46* c.-366A. Together, these observations may explain the incomplete penetrance of Ig-MPGN and C3G in mutation carriers. Analysis of both mutations and susceptibility variants may be useful for defining the risk of developing Ig-MPGN or C3G for patients' relatives during genetic counseling.

Interestingly, the complex pathogenesis of C3G and Ig-MPGN resembles that of another complement-related disorder, the atypical hemolytic uremic syndrome (aHUS), in which multiple hits are necessary for the disease to manifest. Such hits include triggers (infections, pregnancy or drugs), mutations in one or more complement genes, and common at-risk SNPs and haplotypes that act as susceptibility factors for the development of the disease.[293, 325, 326]

Previous studies showed that the *CFH* Y402H common variant is associated with DDD.[264, 327] In the present cohort, the prevalence of the 402H risk allele is also numerically higher in DDD patients compared to controls, however this difference is not statistically significant (Table 3.5). There is only a trend (p=0.063), when a recessive model of inheritance is considered for the risk allele. Notably, the CFH 402H allelic frequencies in the present cohort (0.45 in DDD *vs.* 0.37 in controls) were similar to those observed by Abrera-Abeleda *et al.* (0.48 in DDD vs. 0.37 in controls), therefore the fact that the statistical significance is not reached in this study is probably due to a low statistical power. Indeed, there were only 21 DDD patients included in the present study.

Notably, the prevalence of alternative pathway abnormalities (mutations, C3NeFs, low C3 levels with normal C4) was comparable in Ig-MPGN and C3G (Table 3.2). In particular, 56% and 65% of Ig-MPGN and C3G patients, respectively, carried mutations and/or C3NeFs. This finding, which is consistent with a previous report,[256] suggests that

Ig-MPGN and C3G have more commonalities than previously recognized, and that alternative pathway dysregulation plays a pathogenetic role in both. It is plausible that in some Ig-MPGN patients, immune-complex deposition may trigger the disease in predisposed subjects with genetic or acquired alternative complement pathway dysregulation.[253]

In the present study, nephrotic syndrome at onset and a higher proportion of sclerotic glomeruli and crescents in kidney biopsies negatively impact on renal survival during follow-up (Table 3.7), in accordance with previous studies.[255, 328] We also observe a higher risk of progressing to ESRD in patients without complement gene mutations or C3NeFs. Patients with gene mutations and/or C3NeFs are characterized by very low serum C3 levels and high plasma SC5b-9. In patients with Ig-MPGN or C3G carrying either complement gene mutations or C3NeFs, dysregulation of the alternative pathway in the fluid phase results in massive formation of complement products in blood, such as C3b and SC5b-9.[250] The transfer of these complement products and debris through the fenestrated endothelium to subendothelial region and mesangium triggers glomerular inflammation and leads to disease.[1]

The higher risk of ESRD in patients without mutations or C3NeFs might be explained by the presence of different underlying mechanisms of complement dysregulation resulting in renal damage. These mechanisms could involve solid-phase, kidney-restricted complement activation, rather than fluid phase activation, as suggested by intense glomerular C3 deposits in the face of only modestly altered serum C3 and plasma SC5b-9 levels. This complement pattern resembles that of C3G patients with genomic rearrangements in FH-related genes (*CFHR1-5*).[329] It has been suggested that there is continuous activation of C3 along the glomerular basement membrane, which under normal conditions is controlled by FH.[330] Functional studies revealed that the

abnormal FH-related proteins compete with FH for binding to C3b and to cell surfaces,[330] causing C3 activation products and C3G to accumulate.[85] However, *CFHR* rearrangements are rare, and further studies are warranted to evaluate the mechanism leading to kidney-restricted activation of complement and the worse renal outcome in patients without identified mutations and/or C3NeFs.

The present study has some limitations that must be taken into account. This is a multicenter and retrospective study that may be affected by heterogeneity in the diagnostic process and treatment. However, idiopathic Ig-MPGN and C3G are rare, and large prospective single center studies are extremely difficult. Moreover, multicenter studies have the advantage of increased result generalizability, and they allow more patients to be recruited, increasing statistical power.[331] To reduce the impact of heterogeneity, we reviewed all kidney biopsy reports and clinical files centrally. Next-generation sequencing is not sensitive for the detection of structural chromosomal rearrangements and/or insertions or deletions >10bp,[332] thus we may have missed some genetic abnormalities in the studied genes.

In conclusion, we identify alternative pathway abnormalities, i.e. mutations, C3NeFs, and low C3 levels with normal C4, similarly in Ig-MPGN and C3G. These findings do not support the current classification that restricts the definition of alternative pathway-mediated glomerulopathies to patients with "dominant C3" glomerular staining,[253] since the majority of cases classified as idiopathic Ig-MPGN are associated with abnormalities of the complement alternative pathway. Moreover, the presence of mutations alone does not significantly increase the risk of developing Ig-MPGN or C3G, but they do so when combined with common susceptibility variants, underlining that Ig-MPGN and C3G are complex diseases. Finally, we report that patients without complement gene mutations and/or C3NeFs have a higher risk of progressing to ESRD

than patients with identified mutations and/or C3NeFs. These results may also be interpreted as indicating the existence of at least two different pathogenic mechanisms leading to renal disease progression in Ig-MPGN/C3G.

4. CAN CLUSTER ANALYSIS HELP UNRAVEL THE PATHOGENESIS OF C3G/IC-MPGN AND PREDICT PROGRESSION TO ESRD?

4.1 Introduction

Membranoproliferative glomerulonephritis (MPGN) is a chronic nephropathy characterized by capillary wall thickening and mesangial expansion owing to increased matrix deposition and hypercellularity.[251] Traditionally, MPGN was classified into types I to III based on the presence of subendothelial, highly electron-dense intramembranous or subepithelial deposits on electron microscopy (EM), respectively.[248] Historically, many MPGN cases characterized by glomerular immune-complex deposits were shown to be secondary to infections, autoimmune diseases or malignancies.[248] More recently, MPGN cases with isolated C3 deposits carrying abnormalities of the alternative pathway of complement were observed.[248] To better reflect disease pathogenesis, a new classification based on immunofluorescence (IF) has been proposed that divides MPGN into immune-complex-associated MPGN (IC-MPGN), with significant glomerular immunoglobulin deposition, and C3 glomerulopathy (C3G), with dominant glomerular C3 deposition and little or no immunoglobulin deposition (C3 ≥2 grades of order of magnitude greater than any other immune-reactants).[248, 250, 253] C3G is further divided into dense-deposit disease (DDD) and C3 glomerulonephritis (C3GN), the latter lacking the intramembranous highly electron-dense deposits.[253] Notably, patients with predominant C3 glomerular staining on immunofluorescence and with proliferative patterns of injury other than MPGN are also included in C3G.[253]

The current classification is based on the assumption that IC-MPGN results from the deposition of immune complexes that trigger the complement system through the classical pathway, whereas C3G arises from abnormalities in the control of the complement alternative pathway.[248] Consequently, evaluation for infections, autoimmune diseases or monoclonal gammopathy in IC-MPGN is recommended, and conversely genetic screening in complement genes and assays for C3NeF are recommended in C3G.[248] It is believed that only few IC-MPGN patients remain idiopathic after extensive evaluation and that the disease may be initiated by immunoglobulin deposition and accelerated by complement alternative pathway abnormalities only in few IC-MPGN patients.[248, 253, 257]

The introduction of IF-based classification is an advance to an etiology-based diagnostic approach, but some issues require further investigation. First, the widely accepted "C3 dominant" criterion for the definition of C3G was established using DDD patients as a reference group and has not been validated using patients with known dysregulation of the complement alternative pathway.[300] Second, up to 16% of patients shift from IC-MPGN to C3G and vice versa when a kidney biopsy is repeated.[300, 333] Third, the distinction between DDD and C3GN is not always clear-cut, but there are some borderline cases.[253] Fourth, C3GN and DDD may not reflect different underlying causes, since both may affect different members of the same pedigree.[251, 329, 334, 335] Finally, complement alternative pathway dysregulation is common in idiopathic IC-MPGN patients, since most carry C3NeF and/or mutations in complement alternative pathway genes.[256] Additionally, about half of IC-MPGN patients have low serum C3 with normal C4 levels, indicating activation of the alterative pathway of complement without significant activation of the classical pathway.[256] Altogether, the assumption that immunoglobulin-positive MPGN is immune-complex-mediated and that immunoglobulin-negative MPGN is complement-mediated has yet to be validated.

Recently, unsupervised cluster analysis has been introduced as an objective data-driven grouping method to explore whether patients can be classified into relatively homogeneous groups.[336] Simplistically, subjects with many aspects in common are placed close together and subjects with many dissimilarities are placed further apart. Therefore, the characteristics of the groups are not defined by the operator but arise from the data without *a priori* assumptions. This approach was successfully used to identify subtypes in different diseases.[336-338]

In this study we evaluated whether the "C3-dominant" criterion, currently used to distinguish C3G from IC-MPGN patients, divides patients with genetic or acquired alternative complement pathway abnormalities from patients without. We show that the "C3 dominant" criterion does not effectively capture patients with alternative pathway abnormalities. We also explored whether cluster analysis based on histology findings, serum complement profile, genetic data and clinical features could be a tool for disclosing, in C3G and IC-MPGN patients, distinct disease entities characterized by specific pathophysiological mechanisms. Through cluster analysis, we identify four groups of patients with relatively homogeneous phenotypes. Patients with known alternative pathway abnormalities are separated from those without, the latter clustering into a unique group. Patients with alternative pathway abnormalities are further divided into three groups with distinct underlying disease pathogenetic mechanisms. These clusters are useful for better understanding the pathogenesis of the disease and predicting the risk of progression to ESRD.

4.2 Specific aims

1. To evaluate whether the current IF-based C3G/IC-MPGN classification capture all patients with complement-mediated pathogenesis, studying a group of patients carefully characterized for genetic and biochemical parameters of complement alternative pathway dysregulation.

2. To explore whether cluster analysis based on histologic findings, clinical features, serum complement profile and genetic data could be a tool for better understanding the disease etiology.

4.3 Methods

4.3.1 Patients

We analyzed a cohort of 163 patients recruited through the Italian Registry of MPGN affected by C3G or IC-MPGN. Patients with secondary MPGN, atypical hemolytic uremic syndrome preceding or concomitant to MPGN onset, and patients for whom IF, EM or DNA samples were not available were not included in this cohort. Of the 163 patients, 137 had been described previously.[339] All participants provided informed written consent. The study was approved by the Ethics Committee of the Azienda Sanitaria Locale of Bergamo (Italy). The patients were classified according to the recent C3G/IC-MPGN classification.[250, 253] Briefly, patients with a MPGN pattern at light microscopy and with glomerular immunoglobulin and C3 deposits by IF were considered IC-MPGN, while patients with MPGN or mesangial proliferative patterns and "dominant C3" glomerular staining (intensity≥2 magnitude orders than other immune-reactants) were considered C3G. By EM, C3G was further classified as DDD or C3GN.[253]

Clinical data were recorded using a standardized Case Report Form. Serum C3 and C4 levels were assessed by kinetic nephelometry. Plasma SC5b-9 levels were measured using the MicroVue SC5b-9 Plus EIA commercial kit (SC5b-9 Plus, Quidel).[294] C3NeF activity was determined by purifying IgG from plasma and assessing their ability to stabilize cell-bound C3bBb convertase as previously described.[298] Targeted Next-generation sequencing for the genetic screening of all exons and flanking regions of *CFH*, *CD46*, *CFI*, *CFB*, *C3* and *THBD* genes was performed by highly multiplex PCR using the Ion AmpliSeq™ Library Kit 2.0 followed by template preparation and sequencing on an Ion PGM Sequencer as previously described.[339] Variants fulfilling the following criteria were considered as mutations: 1. variants previously identified in

patients with complement disorders and with functional assays supporting variant pathogenicity; 2. variants with allele frequency ≤0.001 in any subpopulation of the ExAC database and CADD score ≥10.

4.3.2 Variable reduction and Clustering

We used 101 different variables (Table 4.1): 18 for clinical features at onset or during follow-up, 3 for treatment, 30 for pathologic findings (13 for light microscopy, 12 for IF and 5 for EM), 10 regarding the serum or plasma complement profile, 4 regarding genetic or acquired complement abnormalities and 36 regarding common SNPs of complement genes.

To reduce redundancy, in the first step we performed principal component analysis (PCA) of the correlated variables (for example between the 'absolute serum C3 value' and the 'abnormally low serum C3' variables or between closely located SNPs with high linkage disequilibrium) as shown in Table 4.1. To remove noisy variables,[337] we performed a second principal component analysis including all the independent variables and the components derived from the first step. In PCA only significant components (i.e. eigenvalue ≥1) were retained, since components with eigenvalue <1 contribute little to explain the relationships between original variables.[337] Finally, an unsupervised hierarchical clustering was performed based on significant components using the Ward's method with Euclidean distances. For the cluster analysis, imputation of the missing values was performed with the *k-Nearest Neighbor* method and iterative robust regression. Patients and variables with >20% of missing data were removed.

Table 4.1. Variables included in the principal component analysis. The number of principal components with eigenvalue >1 in each step are also reported.

Variable	Principal components obtained in the 1st step	Principal components obtained in the 2nd step
Gender (male/female)	1	18
Age at onset log2 Age at onset Age at onset ≥18 years	1	
Microhematuria at onset Gross hematuria at onset	1	
Proteinuria at onset Nephrotic syndrome at onset Nephrotic syndrome in overall disease course	2	
Renal impairment at onset ESRD at onset Chronic kidney disease in overall disease course ESRD in overall disease course	2	
Trigger event at onset	1	
High blood pressure at onset High blood pressure in overall disease course	1	
Thrombotic microangiopathy in overall disease course	1	
Familiarity for nephropathy	1	
Serum C3 (mg/dl) log2 Serum C3 (mg/dl) Pathologically low serum C3 (<90 mg/dl) Low serum C3 at 1st investigation and normal in later evaluations	2	
Serum C4 (mg/dl) log2 Serum C4 (mg/dl) Pathologically low serum C4 (<10 mg/dl)	1	
Plasma SC5b-9 (ng/ml) log2 plasma SC5b-9 Pathologically high plasma SC5b-9 (>303 ng/ml)	1	
C3NeF positive N° of mutations Presence of mutations and/or C3NeF Mutations *vs.* C3NeF	2	
% Sclerotic glomeruli on LM	1	
% of crescents on LM Presence of crescents on LM	1	

Table 4.1. *Continue.*

Variable	Principal components obtained in the 1st step	Principal components obtained in the 2nd step
Degree of mesangial proliferation on LM Presence of mesangial proliferation on LM	1	...
Degree of endocapillary proliferation on LM Presence of endocapillary proliferation on LM	1	
Degree of interstitial inflammation on LM Presence of interstitial inflammation on LM	1	
Degree of interstitial fibrosis on LM Presence of interstitial fibrosis on LM	1	
Degree of arteriolar sclerosis on LM Presence of arteriolar sclerosis on LM	1	
C3 on IF C3 ≥1+ on IF	1	
IgA on IF IgA ≥1+ on IF	1	
IgG on IF IgG ≥1+ on IF	1	
IgM on IF IgM ≥1+ on IF	1	
C1q on IF C1q ≥1+ on IF	1	
Fibrinogen on IF Fibrinogen ≥1+ on IF	1	
Mesangial deposits on EM	1	
Subepithelial deposits on EM	1	
Subendothelial deposits on EM	1	
Intramembranous granular deposits on EM	1	
Intramembranous highly electron-dense ribbon-like deposits on EM	1	
Corticosteroid treatment Massive immunosuppression treatment ACE-inhibitor/Renin-Angiotensin blocker treatment	1	

Table 4.1. *Continue.*

Variable	Principal components obtained in the 1st step	Principal components obtained in the 2nd step
CFH c.-331C>T genotype	3	...
CFH c.-331C>T carrier C		
CFH c.-331C>T carrier T		
CFH V62I genotype		
CFH V62I carrier V		
CFH V62I carrier I		
CFH Y402H genotype		
CFH Y402H carrier Y		
CFH Y402H carrier H		
CFH E936D genotype		
CFH E936D carrier E		
CFH E936D carrier D		
CD46 c.-652G>A genotype	2	
CD46 c.-652G>A carrier G		
CD46 c.-652G>A carrier A		
CD46 c.-366A>G genotype		
CD46 c.-366A>G carrier A		
CD46 c.-366A>G carrier G		
CD46 c.*783T>C genotype		
CD46 c.*783T>C carrier T		
CD46 c.*783T>C carrier C		
CFB L9H genotype	3	
CFB R32W genotype		
CFB R32Q genotype		
CFB R32Q carrier R		
CFB R32Q carrier Q		
CFB E566A genotype		
C3 P314L genotype	2	
C3 P314L carrier P		
C3 P314L carrier L		
C3 R102G genotype		
C3 R102G carrier R		
C3 R102G carrier G		
THBD A473V genotype	1	
THBD A473V carrier A		
THBD A473V carrier V		

SNP carriers where no variability was observed (i.e. 100% *vs.* 0%) were removed.

4.3.3 Identification of Classification Criteria

For ease of use, quantitative variables were dichotomized by adopting cutoffs. The selection of features that could be used to identify clusters was performed in three phases as previously described.[340] First, the odds ratio was calculated for each feature with the presence or absence of a cluster as dependent variable and the features with statistically significant odds ratios were included in the next phase. Second, logistic regression in a backward stepwise procedure was carried out for each cluster separately as dependent variable and the features that were significantly associated with one of the clusters were selected for the next phase. Third, multinomial logistic regression was performed with clusters as the dependent variable and the features from phase two as the independent variables and those non reaching significance were discarded. This model contained all the finally selected features together.

In the first step of the algorithm that ascribes patients to clusters, we assigned one point to each feature and we summed the points for each patient to obtain a score. ROC curve was drawn for this score to determine the optimal cutoff to distinguish clusters.

4.3.4 Statistical analyses

We used the ANOVA test for continuous variables and the χ^2 test (or Fisher Exact test, where appropriate) for categorical variables. Survival analyses considered as cumulative fractions of patients free of events were estimated using Kaplan-Meier survival curves. Multinomial logistic regression was performed with the *mlogit* package on the R platform. Only variables with a *p*-value <0.05 in the univariate analysis were considered in the multivariate analysis. *P*-values <0.05 were considered statistically significant.

4.4 Results

4.4.1 Features according to the histologic classification

The histological, biochemical, genetic and clinical features of the 163 patients, classified according to the recent C3G/IC-MPGN classification, are reported in Table 4.2. Sixty-six patients (40%) had C3GN, 24 (15%) DDD and 73 (45%) IC-MPGN. A kidney biopsy was performed after a median of 0.4, 1.1 and 0.4 years from the onset for C3GN, DDD and IC-MPGN, respectively, without significant differences between histologic groups. Age of onset and sex distribution did not differ between histologic groups. Complement gene likely pathogenic variants (herein defined as mutations) were identified in 24%, 17% and 16% of patients with C3GN, DDD and IC-MPGN, respectively (Tables 4.2 and 4.3). C3NeF prevalence was higher in DDD patients (77%) compared to C3GN (40%, p=0.003) and IC-MPGN (42%, p=0.005). Overall, in all three histologic groups, the majority of the patients carried mutations and/or C3NeF with a higher prevalence in DDD (82%) compared to IC-MPGN (51%, p=0.012).

Table 4.2. Clinical features, complement assessment, genetic screening and histologic features in patients classified according to the recent histologic classification.

Variable	C3GN	DDD	IC-MPGN	Overall p-value
N	66	24	73	
Gender (% males)	61%	58%	49%	0.387
Data at onset				
Age (yr)	17.7 (±15.5)	14.7 (±11.1)	20.3 (±15.7)	0.259
Microhematuria	86%	92%	84%	0.743
Gross hematuria	38%	46%	32%	0.449
Proteinuria	92%	88%	89%	0.726
Nephrotic syndrome	30%	25%	44%	0.126
Renal impairment	20%	4%	25%	0.079
Trigger event	34%	33%	26%	0.614
Familiarity for nephropathy	18%	13%	14%	0.763
Serum C3 (mg/dl)	51.5 (±41.9)	32 (±35.7)	50.5 (±39.5)	0.104
Serum C4 (mg/dl)	21.5 (±8.8)	24.1 (±9.7)	21.1 (±11)	0.436
Plasma SC5b-9 (ng/ml)	1171 (±1300)	558 (±531)	1029 (±1114)	0.090
Low serum C3 and normal serum C4	75%	83%	67%	0.262
C3NeF positive	40%	77%[a,c]	42%	**0.007**
Mutation carriers	24%	21%	18%	0.648
Mutation carriers and/or C3NeF	60%	82%[c]	51%	**0.042**
Data during follow-up				
Nephrotic syndrome	47%	50%	70%[a]	**0.017**
High blood pressure	35%	25%[c]	51%	**0.041**
Chronic kidney disease	32%	29%	44%	0.239
ESRD	11%	4%	10%	0.763
Thrombotic microangiopathy	6%	0%	4%	0.587
Histological features				
Time Onset to Biopsy (yr)	2.7 (±5.3)	2.8 (±4.1)	2.1 (±4.6)	0.733
Light microscopy				
Sclerotic glomeruli	8% (±17%)	2% (±7%)	8% (±14%)	0.205
Crescents	3% (±11%)	6% (±20%)	7% (±19%)	0.354
Degree of mesangial proliferation*	1.7 (±1)	1.8 (±0.7)	2 (±0.9)	0.207
Degree of endocapillary proliferation*	1 (±1.1)	0.9 (±1.1)	1.3 (±1)	0.244
Degree of interstitial inflammation*	0.5 (±0.8)	0.6 (±0.8)	0.9 (±0.8)[a]	**0.028**
Degree of interstitial fibrosis*	0.4 (±0.7)	0.1 (±0.4)	0.6 (±0.8)	0.059
Degree of arteriolar sclerosis*	0.3 (±0.8)	0 (±0.1)	0.3 (±0.6)	0.209
Immunofluorescence				
C3*	2.7 (±0.5)	2.8 (±0.3)	2.5 (±0.7)	0.070
IgA*	0 (±0.1)	0.1 (±0.3)	0.5 (±0.8)[a,b]	**$2.3x10^{-6}$**
IgG*	0.2 (±0.4)	0.2 (±0.3)	1.6 (±1.2)[a,b]	**$4.6x10^{-20}$**
IgM*	0.3 (±0.5)[b]	0.7 (±0.7)	1.3 (±1)[a,b]	**$1.1x10^{-12}$**
C1q*	0.2 (±0.3)	0.1 (±0.3)	1.1 (±1)[a,b]	**$5.7x10^{-13}$**
Fibrinogen*	0.3 (±0.7)	0.3 (±0.8)	0.3 (±0.7)	0.990
Electron microscopy				
Mesangial deps	73%	50%	56%	0.062
Subepithelial deps	59%[b,c]	8%[c]	38%	**$8.1x10^{-5}$**
Subendothelial deps	70%	8%[a,c]	82%	**$2.8x10^{-10}$**
Intramembr. granular deps	54%	0%[a,c]	46%	**$2.1x10^{-5}$**
Intramembr. highly electron-dense ribbon-like deps	0%	100%[a,c]	4%	**$6.7x10^{-26}$**

[a]significantly different *vs.* C3GN; [b]significantly different *vs.* DDD; [c]significantly different *vs.* IC-MPGN.*Degree of mesangial proliferation, endocapillary proliferation, interstitial inflammation, interstitial fibrosis, and arteriolar sclerosis, as well as IF findings were graded using a scale of 0 to 3, including 0, trace, 1+, 2+ and 3+. Continuous variables are reported as Mean (±S.D.). Intramembr.: intramembranous; deps: deposits. Serum C3: reference 90-180 mg/dl; serum C4: reference 10-40 mg/dl; plasma SC5b-9: reference ≤303 ng/ml.

Table 4.3. Patients with likely pathogenic variants present in the cohort.

Likely pathogenic variant	Zyg.	Hist.	Clust. /Gr.	Global Freq.	Max Freq.	Funct. studies	CADD	C3NeF	Serum C3 (mg/dl)	Plasma SC5-9 (ng/ml)	New/ known
CFH:R2I (chr1:196621252G>T)	Het	DDD	3/3	0	0	No	11	Yes	54	286	(81)
CFH:R78G (chr1:196642281A>G)	Hom	C3GN	1/1	0	0	No	16	No	15	1530	(81, 306-308)
CFH:R78G (chr1:196642281A>G)	Hom	C3GN	1/1	0	0	No	16	No	14	4571	(81, 306-308)
CFH:P88T (chr1:196643004C>A)	Hom	IC	1/1	0	0	No	29	No	48	2074	(81)
CFH:P88T (chr1:196643004C>A)	Hom	IC	1/1	0	0	No	29	No	5	3596	(81)
CFH:R127C (chr1:196645147C>T)	Het	C3GN	1/1	0	0	No	33	NA	72	2789	(81)
CFH G133R (chr1:196645165G>A)	Het	C3GN	1/1	$1.7x10^{-5}$	$3.0x10^{-5}$	No	31	NA	55	355	New
CFH:Y1008X (chr1:196711070delTA)	Hom	C3GN	1/1	0	0	No	32	No	18	NA	(81)
CFH:R1210C (chr1:196716375C>T)	Het	DDD	4/3	$1.7x10^{-4}$	$2.8x10^{-4}$	Yes	12	No	154	368	(81, 256, 294, 301)
CD46:K66N (chr1:207930459A>T)	Het	IC	2/2	$5.3x10^{-4}$	$9.0x10^{-4}$	No	13	No	63	466	(81)
CFI:c.1-4C>T (chr4:110723131G>A)	Het	C3GN	1/1	$8.4x10^{-6}$	$1.5x10^{-5}$	No	17	Yes	17	399	(81)
CFI:G57D (chr4:110687868C>T)	Het	DDD	3/3	$8.2x10^{-6}$	$6.1x10^{-5}$	No	25	Yes	47	321	(81)
CFI G119R (chr4:110685820C>T)	Het	C3GN	1/1	$5.3x10^{-4}$	$9.5x10^{-4}$	Yes	22	NA	79	310	(81, 294)
CFB:G161R (chr6:31914966G>A)	Het	IC	2/1	$6.9x10^{-5}$	$3.0x10^{-4}$	No	27	Yes	36	1600	(81)
CFB:H451R (chr6:31917278A>G)	Het	IC	2/2	0	0	No	25	Yes	40	1874	(81)
CFB:R679W (chr6:31919196C>T)	Hom	IC	2/2	0	0	No	31	No	20	291	(81)
CFI:R317W* (chr4:110670750G>A)	Het	C3GN	1/1	$9.9x10^{-5}$	$2.3x10^{-4}$	No	16	NA	18	NA	(81, 310)

Zyg.: zygosity; Hist.: Histologic group according to the current classification; Global Freq.: variant frequency in all subjects of the ExAC database (v0.3); Max Freq.: maximum frequency among the subpopulations of the ExAC database (v0.3). Funct. studies: Functional studies supporting pathogenicity are available. CADD: CADD Phred score (v1.3). In brackets the reference of the variants already described in MPGN, C3G or aHUS patients. Serum C3 in mg/dl; Plasma SC5-9 in ng/ml. IC: IC-MPGN. Serum C3: reference 90-180 mg/dl; plasma SC5b-9: reference ≤303 ng/ml.
*found in the same patient.

Table 4.3. *Continue.*

Likely pathogenic variant	Zyg.	Hist.	Clust. /Gr.	Global Freq.	Max Freq.	Funct. studies	CADD	C3NeF	Serum C3	Plasma SC5-9	New/ known
C3:R505C (chr19:6710823G>A)	Het	IC	2/1	$8.3x10^{-6}$	$1.5x10^{-5}$	No	25	Yes	18	495	(81)
C3:V619M (chr19:6707931C>T)	Het	C3GN	1/1	$2.9x10^{-4}$	0.001	No	22	No	57	561	(81)
C3:G637R (chr19:6707877C>G)	Het	IC	1/2	$2.2x10^{-4}$	$3.8x10^{-4}$	No	24	Yes	12	1845	(81)
C3:I761del (chr19:6702553delGAT)	Het	C3GN	1/1	0	0	No	14	No	33	845	New
C3:I761del (chr19:6702553delGAT)	Het	C3GN	1/1	0	0	No	14	No	38	888	New
C3:R1042Q* (chr19:6694471C>T)	Het	C3GN	1/1	0	0	No	32	NA	18	NA	(81, 301)
C3:K1051M (chr19:6694444T>A)	Het	IC	1/1	$1.6x10^{-5}$	$3.0x10^{-5}$	Yes	23	NA	41	348	(81, 304)
C3:S1063N (chr19:6693465C>T)	Het	IC	1/2	$1.1x10^{-4}$	$8.1x10^{-4}$	Yes	10	No	70	235	(81)
C3:R1303H (chr19:6685060C>T)	Het	C3GN	1/1	$8.3x10^{-6}$	$1.5x10^{-5}$	No	28	No	25	413	(81)
C3:R1320Q (chr19:6685009C>T)	Het	C3GN	1/1	0	0	No	28	No	15	1087	(81)
C3:D1362N (chr19:6684607C>T)	Het	IC	4/4	$4.9x10^{-5}$	$1.2x10^{-4}$	No	11	NA	112	337	(81)
C3:C1518R (chr19:6679214A>G)	Het	C3GN	2/2	0	0	No	26	Yes	6	2520	(81)
C3:D1625H (chr19:6678010G>C)	Het	C3GN	1/1	0	0	No	12	Yes	2	3750	(81)
THBD:D486Y (chr20:23028686G>T)	Het	C3GN	1/1	0	0	Yes	12	Yes	40	1298	(81, 309)
THBD:P495S (chr20:23028659G>A)	Het	DDD	3/3	$5.8x10^{-4}$	0.001	Yes	6	Yes	18	137	(81, 309)

Zyg.: zygosity; Hist.: Histologic group according to the current classification; Global Freq.: variant frequency in all subjects of the ExAC database (v0.3); Max Freq.: maximum frequency among the subpopulations of the ExAC database (v0.3). Funct. studies: Functional studies supporting pathogenicity are available. CADD: CADD Phred score (v1.3). In brackets the reference of the variants already described in MPGN, C3G or aHUS patients. Serum C3 in mg/dl; Plasma SC5-9 in ng/ml. IC: IC-MPGN. Serum C3: reference 90-180 mg/dl; plasma SC5b-9: reference ≤303 ng/ml.
*found in the same patient.

4.4.2 Efficacy of the "C3 dominant" criterion for identifying DDD and alternative pathway dysregulation

We evaluated whether the "C3 dominant" criterion, widely used to define C3G, efficiently identifies patients with intramembranous highly electron-dense deposits and patients with alternative pathway dysregulation. As shown in Table 4.4, “C3 dominant” staining captured 24 out of 27 patients with intramembranous highly electron-dense deposits (sensitivity=89%). On the other hand, “C3 dominant” staining had low sensitivity (57%) for identifying patients with complement gene mutations and/or C3NeF (Table 4.4). Consistently, the "C3 dominant" criterion had low sensitivity (56%) for identifying patients with low C3 and normal C4. In addition, compared to IC-MPGN, C3G patients did not show a significantly higher prevalence of mutations and/or C3NeF (64% *vs.* 53%, p=0.154), or a prevalence of low C3 and normal C4 (76% *vs.* 68%, p=0.240).

Variability in the interpretation of immune-reactant intensity may lead to misclassification of C3G and IC-MPGN.[329] To overcome bias due to immune-reactant intensity misevaluation, we repeated the analyses, including only the extremes of the phenotypic continuum in the 163 patients: C3G patients with absent immunoglobulin or C1q staining (n=31) *vs.* IC-MPGN patients with 3+ staining in either at least one of the immunoglobulin classes or C1q (n=33) (Table 4.4). Even with this setting, the prevalence of mutations and/or C3NeF, as well as the prevalence of low C3 and normal C4, were not significantly higher in C3G compared to IC-MPGN (56% *vs.* 50%, p=0.670, and 77% *vs.* 66%, p=0.333, respectively). Accordingly, with the same setting, the sensitivity of IF for identifying patients with complement gene mutations and/or C3NeF or for identifying patients with low C3 but normal C4 did not improve (48% and 50%, respectively).

Table 4.4. Sensitivity and specificity of the C3-dominant criterion on IF to capture patients with intramembranous electron-dense deposits, mutations and/or C3NeF and low serum C3 and normal C4.

	C3G	IC-MPGN	*p*	"C3 dominant" criterion	
				Sensitivity	Specificity
All patients					
Intramembranous electron-dense deposits	0.27	0.04	**0.0001**	0.89	0.51
Mutations and/or C3NeF	0.64	0.53	0.154	0.57	0.55
Low serum C3 and normal C4	0.76	0.68	0.240	0.56	0.55
Only C3G patients with isolated C3 and IC-MPGN patients with 3+ staining in any of Ig or C1q*					
Intramembranous electron-dense deposits	0.23	0.03	**0.025**	0.83	0.49
Mutations and/or C3NeF	0.56	0.50	0.670	0.48	0.57
Low serum C3 and normal C4	0.77	0.66	0.333	0.50	0.63

*C3G patients with no immunoglobulin or C1q staining on IF and IC-MPGN patients with 3+ staining in at least one of the immunoglobulins or C1q. C3G, n=90; IC-MPGN, n=73.

4.4.3 Clustering analysis

To investigate the presence of relatively homogeneous groups of patients based on histologic, biochemical, genetic and clinical features, we performed unsupervised hierarchical cluster analysis using 101 variables reduced to 18 principal components (Table 4.1).

We identified 4 groups (clusters) as shown in Figure 4.1 and Table 4.5. Initially, cluster 4 was separated from the others. It includes patients with a lower prevalence of mutations and/or C3NeF, normal or mildly altered serum C3 (s-C3) levels and later onset. Subsequently, cluster 3, which was characterized by a higher prevalence of intramembranous electron-dense deposits, was divided from clusters 1 and 2, which showed highly increased plasma SC5b-9 levels. Finally, compared to cluster 1, cluster 2 included patients with stronger IgG and C1q staining.

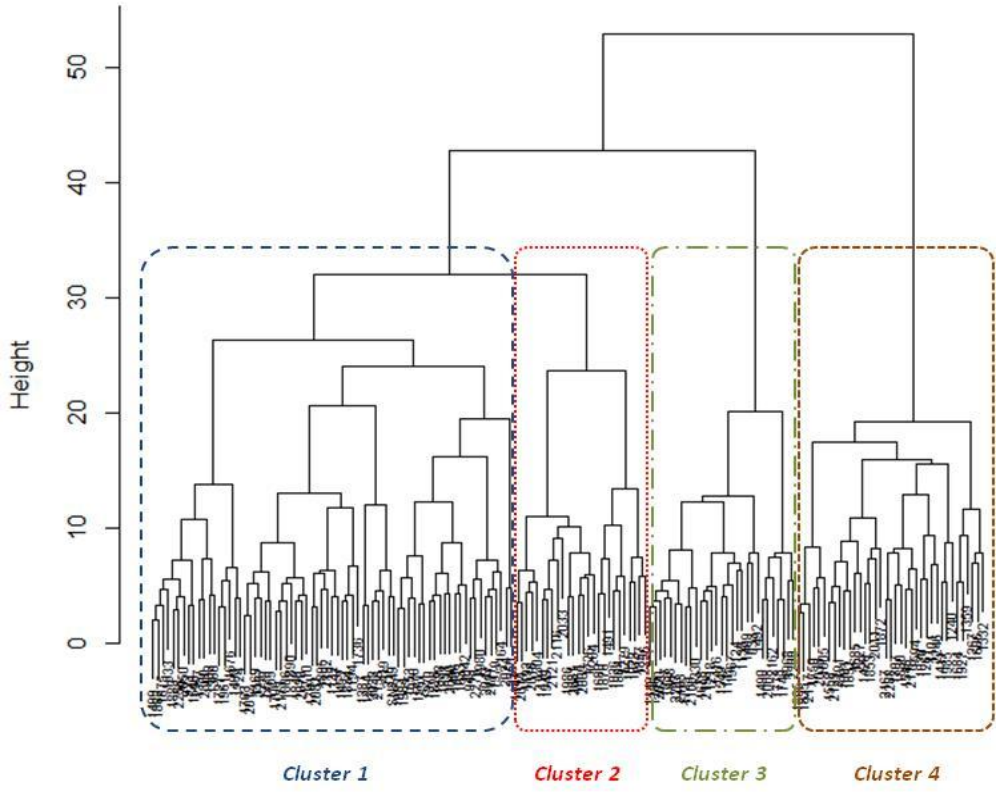


Figure 4.1. Dendrogram illustrating the results of the cluster analysis in 163 patients with C3G or IC-MPGN. Each vertical line at the extremity of the dendrogram (bottom) represents a patient and the length of the vertical lines represents the degree of dissimilarity between patients or groups of patients. The coloured boxes define the four clusters.

Cluster 1 was composed of 71 patients with a high prevalence of mutations and/or C3NeF (70%), low s-C3 (38±30 mg/dl) and very high SC5b-9 levels (1315±1300 ng/ml). The mean age of onset was 13.6 years. These patients were characterized by few crescents (1±4%, p=0.002) and low interstitial inflammation (0.4±0.5, p=2x10^{-5}) in the kidney biopsy compared to the other clusters. On EM, subendothelial (82%) deposits were higher compared to clusters 3 and 4 (p=8x10^{-10}).

Cluster 2 was composed of 26 patients characterized by a high prevalence of mutations and/or C3NeF (76%), low s-C3 levels (28±22 mg/dl) and high SC5b-9 levels (1681±1334 ng/ml). Compared to the other clusters, this cluster included patients with stronger staining of IgG (1.9±1.1, $p=8x10^{-8}$) and C1q (1.4±1.0, $p=5x10^{-6}$) in the kidney biopsy. On EM, subendothelial deposits were more frequent (88%) compared to clusters 3 and 4 ($p=8x10^{-10}$). Patients in this group had a higher prevalence of nephrotic syndrome at onset (65%, $p=0.007$) compared to the other groups.

Patients in cluster 3 (n=29) showed a high prevalence of mutations and/or C3NeF (81%) and low s-C3 levels (33±29 mg/dl). This cluster was characterized by a higher prevalence of intra-membranous highly electron-dense deposits (83%, $p=2x10^{-20}$), and lower subepithelial (7%, $p=7x10^{-4}$) and subendothelial (14%, $p=8x10^{-10}$) deposits compared to the other three clusters.

Finally, cluster 4 was composed of 37 patients with a lower prevalence of mutations and/or C3NeF (9%, $p=9x10^{-10}$), higher s-C3 (93±42 mg/dl, $p=2x10^{-16}$) and later age of onset (31±19 years, $p=5x10^{-9}$) compared to the other three clusters. Interestingly, cluster 4 included all 7 patients in the present cohort that initially presented C3GN or IC-MPGN with bright glomerular C3 staining and later developed thrombotic microangiopathy in a median time of 3.2 years (IQR, 2.3-5.9 years) from onset.

Table 4.5. Clinical features, complement assessment, genetic screening and histologic features in patients classified according to the clusters obtained through cluster analysis.

Variable	1	2	3	4	Overall p-value
N	71	26	29	37	
Gender (% males)	55%	46%	52%	65%	0.493
Data at onset					
Age (yr)	13.6 (±11.1)	20.5 (±13.2)[a,c]	12.6 (±8.1)	30.8 (±19.3)[a,b,c]	**$4.7x10^{-9}$**
Microhematuria	90%	83%	96%	72%[a,c]	**0.026**
Gross hematuria	38%	27%	50%	30%	0.257
Proteinuria	94%	92%	90%	81%	0.198
Nephrotic syndrome	31%	65%[a,c,d]	28%	30%	**0.007**
Renal impairment	14%	27%	3%[b]	38%[a,c]	**0.002**
Trigger event	32%	24%	23%	36%	0.612
Familiarity for nephropathy	24%[b,c]	4%	7%	14%	**0.048**
Serum C3 (mg/dl)	38 (±30)	27.7 (±21.8)	33.1 (±29.2)	93.2 (±42.1)[a,b,c]	**$2.3x10^{-16}$**
Serum C4 (mg/dl)	21.3 (±10.2)	17.6 (±8.6)[c,d]	22.6 (±8.7)	24.9 (±10.6)	**0.037**
Plasma SC5b-9 (ng/ml)	1315 (±1300)[c,d]	1681 (±1334)[c,d]	519 (±501)	385 (±270)	**$5.8x10^{-7}$**
Low serum C3 & normal C4	81%	81%	86%	39%	**$5.4x10^{-5}$**
C3NeF positive	48%[c]	68%	81%	0%[a,b,c]	**$1.0x10^{-9}$**
Mutation carriers	28%	27%	17%	5%[a,b]	**0.037**
Mutation carriers and/or C3NeF	70%	76%	81%	9%[a,b,c]	**$8.5x10^{-10}$**
Data during follow-up					
Nephrotic syndrome	46%	85%[a,c]	52%	65%	**0.006**
High blood pressure	35%	31%	28%	68%[a,b,c]	**0.002**
Chronic kidney disease	23%	50%[a]	31%	59%[a,c]	**$8.0x10^{-4}$**
ESRD	3%	8%	7%	22%	**0.013**
Thrombotic microangiopathy	0%	0%	0%	19%[a]	**$5.2x10^{-5}$**
Histological features					
Time Onset to Biopsy (yr)	2.6 (±5.1)	2.2 (±4.6)	2.2 (±3.7)	2.6 (±5.2)	0.977
Light microscopy					
Sclerotic glomeruli	5% (±12%)	6% (±10%)	2% (±5%)	18% (±22%)[a,b,c]	**$1.3x10^{-5}$**
Crescents	1% (±4%)[b,c,d]	6% (±15%)	14% (±27%)	8% (±19%)	**0.002**
Degree of mesangial proliferation*	2 (±0.8)[b,d]	1.5 (±1.2)	2.1 (±0.7)[b,d]	1.6 (±1)	**0.020**
Degree of endocapillary proliferation*	1 (±1.1)	1.6 (±1.1)	1.3 (±1.1)	0.9 (±1.1)	0.067
Degree of interstitial inflammation*	0.4 (±0.5)[b,c,d]	0.8 (±0.7)	0.8 (±0.8)	1.2 (±1)	**$2.0x10^{-5}$**
Degree of interstitial fibrosis*	0.2 (±0.4)	0.7 (±0.9)[a,c]	0.2 (±0.5)	1 (±0.9)[a,c]	**$2.4x10^{-7}$**
Degree of arteriolar sclerosis*	0.1 (±0.4)	0.3 (±0.6)[c]	0.1 (±0.2)	0.6 (±1)[a,c]	**$8.0x10^{-4}$**
Immunofluorescence					
C3*	2.6 (±0.7)	2.8 (±0.3)	2.7 (±0.4)	2.6 (±0.5)	0.186
IgA*	0.2 (±0.5)	0.6 (±0.7)[a,c]	0.1 (±0.3)	0.4 (±0.8)	**0.009**
IgG*	0.6 (±0.9)	1.9 (±1.1)[a,c,d]	0.4 (±0.7)[d]	0.9 (±1.2)	**$7.8x10^{-8}$**
IgM*	0.8 (±0.9)	1.1 (±0.9)	0.7 (±0.7)	0.7 (±0.9)	0.272
C1q*	0.5 (±0.8)	1.4 (±1)[a,c,d]	0.2 (±0.5)	0.6 (±1)	**$4.9x10^{-6}$**
Fibrinogen*	0.3 (±0.7)	0.8 (±1.1)[a,d]	0.3 (±0.7)	0.1 (±0.3)	**0.001**
Electron microscopy					
Mesangial deps	73%[c,d]	68%	50%	44%	**0.016**
Subepithelial deps	49%	52%	7%[a,b,d]	49%	**$7.0x10^{-4}$**
Subendothelial deps	82%	88%	14%[a,b,d]	63%[a,b]	**$7.9x10^{-10}$**
Intramembr. granular deps	55%[c,d]	64%[c,d]	11%	29%	**$4.0x10^{-5}$**
Intramembr. highly electron-dense ribbon-like deps	2%	0%	83%[a,b,d]	6%	**$1.9x10^{-20}$**

[a]significantly different *vs.* Cluster 1; [b]significantly different *vs.* Cluster 2; [c]significantly different *vs.* Cluster 3; significantly different *vs.* Cluster 4. *Degree of mesangial proliferation, endocapillary proliferation, interstitial inflammation, interstitial fibrosis, and arteriolar sclerosis, as well as IF findings were graded using a scale of 0 to 3, including 0, trace, 1+, 2+ and 3+. Continuous variables are reported as Mean (±S.D.). Intramembr.: intramembranous; deps: deposits. Serum C3: reference 90-180 mg/dl; serum C4: reference 10-40 mg/dl; plasma SC5b-9: reference ≤303 ng/ml.

Clusters 1 and 4 were composed prevalently by C3GN and IC-MPGN patients (Table 4.6). Cluster 2 included mostly IC-MPGN patients (77%), while cluster 3 was mainly composed of DDD patients. Interestingly, 21 out of 24 DDD patients and all 3 IC-MPGN patients with intramembranous highly electron-dense deposits were included in cluster 3.

Table 4.6. Overlapping between histologic groups and clusters.

Histologic diagnosis	*Cluster 1*	*Cluster 2*	*Cluster 3*	*Cluster 4*	Overall *p*-value
C3GN	41	6	2	17	
DDD	1	0	21	2	
IC-MPGN	29	20	6	18	$<2.2x10^{-16}$

C3GN, n=66; DDD, n=24; IC-MPGN, n=73.

4.4.4 A clinically applicable algorithm for cluster identification

Then we selected the minimum set that could be used in clinical practice to ascribe patients to the new clusters from the histologic, biochemical, genetic and clinical features available at onset. The selection of the features that independently correlate with the clusters was performed in three phases using binomial and multinomial logistic regression (Tables 4.7 and 4.8). We observed that 11 different features associated with one or more clusters: mutations and/or C3NeF, serum C3 ≤70 mg/dl, mild-to-severe interstitial fibrosis, C1q glomerular staining ≥1+, fibrinogen glomerular staining ≥1+, mesangial deposits, intramembranous highly electron-dense deposits, adult-onset, the CFH H402Y, the C3 R102G and the THBD A473V variants.

To finally make up a short, simple and more robust list of criteria, we repeated the above analyses, adopting a significance threshold of 0.001 (Tables 4.9 and 4.10). Multinomial logistic regression showed that mutations and/or C3NeF, and serum C3

≤70mg/dl reduced the probability of belonging to cluster 4 (Relative Risk, RR=0.05, p=5.7x10^{-4}, and RR=0.03, p=1.3x10^{-5}, respectively), while age of onset ≥18 years increased this probability (RR=15.9, p=5.5x10^{-4}). The presence of C1q glomerular deposits (≥1+) increased the probability of belonging to cluster 2 (RR=8.7, p=1.5x10^{-4}) and intramembranous electron-dense deposits highly increased the probability of belonging to cluster 3 (RR=861, p=1.5x10^{-6}) (Table 4.10).

We empirically combined the results of the above analysis in a 3 step algorithm (Figure 4.2). Briefly, in the first step, we considered mutations and/or C3NeF, s-C3 ≤70mg/dl, age of onset <18 years and the intramembranous dense-deposits. Based on ROC curve analysis, patients fell into cluster 4 when they carried fewer than two of the above features. Second, patients with intramembranous dense-deposits were classified in cluster 3. Third, patients with ≥1+ C1q deposits were assigned to cluster 2 and the remaining in cluster 1. The characteristics of these algorithm-identified clusters are shown in Table 4.11. The concordance of the algorithm-based classification with the cluster analysis is 72% and with the multinomial logistic regression predictions is 88%.

Table 4.7. Univariate and backward multivariate binomial logistic regression analysis to identify the criteria for patients' classification adopting the 0.05 significance threshold.

Cluster	Feature	Univariate analysis		Multivariate analysis	
		OR	*p*	OR	*p*
1	Age of onset ≥18 years	0.29	7.4E-04		
1	Familiarity for nephropathy	3.3	0.010		
1	Presence of sclerotic glomeruli on LM	0.36	0.004		
1	Serum C3 ≤70 mg/dl	2.8	0.008		
1	Plasma SC5b-9 >700 ng/ml	3.1	0.001		
1	Mutation carriers and/or C3NeF	2.3	0.013	6.0	0.004
1	Mutation carriers	2.2	0.046		
1	Presence of crescents on LM	0.10	0.000	0.02	$1.0x10^{-4}$
1	Interstitial inflammation (≥+1) on LM	0.39	3.7E-03		
1	Interstitial fibrosis (≥+1) on LM	0.22	2.0E-04	0.03	<0.0001
1	Arteriolar sclerosis (≥+1) on LM	0.31	0.029		
1	Mesangial deposits on EM	2.3	0.020	10.7	0.001
1	Subendothelial deposits on EM	4.2	4.1E-04		
1	Intramembranous granular deposits on EM	2.8	0.005		
1	Intramembranous highly electron-dense ribbon-like deposits on EM	0.03	0.001	3E-04	<0.0001
1	CFH H402Y genotype	0.66	0.046		
1	CFH H402Y HH & HY vs. YY	2.3	0.011	8.5	0.002
1	CD46 c.-652G>A genotype	0.59	0.024		
1	CD46 c.-652G>A GG & GA vs. AA	2.2	0.014	9.7	0.001
1	CFB R32Q genotype	0.34	0.022		
1	CFB R32Q QQ & RQ vs. RR	0.31	0.018	0.08	0.009
1	C3 P314L genotype	0.42	0.008		
1	C3 R102G genotype	0.39	0.002		
1	C3 P314L LL & PL vs. PP	0.42	0.018		
1	C3 R102G GG & RG vs. RR	0.40	0.009	0.07	$2.0x10^{-4}$
1	THBD A473V genotype	0.39	0.016		
1	THBD A473V VV & AV vs. AA	0.39	0.024	0.11	0.002
2	Nephrotic syndrome at onset	4.4	0.001		
2	Serum C3 ≤70 mg/dl	11.8	0.017		
2	Sporadically low serum C3	3.1	0.022		
2	Plasma SC5b-9 >700 ng/ml	3.3	0.008	7.6	0.005
2	C3NeF positive	3.1	0.015		
2	Mesangial proliferation (≥+1) on LM	0.22	0.002		
2	Endocapillary proliferation (≥+1) on LM	3.7	0.013		
2	Interstitial fibrosis (≥+1) on LM	2.6	0.030		
2	IgA deposits (≥1+) on IF	4.0	0.005		
2	IgG deposits (≥1+) on IF	7.0	1.0E-04		
2	C1q deposits (≥1+) on IF	9.4	<0.0001	76.0	$2.0x10^{-4}$
2	Fibrinogen* on IF	4.4	2.0E-03	54.5	$5.0x10^{-4}$
2	Subendothelial deposits on EM	4.6	0.024		
2	Intramembranous granular deposits on EM	3.6	0.009		
2	CFH c.331C>T genotype	1.8	0.047		
2	CFH c.331C>T CC & CT vs. TT	0.22	0.003		
2	CFH H402Y genotype	4.9	3.2E-04		
2	CFH H402Y HH & HY vs. YY	0.13	1.0E-04		
2	CFH E936D genotype	3.2	3.2E-04		
2	CFH E936D EE & ED vs. DD	0.05	3.5E-05	0.01	$4.0x10^{-4}$
2	CFH E936D DD & ED vs. EE	2.6	0.031		

Table 4.7. *Continue.*

Cluster	Feature	Univariate analysis		Multivariate analysis	
		OR	*p*	Cluster	Feature
2	CD46 c.-652G>A genotype	4.9	0.001		
2	CD46 c.-652G>A GG & GA vs. AA	0.13	4.1E-04	0.01	$5.0x10^{-4}$
2	CD46 c.-366G>A genotype	3.1	0.007		
2	CD46 c.-366G>A GG & GA vs. AA	0.31	0.014		
2	CD46 c.*783T>C genotype	0.27	0.003		
2	CD46 c.*783T>C CC & TC vs. TT	0.25	0.005		
3	Age of onset ≥18 years	0.25	0.015	0.05	0.027
3	Renal impairment at onset	0.12	0.040		
3	Serum C3 ≤70 mg/dl	4.0	0.031		
3	Plasma SC5b-9 >700 ng/ml	0.31	0.018		
3	Mutation carriers and/or C3NeF	3.1	0.021		
3	C3NeF positive	5.9	3.0E-04		
3	IgG deposits (≥1+) on IF	0.26	0.011		
3	C1q deposits (≥1+) on IF	0.25	0.015		
3	Subepithelial deposits on EM	0.06	6.2E-04	0.06	0.032
3	Subendothelial deposits on EM	0.06	1.1E-07		
3	Intramembranous granular deposits on EM	0.10	0.002		
3	Intramembranous highly electron-dense ribbon-like deposits on EM	244	3.0E-12	453.3	<0.0001
3	CFH c.331C>T genotype	0.46	0.026		
3	CFH c.331C>T TT & CT vs. CC	0.36	0.016		
3	CFH H402Y YY & HY vs. HH	0.39	0.040		
3	CFH E936D genotype	0.42	0.050		
3	CFB R32Q genotype	3.0	0.015		
3	CFB R32Q QQ & RQ vs. RR	2.9	0.025		
4	Age of onset ≥18 years	7.6	1.2E-06	31.6	0.003
4	Microhematuria at onset	0.28	0.009		
4	Proteinuria at onset	0.33	0.041		
4	Renal impairment at onset	3.7	0.002		
4	Serum C3 ≤70 mg/dl	0.04	<0.0001	0.01	$4.0x10^{-4}$
4	Plasma SC5b-9 >700 ng/ml	0.12	2.0E-04		
4	Mutation carriers and/or C3NeF	0.04	2.8E-08	5E-03	$6.0x10^{-4}$
4	C3NeF positive	0.02	1.0E-04		
4	Mutation carriers	0.17	0.018		
4	Presence of sclerotic glomeruli on LM	6.1	<0.0001		
4	Interstitial inflammation (≥+1) on LM	3.3	2.7E-03		
4	Interstitial fibrosis (≥+1) on LM	5.3	<0.0001	52.7	0.004
4	Arteriolar sclerosis (≥+1) on LM	3.2	0.013		
4	Mesangial deposits on EM	0.35	0.009	0.06	0.011
4	Intramembranous granular deposits on EM	0.31	0.017	0.08	0.024
4	CFB R32W genotype	0.27	0.043	0.01	0.006
4	C3 P314L genotype	2.3	0.007		
4	C3 P314L PP & PL vs. LL	0.10	0.008		
4	C3 P314L LL & PL vs. PP	2.2	0.039		
4	C3 R102G genotype	2.7	5.8E-04		
4	C3 R102G RR & RG vs. GG	0.14	0.003		
4	C3 R102G GG & RG vs. RR	3.1	0.004		
4	THBD A473V genotype	3.4	9.2E-04		
4	THBD A473V VV & AV vs. AA	3.8	0.001	21.1	0.025

Table 4.8. Features that independently predict clusters obtained by multivariate multinomial logistic regression and adopting a 0.05 significance threshold.

Feature	Prevalence	Group vs. reference	*β*	RR (e^{β})	*p*
Mutations and/or C3NeF	59%	4	-4.9	0.01	0.009
Serum C3 ≤70 mg/dl	73%	4	-4.8	0.01	0.002
Interstitial fibrosis (≥1+) on LM	29%	2	2.8	16	0.003
		4	4.1	58	0.008
C1q deposits (≥1+) on IF	40%	2	2.3	10	0.004
Fibrinogen deposits (≥1+) on IF	17%	2	2.8	16	0.006
Mesangial deposits on EM	62%	4	-4.4	0.01	0.004
Intramembranous highly electron dense deposits	17%	3	7.1	1240	4.0×10^{-6}
Age of onset ≥18 years	34%	4	5.5	247	0.002
CFH H402Y: HH & HY vs. YY	58%	2	-3.2	0.04	5.9×10^{-4}
C3 R102: GG & RG vs. RR	34%	4	3.6	35	0.004
THBD A473V: VV & AV vs. AA	23%	4	3.0	19.8	0.046

[a]The group with the greatest number of patients (cluster 1) was taken as reference group. [b]Relative risk.

Table 4.9. Univariate and backward multivariate binomial logistic regression analysis to identify the criteria for patients' classification adopting the 0.001 significance threshold.

Cluster	Feature	Univariate analysis		Multivariate analysis	
		OR	*p*	OR	*p*
1	Age of onset ≥18 years	0.29	$7.4x10^{-4}$		
1	Plasma SC5b-9 >700 ng/ml	3.1	$6.0x10^{-4}$		
1	Presence of crescents on LM	0.10	$2.0x10^{-4}$		
1	Interstitial fibrosis (≥+1) on LM	0.22	$2.0x10^{-4}$	0.17	<0.0001
1	Subendothelial deposits on EM	4.2	$4.1x10^{-4}$	4.93	$1.0x10^{-4}$
2	IgG deposits (≥1+) on IF	7.0	$1.0x10^{-4}$		
2	C1q deposits (≥1+) on IF	9.4	<0.0001	12.37	<0.0001
2	CFH H402Y genotype	4.9	$3.2x10^{-4}$		
2	CFH H402Y HH & HY vs. YY	0.13	$1.0x10^{-4}$		
2	CFH E936D genotype	3.2	$3.2x10^{-4}$		
2	CFH E936D EE & ED vs. DD	0.05	$3.5x10^{-5}$	0.07	0.001
2	CD46 c.-652G>A GG & GA vs. AA	0.13	$4.1x10^{-4}$	0.11	$8.0x10^{-4}$
3	C3NeF positive	5.9	$3.0x10^{-4}$		
3	Subepithelial deposits on EM	0.06	$6.2x10^{-4}$		
3	Subendothelial deposits on EM	0.06	$1.1x10^{-7}$		
3	Intramembranous highly electron-dense ribbon-like deposits on EM	244	$3.0x10^{-12}$	209.6	<0.0001
4	Age of onset ≥18 years	7.6	$1.2x10^{-6}$	14.2	$5.0x10^{-4}$
4	Serum C3 ≤70 mg/dl	0.04	<0.0001	0.02	<0.0001
4	Plasma SC5b-9 >700 ng/ml	0.12	$2.0x10^{-4}$		
4	Mutation carriers and/or C3NeF	0.04	$2.8x10^{-8}$	0.06	$3.0x10^{-4}$
4	C3NeF positive	0.02	$1.0x10^{-4}$		
4	Presence of sclerotic glomeruli on LM	6.1	<0.0001		
4	Interstitial fibrosis (≥+1) on LM	5.3	<0.0001		
4	C3 R102G genotype	2.7	$5.8x10^{-4}$		
4	THBD A473V genotype	3.4	$9.2x10^{-4}$		

Table 4.10. Features that independently predict clusters obtained by multivariate multinomial logistic regression and adopting a 0.001 significance threshold.

Feature	Prevalence	*Group vs.* reference[a]	β	RR[b] (e^{β})	*p*
Mutations and/or C3NeF	59%	4	-2.9	0.05	$5.7x10^{-4}$
Serum C3 ≤70 mg/dl	73%	4	-3.6	0.03	$1.3x10^{-5}$
Intramembranous highly electron dense deposits	17%	3	6.8	861	$1.5x10^{-6}$
Presence of C1q deposits (≥1+)	40%	2	2.2	8.7	$1.5x10^{-4}$
Age of onset ≥18 years	34%	4	2.8	15.9	$5.5x10^{-4}$

[a]The group with the greatest number of patients (cluster 1) was taken as reference group. [b]Relative risk. A significance *p*-value threshold of 0.001 was adopted.

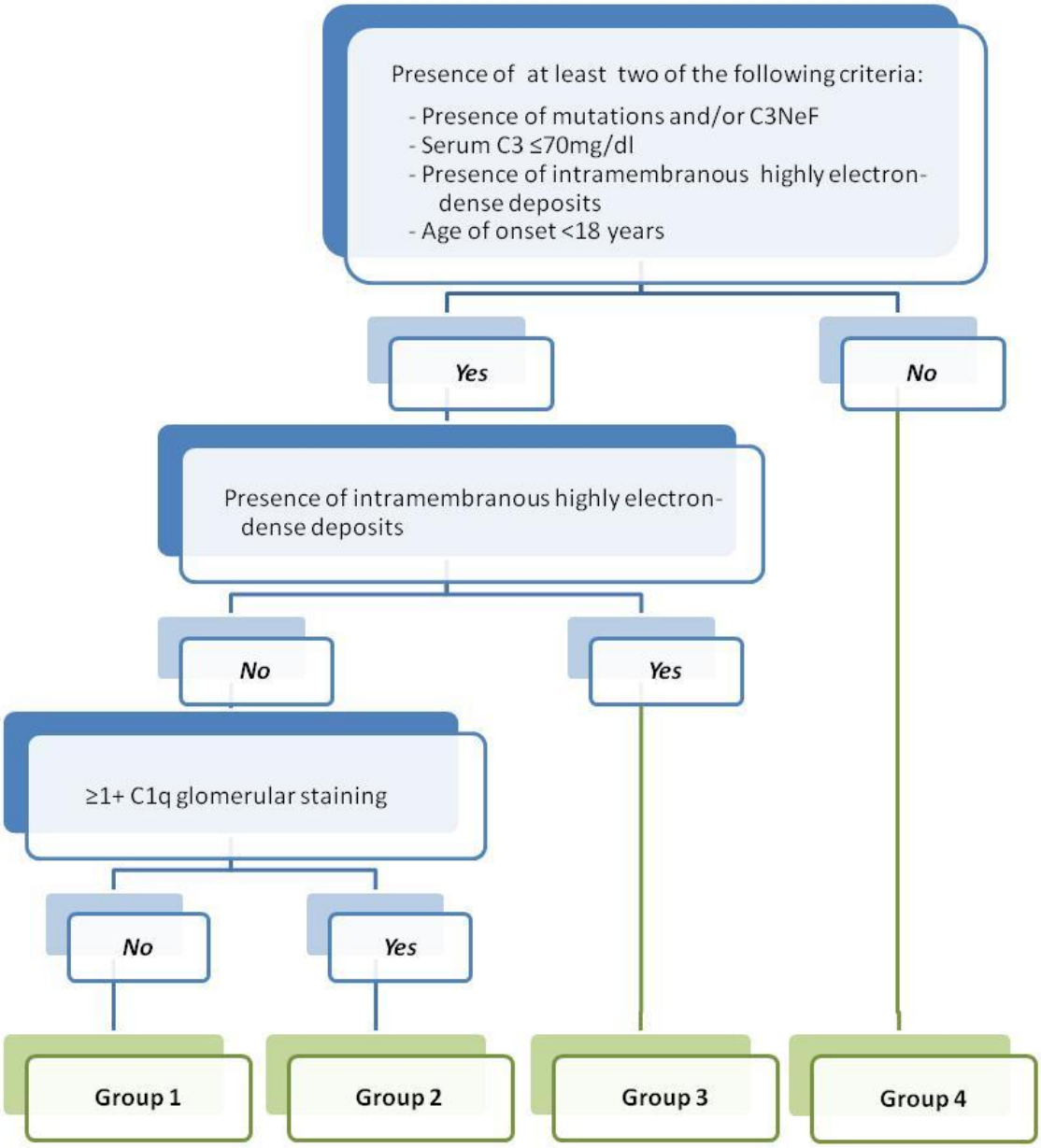


Figure 4.2. Criteria for cluster definition.

Table 4.11. Clinical features, complement assessment, genetic screening and histologic features in patients classified according to the 3-step algorithm clusters.

Variable	1	2	3	4	Overall *p*-value
N	53	32	26	52	
Gender (% males)	53%	44%	62%	62%	0.379
Data at onset					
Age (yr)	11.1 (±7.6)	15.7 (±11.1)	14.3 (±10.8)	29.6 (±18.3)[a,b,c]	**$6.8x10^{-11}$**
Microhematuria	87%	93%	92%	78%	0.240
Gross hematuria	42%	25%	46%	33%	0.266
Proteinuria	92%	97%	85%	87%	0.333
Nephrotic syndrome	30%	69%[a,c,d]	23%	27%	**$2.2x10^{-4}$**
Renal impairment	15%	19%	4%	33%[a,c]	**0.015**
Trigger event	35%	16%	30%	34%	0.276
Familiarity for nephropathy	21%	13%	12%	14%	0.705
Serum C3 (mg/dl)	36.1 (±27.6)[b]	23.4 (±25.1)	32.2 (±34.6)	84.1 (±38)[a,b,c]	**$2.3x10^{-16}$**
Serum C4 (mg/dl)	22.3 (±9.5)	17.4 (±10.4)[a,c,d]	23.8 (±9.4)	22.8 (±10)	**0.047**
Plasma SC5b-9 (ng/ml)	1284 (±1306)[c,d]	1834 (±1293)[c,d]	546 (±523)	486 (±613)	**$5.4x10^{-8}$**
Low serum C3 & normal C4	90%	69%	85%	51%	**$5.4x10^{-5}$**
C3NeF positive	60%	72%	78%	0%[a,b,c]	**$1.1x10^{-12}$**
Mutation carriers	38%	22%	23%	2%[a,b,c]	**$1.3x10^{-4}$**
Mutation carriers and/or C3NeF	85%	81%	83%	4%[a,b,c]	**$2.3x10^{-17}$**
Data during follow-up					
Nephrotic syndrome	47%	88%[a,c,d]	46%	56%	**0.001**
High blood pressure	30%	41%	23%	60%[a,c]	**0.004**
Chronic kidney disease	23%	34%	27%	58%[a,b,c]	**0.001**
ESRD	4%	9%	4%	15%	0.177
Thrombotic microangiopathy	0%	0%	0%	13%[a,b]	**0.002**
Histological features					
Time Onset to Biopsy (yr)	2.6 (±4.6)	1.5 (±2.9)	2.7 (±4)	2.9 (±6.1)	0.600
Light microscopy					
Sclerotic glomeruli	3% (±7%)	6% (±13%)	2% (±6%)	15% (±21%)[a,b,c]	**$9.1x10^{-5}$**
Crescents	5% (±17%)	1% (±4%)	6% (±20%)	8% (±18%)	0.378
Degree of mesangial proliferation*	1.9 (±0.9)	1.8 (±1.1)	1.9 (±0.7)	1.7 (±1)	0.564
Degree of endocapillary proliferation*	1.2 (±1.2)	1.5 (±1)	1 (±1.1)	0.9 (±1)	0.082
Degree of interstitial inflammation*	0.4 (±0.6)[b,d]	0.8 (±0.7)	0.7 (±0.8)	1 (±0.9)	**0.005**
Degree of interstitial fibrosis*	0.2 (±0.5)	0.5 (±0.8)[a]	0.2 (±0.4)	0.8 (±0.9)[a,c]	**$1.8x10^{-4}$**
Degree of arteriolar sclerosis*	0.1 (±0.3)	0.2 (±0.5)	0 (±0.1)	0.6 (±1)[a,c]	**$1.8x10^{-4}$**
Immunofluorescence					
C3*	2.7 (±0.5)	2.8 (±0.5)	2.8 (±0.3)	2.5 (±0.7)	0.071
IgA*	0.1 (±0.4)	0.5 (±0.6)[a,c]	0.1 (±0.3)	0.4 (±0.8)[a]	**0.004**
IgG*	0.3 (±0.7)	1.6 (±1)[a,c,d]	0.3 (±0.7)	1.1 (±1.2)[a,c]	**$4.8x10^{-9}$**
IgM*	0.6 (±0.8)	1.3 (±0.9)[a,c,d]	0.7 (±0.7)	0.8 (±0.9)	**0.004**
C1q*	0 (±0.1)[c,d]	1.7 (±0.7)[a,c,d]	0.2 (±0.5)[d]	0.6 (±0.9)	**$6.9x10^{-21}$**
Fibrinogen*	0.4 (±0.9)	0.5 (±0.8)	0.3 (±0.8)	0.1 (±0.3)[a,b]	**0.037**
Electron microscopy					
Mesangial deps	68%	67%	54%	56%	0.473
Subepithelial deps	48%	48%	12%[a,b,d]	48%	**0.008**
Subendothelial deps	76%	89%	8%[a,b,d]	73%	**$1.2x10^{-10}$**
Intramembr. granular deps	56%	60%	0%[a,b,d]	40%	**$6.1x10^{-6}$**
Intramembr. highly electron-dense ribbon-like deps	0%	0%	100%[a,b,d]	2%	**$5.7x10^{-29}$**

[a]significantly different *vs.* Cluster 1; [b]significantly different *vs.* Cluster 2; [c]significantly different *vs.* Cluster 3; [d]significantly different *vs.* Cluster 4. *Degree of mesangial proliferation, endocapillary proliferation, interstitial inflammation, interstitial fibrosis, and arteriolar sclerosis, as well as IF findings were graded using a scale of 0 to 3, including 0, trace, 1+, 2+ and 3+. Continuous variables are reported as Mean (±S.D.). Intramembr.: intramembranous; deps: deposits. Serum C3: reference 90-180 mg/dl; serum C4: reference 10-40 mg/dl; plasma SC5b-9: reference ≤303 ng/ml.

4.4.5 Prognostic significance of the new cluster-based classification

We evaluated whether the four clusters identified by the algorithm were characterized by a different renal outcome. Kaplan-Meier analyses showed that patients in the fourth cluster of the algorithm-based classification have a higher risk of ESRD during follow-up (Figure 4.3A). Similar results were obtained with Kaplan-Meier analyses of renal outcome in the groups identified by the unsupervised hierarchical cluster analysis (Figure 4.3B), while there were no differences between histologic groups or between C3G and IC-MPGN (Figure 4.3).

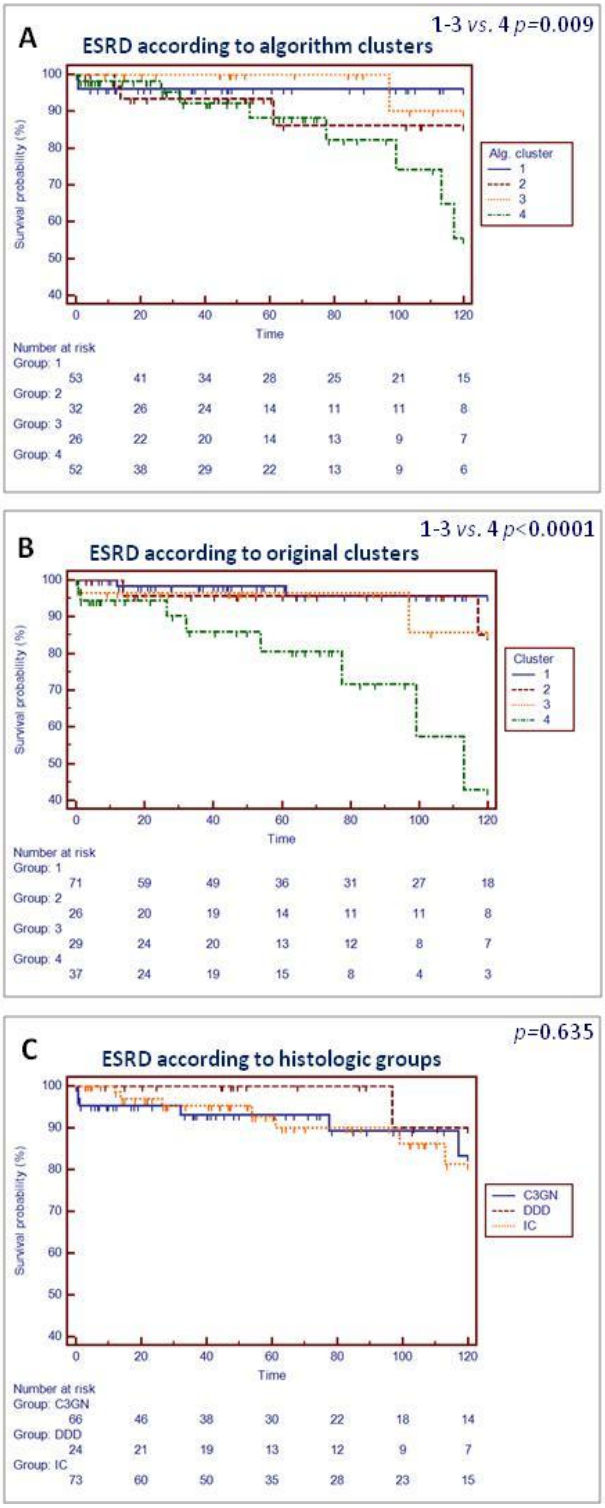


Figure 4.3. Kaplan-Meier renal survival analysis according to the clusters defined by the 3-step algorithm (panel A), the clusters obtained by the cluster analysis (panel B) and the histologic groups (panel C).

4.4.6 Differences in complement gene mutation distribution and C3NeF activity between clusters

We evaluated the distribution of the mutations in the different algorithm-identified clusters. We observed that mutations affecting *C3* and *CFB*, the two components of the alternative pathway C3 and C5 convertases, were more frequent in clusters 1 and 2 (14% and 23% of patients, respectively) compared to cluster 3 (0%) (Figure 4.4A). At variance, no difference was observed between histologic groups or between C3G and IC-MPGN (Figure 4.4B). Interestingly, cluster 1 included all 5 patients with homozygous *CFH* mutations.

To evaluate whether there are differences between C3NeFs characteristics in algorithm-identified clusters, we analyzed the efficiency of C3NeF to stabilize C3 convertase in C3NeF-positive patients. C3NeFs of patients in cluster 1 showed lower C3 convertase stabilizing activity (57±34%) compared to cluster 3 (87±23%, p=0.008) (Figure 4.4C), while C3NeF activity did not differ between histologic groups or between C3G and IC-MPGN (Figure 4.4D).

Comparable results were obtained when we performed the above analyses using the groups identified by the unsupervised hierarchical cluster analysis (Figure 4.5).

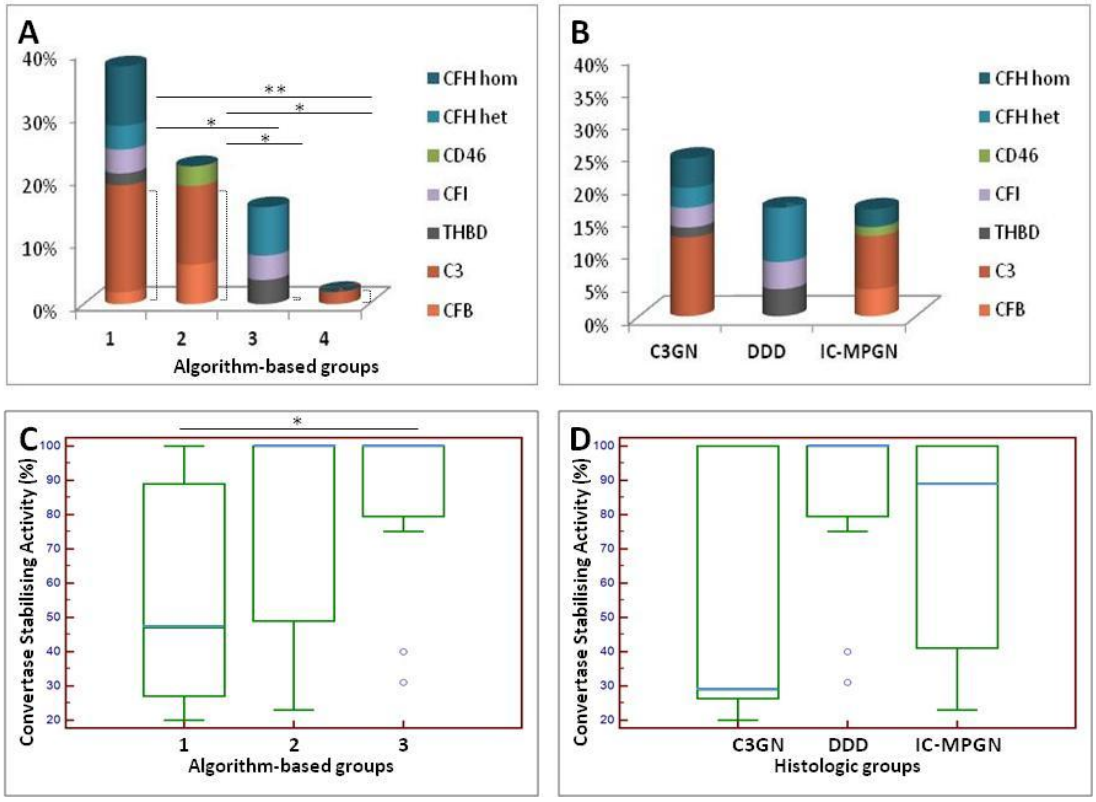


Figure 4.4. Panels A and B. Distribution of the mutations according to the algorithm-based groups (panel A) and the current histologic classification (panel B). Mutations in *C3* and *CFB* are significantly overrepresented in clusters 1 and 2 compared to clusters 3 and 4 (panel A) but no statistically significant difference are observed between histologic groups (panel B) (**p<0.01, *p<0.05). **Panels C and D.** C3NeF residual activity evaluated by hemolytic assay in C3NeF positive patients mutations according to the algorithm-based groups (panel C) and the current histologic classification (panel D) (*p<0.05). Group 4 is not represented in panel C, since it included no C3NeF positive patients. The central box represents the values from 25° to 75° percentile. The blue lines represent the medians. A line extends from the minimum to the maximum value, excluding "outsiders" (circles) defined as a value that is smaller than the 25° percentile minus 1.5 times the interquartile range, or larger than the 75° percentile plus 1.5 times the interquartile range.

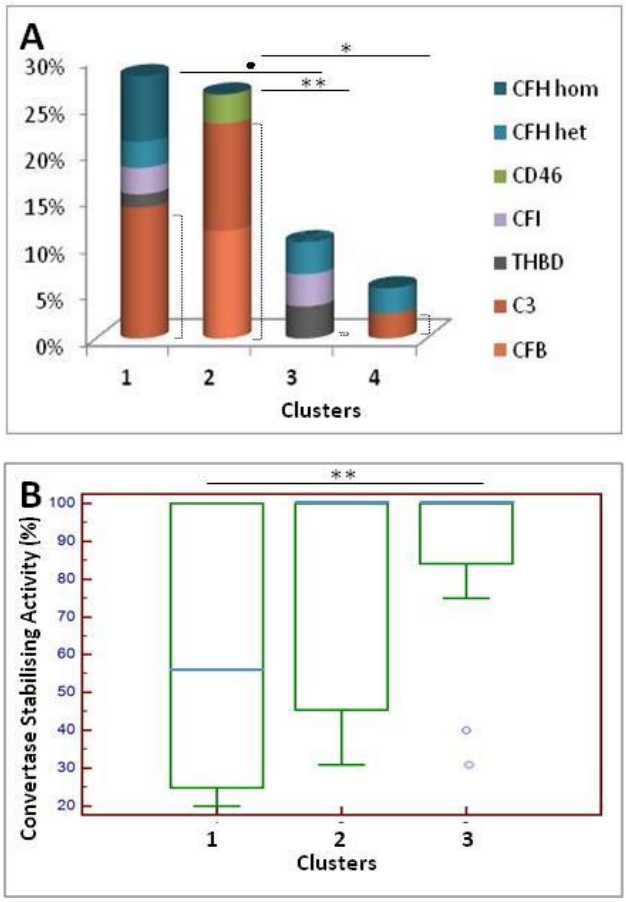


Figure 4.5. A) Distribution according to the newly identified clusters of the mutations. Mutations in *C3* (dark red) and *CFB* (light red) are significantly overrepresented in clusters 1 and 2 compared to clusters 3 and 4 (**p<0.01, *p<0.05, •p=0.059). **B)** Distribution according to the newly identified clusters of the C3NeF residual activity evaluated by hemolytic assay in C3NeF positive patients (**p<0.01). The central box represents the values from 25° to 75° percentile. The blue lines represent the medians. A line extends from the minimum to the maximum value, excluding "outsiders" (circles) defined as a value that is smaller than the 25° percentile minus 1.5 times the interquartile range, or larger than the 75° percentile plus 1.5 times the interquartile range.

4.5 Discussion

Here, in 163 patients with C3G or IC-MPGN, we first investigated whether the IF-based C3G/IC-MPGN classification fits with the underlying pathogenesis leading to different patterns of complement activation. We show that a significant number of IC-MPGN patients present alternative pathway dysregulation. To better understand the pathogenesis of the diseases, we explored the presence of relatively homogeneous subgroups through unsupervised cluster analysis based on histological, biochemical, genetic and clinical data. C3G/IC-MPGN patients could be divided into 4 clusters characterized by specific clinical phenotypes, IF and EM features and complement abnormalities. Finally we provide an easy-to-use algorithm to apply in clinical practice to ascribe patients into clusters.

The current C3G/IC-MPGN classification is based on two assumptions. First, in idiopathic IC-MPGN, an unidentified antigen induces immunoglobulin production, with immune-complex formation, subsequent glomerular deposition and complement activation through the classical pathway.[258] Second, in C3G, C3 deposits with scanty immunoglobulins implicate complement activation through the alternative (or the lectin) pathway.[248, 252] Based on these assumptions, C3G patients should show a high prevalence of alternative pathway abnormalities, while IC-MPGN should show low prevalence.[258] However, our study provides different evidence. The "C3-dominant" staining, which defines C3G, has low sensitivity for capturing patients with complement gene mutations and/or C3NeF, or patients with low serum C3 and normal C4 (Table 4.4). There is no difference in alternative pathway dysregulation prevalence even when we compare patients with isolated C3 staining to those with strong immunoglobulin/C1q staining.

Failure to observe a higher prevalence of alternative pathway abnormalities in C3G

compared to IC-MPGN could be attributed to the fact that a large number of idiopathic IC-MPGN cases carry alternative pathway abnormalities, which are not sufficient to cause the disease, but require the glomerular immunoglobulin deposition as a second "hit". Alternatively, some C3G cases without alternative pathway abnormalities may be secondary to “masked" immunoglobulin deposits, which are not detectable by standard immunofluorescence techniques.[341] In addition, in some IC-MPGN cases carrying alternative pathway abnormalities, the presence of immunoglobulins may be a non-specific finding caused by immunoglobulin entrapment secondary to proteinuria in areas of sclerosis or in podocytes.[85, 342] Independently of the explanation, our findings and previous studies document that the presence of significant immunoglobulin/C1q staining alone cannot be used to exclude alternative pathway dysregulation.[256, 339]

To investigate whether C3G/IC-MPGN patients can be subdivided in homogeneous groups, we used a statistical approach where the categorization criteria are not defined by the operator but arise from the data. Simplistically, patients who have many features in common are placed in the same group and those with many differences are placed further apart. The analyses show that patients without alternative pathway dysregulation (cluster 4) should be placed separately from those with such abnormalities (Table 4.5 and Figure 4.1). Interestingly, the latter patients are further subdivided into 3 clusters. Clusters 1 and 2, with very high SC5b-9 levels and a high prevalence of subendothelial deposits, are distinguished from cluster 3, characterized by high prevalence of intramembranous electron-dense deposits. Then, cluster 1 is distinguished from cluster 2, the latter presenting the strongest C1q and IgG staining, and the highest prevalence of nephrotic syndrome at onset. We show that these phenotypic-defined clusters can partially be explained by differences in mutation distribution in complement genes and in C3NeF efficiency to stabilize C3 convertase (Figures 4.4 and 4.5). Patients in clusters 1 and 2 have higher prevalence of mutations in *C3* and *CFB*, the two components of the

alternative pathway C3 and C5 convertases, compared to cluster 3. C3NeFs in cluster 3 stabilize the C3 convertase more efficiently than those in cluster 1. Remarkably, neither the distribution of the mutations, nor the C3NeF efficiency to stabilize the C3 convertase, were variables that were included in the cluster analysis. This further emphasizes the solidity of our findings.

Zhang and colleagues proposed that C3GN and DDD can be differentiated on the basis of different degrees of complement activation at C3 or C5 levels; dysregulation of the C5 convertase is greater in C3GN, while dysregulation of the C3 convertase is greater in DDD.[343] This model can be applied to clusters identified here. C5 convertase activity would be greater in clusters 1 and 2, as suggested by the higher SC5b-9 levels, leading to less electron-dense and more amorphous deposits all along the glomerular membrane layer, whereas C3 convertase activity would be greater in cluster 3, leading to denser midlayer deposits. This hypothesis is in line with findings here that *C3* and *CFB* mutations, which likely affect both the C3 convertase (C3bBb) and the C5 convertase (C3bBbC3b), are prevalent in clusters 1 and 2, and that C3NeFs in cluster 3 strongly stabilize the C3 convertase. We hypothesize that C3NeFs in clusters 1 and 2 may stabilize the C5 convertase more effectively than C3NeFs in cluster 3, similarly to previously described C3NeFs[326]. However, further functional studies are needed to address this issue. Finally, the combination in cluster 2 of alternative pathway abnormalities with C1q deposits, which are always associated with IgG and/or IgM deposits, strongly suggests a concomitant activation of both classical and alternative pathways.

At variance with clusters 1-3, cluster 4 is characterized by a low prevalence of complement gene mutations and/or C3NeF, and likely normal s-C3 and SC5b-9 (Table 4.5). Patients present later onset, more advanced interstitial and arteriolar lesions, and a

higher risk of ESRD. The latter finding is in line with previous data from our group showing a higher risk of ESRD in C3G/IC-MPGN patients without mutations or C3NeF.[339] The higher risk of ESRD in cluster 4 might be explained by the presence of peculiar underlying mechanisms of complement dysregulation resulting in renal damage. Bright C3 staining with a normal serum complement profile in this cluster implies local complement activation on glomerular cells or along the glomerular basement membrane,[85] leading to direct injury of glomerular cells by complement effector molecules. The fact that all 7 patients who developed thrombotic microangiopathy, an event associated with local complement activation on endothelial glomerular cell surface, fall into cluster 4, would support the above hypothesis. On the other hand, we would speculate that in patients from clusters 1-3, massive fluid phase complement activation results in accumulation of degradation products in the glomerulus, but at the same time the associated systemic complement consumption prevents direct complement-mediated injury. Alternatively, in some patients from cluster 4 the C3 glomerular deposits may be due to a non-specific inflammatory process secondary to a glomerulopathy not directly related to complement, as suggested by the finding that some MPGN patients carry mutations in podocytopathy-associated genes.[152]

Our approach has multiple advantages. We use a large cohort of C3G and IC-MPGN patients. We do not limit our observations to histologic features but integrate them with clinical features and markers of the underlying molecular mechanisms. We use a statistical approach, where the classification criteria are operator-independent. Patients within clusters share common phenotypes and underlying causes. The newly identified clusters are useful for predicting the risk of ESRD. Finally, clusters may also be useful for identifying the patients who would most benefit from anti-C5 treatment. In particular, patients with alternative pathway dysregulation and high SC5b-9, the main

features in clusters 1 and 2, may be more likely to respond to this therapy.[343, 344]

We also provide a simple algorithm for patient assignment into clusters. It combines complement features, like complement gene mutations, C3NeFand s-C3 ≤70mg/dl, with histologic features, like intramembranous electron-dense deposits and C1q deposits, and with age of onset. This algorithm may be useful both in basic and clinical research to better define the pathogenesis of C3G/IC-MPGN, and to predict the progression to ESRD.

Some limitations of the present study must be taken into account. Some abnormalities previously reported in association with IC-MPGN/C3G, such as *DGKE* and *CFHR* mutations and rearrangements, as well as autoantibodies against FH and FB, were not investigated. Other features relevant to C3G/IC-MPGN pathogenesis that could have increased the ability to identify clusters may have been missed. Additionally, this is a multicenter and retrospective study that may be affected by heterogeneity in the diagnostic process and treatment. Nonetheless, multicenter studies improve patient recruitment, increase statistical power and maximize the generalizability of results. Treatments might have modified disease presentation and course, affecting our results. Conversely, treatment responses may depend on clusters. Finally further future discovery of the disease causes will make the use of indirect evidence of alternative pathway dysregulation, such as low s-C3 levels, obsolete in the algorithm.

In conclusion, we identify clusters of C3G/IC-MPGN patients with distinct underlying mechanisms and phenotypes. These clusters improve the ability to define the risk of ESRD and may be useful to basic and clinical researchers. Finally, we provide an easy-to-use algorithm to identify clusters by integrating IF and EM features with the complement component profile and clinical features.

5. DESIGN OF A GENETIC TEST FOR IDIOPATHIC NEPHROTIC SYNDROME AND SEARCH FOR NOVEL GENES INVOLVED IN THE PATHOGENESIS OF SRNS

5.1 Introduction

Nephrotic syndrome (NS) is characterized by increased permeability of the glomerular barrier to plasma macromolecules resulting in heavy proteinuria that may lead to progressive renal function loss and exposes patients to the risk of life threatening infections, thromboembolic episodes, and dyslipidemia. NS affects 3-4 children every 100,000 and accounts for about 30% of primary forms of glomerulonephritis in adults.[345]

Glucocorticoids are first-line treatment and the rationale for immunosuppressive therapy relies on the evidence that immune alterations have been identified among affected patients.[107] Specifically, it appears that in some patients T cells promote the production of a circulating factor that alters the glomerular permeability of the filtration barrier.[108] Most patients respond to corticosteroid therapy with remission of proteinuria (steroid sensitive NS, SSNS), while about 20% of them do not (steroid resistant NS, SRNS). More than 40% of steroid-resistant cases develop end stage renal disease (ESRD) within 10 years from diagnosis.[345] Some of these cases might benefit from other immunosuppressive treatments, such as alkylating agents, calcineurin inhibitors or Rituximab, which however have significant toxicity profiles.[110]

Most SSNS patients show minimal change disease (MC) on renal biopsy, while among

patients with SRNS the more common histological findings are focal segmental glomerulosclerosis (FSGS), MC, and membranous nephropathy (MN) (Table 5.1). Other findings in SRNS may be membranoproliferative glomerulonephritis (MPGN), IgA nephropathy and amyloidosis. Recognition of the specific conditions underlying SRNS is important as they differ with regard to their treatment and outcome.

SRNS is thought to be caused by a primary defect in the glomerular filtration barrier. Along the past 20 years, genetic studies of familial cases of SRNS and genetically modified animal models have shown the crucial role of genetic background in the pathogenesis of the disease. These studies have led to the identification of the podocyte and its interactions with the glomerular basement membrane as a central player in the pathogenesis of SRNS. Alterations in several podocyte proteins have been recognized as the cause of SRNS.[144] Interestingly, defects in these proteins can cause isolated glomerulopathy, complex syndromes, where the nephropathy is associated with extra-renal manifestations, or both. There is evidence that mutations affecting proteins that make part of the podocyte slit diaphragm, such as nephrin (*NPHS1*), podocin (*NPHS2*), calcium channel Transient Receptor Potential cation Channel 6 (*TRPC6*), phospholipase C-epsilon 1 (*PLCE1*) and CD2 Associated Protein (*CD2AP*), cause SRNS.[145, 146] SRNS can also be caused by mutations in proteins involved in the organization of the podocyte cytoskeleton such as alpha-actinin-4 (*ACTN4*), inverted formin 2 (*INF2*), myosin 1E (*MYO1E*), Rho GDP dissociation inhibitor alpha (*ARHGDIA*), and RhoA-activated Rac1 GTPase-activating protein (*ARHGAP24*).[145, 146, 179, 180] Another group of genes associated with SRNS are those encoding proteins involved in focal adhesions that tether basal surface of the podocytes to the underlying glomerular basement membrane including integrins *ITGA3* and *ITGB4*, *EMP2*, *LAMB2*, and collagens *COL4A3*, *COL4A4* and *COL4A5*.[346] Mutations are also found in genes encoding transcription factors that regulate podocyte development and homeostasis including *WT1*, *PAX2* and

LMX1B.[110] The *PTPRO* gene encoding for the transmembrane podocyte protein Glepp1, which regulates the glomerular pressure/filtration rate, has also been associated to SRNS.[239] Moreover, mutations in genes encoding for proteins involved in the biosynthesis of CoQ10, such as *COQ2*, *COQ6*, *ADCK4* and *PDSS2*, have also been described.[214] Recently, mutations of nucleoporin genes, such as *NUP93*, *NUP107*, *NUP205* and *XPO5*, have been described to cause SRNS. Nucleoporins make part of the nuclear pore complex, a huge protein complex embedded in the nuclear envelope, which is involved in nucleocytoplasmic transport, nuclear framework, and gene regulation.[237, 238]

Noteworthy, *PAX2* and *LMX1B* are also associated with congenital anomalies of the kidney and the urinary tract (CAKUT). CAKUT is the most frequent cause of chronic kidney disease in childhood. It comprise a spectrum of malformations that occur at the level of the kidney (small kidneys, renal dysplasia, multicystic kidneys, horseshoe kidney, renal agenesis), collecting system (megaureter, hydronephrosis, duplex ureter), bladder (e.g. ureterocele, vesicoureteral reflux), or urethra (e.g. posterior urethral valves).[219] Proteinuria is also frequently observed.[219] CAKUT are associated with genes involved in the kidney development from the ureteric bud induction stages (*GDNF*) to ureteric bud branching (*PAX2*, *EYA1*, *SIX1*, *ROBO2*) to glomerular and podocyte differentiation (*PAX2*, *LMX1B*).[219]

Other molecular mechanisms have been identified in diseases characterized by nephrotic syndrome. As detailed in results of chapters 3 and 4 of this thesis, in patients affected by MPGN and C3 glomerulopathy mutations have been described in genes encoding proteins involved in the alternative pathway of complement system (*CFH*, *CFHR5*, *CD46*, *CFI*, *CFB*, *C3* and *THBD*).[81] Moreover, mutations in DGKE, the gene encoding for diacylglycerol kinase epsilon, have been found in MPGN, atypical

hemolytic uremic syndrome and SRNS.[152, 347, 348]

Membranous nephropathy (MN) is characterized by the deposition along the capillary walls of IgG auto-antibodies, commonly against the M-type phospholipase A_2 receptor. Although it is an immune-mediated disease, a genetic background has been demonstrated. Indeed, a strong association between *HLA-DQA1* and *PLA2R1* gene variants and idiopathic MN has been shown.[273]

IgA nephropathy (IgAN) is another disorder affecting kidney glomeruli. Its typical presentation is hypertension or mild oedema in young oligosymptomatic adults.[134] Hematuria and chronic kidney disease are the main characteristic, while NS is observed in 10% of patients.[278] Among patients with NS 48% show complete remission, 32% partial remission and 20% no response to treatment.[278]

IgAN is an immune-mediated disease caused by abnormal deposition of deficient O-linked glycans IgA1 antibodies. Nonetheless, multiple susceptibility loci for IgAN have been identified implicating defects in adaptive immunity (HLA genes and), innate immunity (*VAV3, DEFA*, *TNFSF*, *CARD9*, *ITGAM*, *ITGAX* and *HORMAD2*) and the alternative pathway of complement (*CFHR1*).[279, 282] Interestingly, the presence of components of the complement alternative pathway, but not of the classical pathway, in the kidney specimens and the increased susceptibility for IgAN in *CFHR1* deleted subjects underline the role of the complement alternative pathway.[282] This links IgAN to MPGN since the alternative pathway of complement is the milestone of the MPGN pathogenesis.[81, 349]

Table 5.1. Frequency of primary glomerulopathies.

Type of Primary Glomerulopathy	Frequency (%)		No remission (%)	ESRD (%)
	Childhood	Adults		
Minimal change nephropathy	80-95	10-30	10-35	<5
Focal segmental glomerulosclerosis	4-10	10-30	40-89	0-55
Membranoproliferative glomerulonephritis	4-7	4-15	95	40-90
Membranous nephropathy	2-4	23-30	50-67	30-50
Other (mesangioproliferative glomerulonephritis, IgA nephropathy etc.)	4-5	15-47	<90	20

Modified from Rugarli (MASSON, 1997),[350] McGrogan (NDT, 2011)[135] and eMedicine (emedicine.medscape.com).

Recurrence of proteinuria after renal transplantation is observed in approximately 30 to 50% of patients with NS and negatively impacts graft survival. Genetic forms rarely recur after transplantation,[351] providing further evidence that the disease originates from a primary glomerular injury. So far, abnormalities in known genes account for about 40% of familial and about 20% of sporadic cases with primary SRNS. Hence, other genes are probably involved in the pathogenesis of SRNS. Other loci have been linked to familial SRNS, even if the corresponding genes have not been identified yet, further supporting a highly genetic heterogeneity of this disease.[352]

Genetics have been recently experienced a dramatic change with the introduction of next generation sequencing (NGS), which allows rapid and low cost sequencing of genomic regions, which would require months of work with classic Sanger sequencing. NGS has been recently applied to the discovery of new causal variants and candidate genes for various Mendelian disorders including SRNS.[179] In addition, de novo variants were also identified for sporadic cases, which would have not been possible without NGS.[353] However its application to clinical diagnosis is still in its infancy, since the accuracy of mutation detection should be improved in order to become a viable option.[354-356]

Understanding of the genetic factors that contribute to the development of SRNS would provide invaluable information that could be used for early diagnosis, more precise treatment decisions and genetic counselling in patients and families. Creating a NGS panel that comprises genes associated with different glomerulonephritides could be helpful to reach diagnosis in a part of NS patients, in whom renal biopsy is often contraindicated (e.g. due to coagulation disorders). It may also improve the accuracy of histologic diagnosis and contribute to more targeted therapies. Here is presented the design and the set-up of a NGS panel for glomerular disorders and nephrotic syndrome. Its application in a SRNS cohort resolved a significant percentage of cases and suggests a role of the CAKUT-associated genes in the pathogenesis of SRNS. Finally, research of novel SRNS-genes identified *EPB41L5* as a candidate gene for SRNS.

5.2 Specific aims

1. To create a NGS based genetic test containing genes associated with the most common primary proteinuric glomerular disorders.

2. To search for mutations in known SRNS-associated genes in patients with the sporadic and the familial form of SRNS, in order to evaluate the prevalence of mutations in patients.

3. To evaluate whether SRNS patients carry mutations in genes associated with other primary proteinuric glomerular disorders.

4. To search for new gene/s involved in the pathogenesis of SRNS by screening selected candidate genes in the entire cohort.

5. To search for new gene/s involved in the pathogenesis of SRNS in familial cases without mutations in the known genes using autozygosity and linkage studies all along the genome and whole exome studies.

5.2 Methods

5.2.1 Patients

Patients have been recruited by the Registry of Steroid Resistant Nephrotic Syndrome established at the Clinical Research Centre for Rare Diseases *Aldo e Cele Daccò* (CRCRD). Clinical data for all patients have been collected either by direct interview or by examining the clinical files sent by external nephrology units. All participants provided informed written consent. The study was approved by the Ethics Committee of the Azienda Sanitaria Locale of Bergamo (Italy).

5.2.2 Search strategy and selection criteria to identify genes associated with kidney glomerular disorders and nephrotic syndrome

We searched PubMed, Google Scholar and the OMIM database between April, 2015, and July 31, 2015, for the following search terms:

"nephrotic syndrome", "steroid resistant nephrotic syndrome", "focal segmental glomerulosclerosis", "FSGS", "minimal change disease", "minimal change nephrotic syndrome", "congenital anomalies of the kidney and the urinary tract", "CAKUT", "membranoproliferative glomerulonephritis", "mesangiocapillary glomerulonephritis", "mesangioproliferative glomerulonephritis", "dense deposit disease", "C3 nephropathy", "C3 glomerulonephritis", "C3 glomerulopathy", "IgA-nephropathy", "IgA glomerulonephritis", "IgA nephritis", "membranous nephropathy", "membranous glomerulonephritis", "C1q nephropathy", and "fibrillary glomerulonephritis".

The above terms were searched at the OMIM database as is. At PubMed each one was combined with the following terms: "... AND (gene OR genet* OR mutat*)", while at Google Scholar was combined with: "... AND (gene OR genetic OR mutation)"

Restriction for English language was applied.

Original research and reviews on the genetic causes of the above disorders were selected. The references of the reviews were also evaluated to identify more genes. Only genes causing glomerular disorders and/or nephropathy in human beings were selected. Among the CAKUT-associated genes, only those associated with proteinuria were selected.

5.2.3 Selection of the candidate genes

To select the candidate genes different databases and software were used:

Gene expression profiling by next generation sequencing (RNA-seq) in normal tissues were obtained from the RNA Seq Atlas (http://medicalgenomics.org/rna_seq_atlas).[357] To evaluate a prevalent expression in kidney, for each gene the median tissue expression was subtracted from the kidney expression value.

The Mouse Genome Informatics (MGI) database that enables access to spontaneous, induced, and genetically-engineered gene mutations and their strain-specific phenotypes and link them to the human homologues. Phenotypic terms present in MGI were obtained for all available genes (http://www.informatics.jax.org/) using the Batch query search engine. The phenotypic terms present in the Mouse Genome Informatics (MGI) database were divided in 3 subgroups: terms related to glomerulopathy and nephrotic syndrome (n=99, Table 5.2), terms related to other kidney and urinary tract abnormalities (n=263) and terms not related to kidney (n=7,796).

Disease-causing genes are not expected to tolerate mutations. Two indexes gene mutation tolerance were used: the Gene-level Integrated Metric of negative Selection, GIMS (data were obtained from http://glom.sph.umich.edu/GIMS/),[358] and the Residual

Variation Intolerance Score, RVIS (data were obtained from http://genic-intolerance.org/).[359] For GIMS a lower number indicates stronger negative selection. For RVIS a gene with a score close to 0 has less functional variants and is defined as "intolerant".

The Podocyte Prediction Score, a score for defining the importance of a gene for the podocyte, was obtained through Support Vector Machine learning using genome-wide expression data of normal murine podocytes, mesangial, and endothelial cells.[359, 360]

Prioritization was performed using 3 different tools (Endeavour, ToppGene e GeneDistiller). For technical limitations during the analyses with ToppGene and GeneDistiller, the Test gene set was divided in 3 parts. The following SRNS, MPGN and MN associated genes were used as training set.

- SRNS: NPHS2, NPHS1, PLCE1, LAMB2, PTPRO, MYO1E, WT1, TRPC6, CD2AP, INF2, ARHGDIA, ARHGAP24, ACTN4, PAX2, ADCK4, EMP2, TTC21B, CRB2, CUBN, SMARCAL1, COQ2, COQ6, LMX1B, COL4A4, COL4A3, COL4A5, MYH9, ANLN

- MPGN: CFB, CFH, C3, CFI, CD46, THBD, DGKE, CFHR5

- MN: *PLA2R1*

Table 5.2. Phenotypic terms present in the Mouse Genome Informatics database that were considered as related to glomerulopathy and nephrotic syndrome.

glomerulonephritis	delayed kidney development
abnormal renal glomerulus basement membrane morphology	abnormal kidney cortex morphology
increased renal glomerulus basement membrane thickness	abnormal kidney development
abnormal renal glomerulus morphology	kidney failure
increased renal glomerulus apoptosis	abnormal basement membrane morphology
renal glomerulus hypertrophy	impaired basement membrane formation
abnormal renal filtration rate	albuminuria
cortical renal glomerulopathies	erythruria
renal fibrosis	hemoglobinuria
renal glomerulus atrophy	erythrocyturia
renal necrosis	decreased circulating serum albumin level
decreased renal plasma flow rate	albuminuria
abnormal renal glomerular filtration rate	abnormal circulating serum albumin level
decreased renal glomerular filtration rate	increased glomerular capsule space
renal glomerulus fibrosis	glomerular crescent
absent renal glomerulus	glomerulosclerosis
renal glomerular synechia	abnormal glomerular mesangium morphology
increased renal glomerulus lobularity	abnormal glomerular capillary endothelium morphology
abnormal renal corpuscle morphology	abnormal glomerular capillary morphology
abnormal renal glomerulus basement membrane thickness	abnormal glomerular endothelium fenestra morphology
abnormal renal plasma flow rate	dilated glomerular capillary
renal glomerulus lipidosis	absent tubuloglomerular feedback response
decreased renal glomerulus basement membrane thickness	abnormal juxtaglomerular apparatus morphology
abnormal renal protein reabsorption	abnormal juxtaglomerular cell morphology
abnormal podocyte morphology	absent glomerular endothelium fenestra
podocyte foot process effacement	abnormal glomerular filtration barrier function
abnormal podocyte foot process morphology	abnormal glomerular capsule parietal layer morphology
fused podocyte foot processes	decreased glomerular capsule space
absent podocytes	juxtaglomerular cell hyperplasia
absent podocyte slit diaphragm	decreased tubuloglomerular feedback response
podocyte hypertrophy	decreased glomerular capillary number
podocyte microvillus transformation	abnormal tubuloglomerular feedback response
increased podocyte apoptosis	glomerulus hemorrhage
absent podocyte foot process	glomerular capillary congestion
decreased podocyte number	glomerular capillary thrombosis
abnormal podocyte slit diaphragm morphology	abnormal glomerular capsule space morphology
detached podocyte	abnormal extraglomerular mesangial cell morphology
abnormal podocyte slit junction morphology	abnormal glomerular filtration barrier morphology
abnormal podocyte adhesion	expanded mesangial matrix
abnormal podocyte motility	mesangial cell hyperplasia
abnormal podocyte polarity	mesangial cell interposition
decreased urine urea nitrogen level	absent mesangial cell
increased urine protein level	mesangiolysis
decreased urine creatinine level	abnormal mesangial cell morphology
abnormal urine protein level	abnormal mesangial matrix morphology
increased urine microalbumin level	mesangial cell hypoplasia
increased urine major urinary protein level	increased circulating creatinine level
increased urine microglobulin level	decreased creatinine clearance
increased urine beta2-microglobulin level	increased blood urea nitrogen level
hematuria	

Manual evaluation of the genes emerged from the above databases and software was performed based on the scientific literature obtained through PubMed and Google Scholar using the following search terms:

(‘gene name’ OR ‘alternative gene name 1’ OR ... OR ‘alternative gene name *n*’ OR ‘gene symbol’ OR ‘alternative gene symbol 1’ OR ... OR ‘alternative symbol *n*’)

AND (nephrotic OR proteinuria OR albuminuria OR glomerular OR glomerulopathy OR podocyte OR “slit diaphragm” OR “focal segmental glomerulosclerosis” OR “minimal change disease” OR “membranoproliferative glomerulonephritis” OR “membranous nephropathy”)

5.2.4 Ion Torrent PGM DNA Sequencing

Genomic DNA was extracted from peripheral blood using the NucleoSpin® Blood kit (Macherey-Nagel). Analysis of DNA regions from 112 genes was performed by highly multiplex PCR using the Ion AmpliSeq™ Library Kit 2.0 and the Ion AmpliSeq Custom WG_IAD76560 panel (subdivided into two pools), designed by the White-Glove service of Life Technologies. After clonal amplification using the Ion PGM™ Template HiQ Kit, sequencing was performed on Ion PGM Sequencer (Life Technologies) using a Ion 318 chip v2. Variants were identified using TorrentSuite Software 5.2 which includes the alignment plugin, coverage plugin, and variantCaller plugin. In addition, for coverage analysis a home-made R script was created and used to manage the Samtools and Bedtools software. Regions with a depth of amplicon coverage ≥10x reads were considered appropriately analyzed.

Likely pathogenic variants were confirmed with Sanger sequencing.

5.2.5 Whole Exome Sequencing

Whole Exome Sequencing (WES) with the Illumina platform for the pedigrees of SN091 and SN140 was outsourced at the Department of Biotechnology, University of Verona, Italy. The library was enriched for exonic regions in the human genome using the TruSeq Exome Enrichment Kit (Illumina) and subsequently sequenced on a HiSeq 2000 Illumina sequencer.

Whole Exome Sequencing (WES) with the SOLiD platform for patient IR35 was outsourced at the Genomnia laboratories (Italy). The library was enriched for exonic regions in the human genome using the SureSelect Human All Exon kit (Agilent). Then it was analyzed on a SOLiD Sequencer (Applied Biosystems).

5.2.6 Variant annotation

Functional annotation of the variants was performed using the ANNOVAR software (June 17th 2015 release).

The allele frequencies of the variants were obtained from the following public databases: "1000 Genomes Project" (v5a.20130502 release), "NHLBI GO Exome Sequencing Project" (v.0.0.30, Nov. 3, 2014, release) and Exome Aggregate Consortium (ExAC v0.3).

To perform the comparisons between patients and controls, 404 unphenotyped subjects of European non-Finnish origin from the 1000 Genomes Project (CEU, TSI, GBR and IBS subpopulations) for whom individual genotypes were available was used as a control group. To avoid bias in assessing allele frequency, the independent 6500 Exome Sequencing Project (ESP) was considered as reference to assess allelic frequencies for case-control comparison. Missense variants, insertions/deletions in the coding regions,

splicing variants affecting the first two nucleotides flanking exons, or variants lying within the Kozak consensus sequence, were considered functional. Common variants that exceeding an allelic frequency of 0.01 in any subpopulation from all databases were filtered out.

The neutral or damaging effect of a variant on the protein function was predicted using the following software and algorithms available through the Annovar software: SIFT, Polyphen2, Mutation Taster, Mutation Assessor, LRT, FATHMM, CADD v1.3 (damaging if ≥10), GERP (damaging if ≥2) e SiPhy (damaging if ≥10). Since the CADD score of insertion and deletions is not precomputed in Annovar, it was calculated directly at CADD website (http://cadd.gs.washington.edu/score).

Likely pathogenic variants in known genes were defined on the basis of inheritance, type of mutation, allele frequency and CADD pathogenic score. The presence of at least one allele was required for genes with autosomal dominant or X-linked inheritance, and the presence of one homozygous or two heterozygous alleles for genes with autosomal recessive inheritance. Nonsense variants (due to stop codons) and frameshift variants, which lead to truncated protein products and thereby impair protein function, were considered pathogenic, if the inheritance criterion was fulfilled. Missense and in frame insertions or deletions were evaluated as follows: 1) allelic frequency ≤0.0001 for genes with autosomal dominant or X-linked inheritance, and ≤0.001 for genes with autosomal recessive inheritance; 2) CADD pathogenicity score ≥15 (a more stringent CADD score cut-off was used compared to the one in the 10 algorithm prediction).

In the candidate genes, the variants with allelic frequency ≤0.001 and CADD score ≥15 were defined as possibly pathogenic variants.

5.2.7 Genome-wide SNP genotyping

Genotyping of about 1 million SNPs all along the genome was performed using the Genome-Wide Human SNP Array 6.0 (Affymetrix) in collaboration with the Department of Biomedical Sciences, University of Brescia.

5.2.8 Autozygosity mapping software

A way to identify regions where both alleles are inherited by a common ancestor of the parents (identity by descent) is the autozygosity mapping. For this purpose, based on the Broman and Weber approach,[361] the autozygosity LOD score is calculated for a genomic variant *i* using the following formula:

$$LOD(j,k) = \sum_{i=j}^{k} \log_{10}\left[\frac{\Pr\,(g_i|\text{autozygous at i})}{\Pr\,(g_i|\text{not autozygous at i})}\right]$$

where the probabilities of the observed genotype at a marker, given the autozygosity status of the surrounding segment, are obtained through the formulas in the following table:

Observed Genotype	Probability that the segment is		Ratio
	Autozygous	Not Autozygous	
AA	$(1-\varepsilon)p_A + \varepsilon p_A^2$	p_A^2	$(1-\varepsilon)/p_A + \varepsilon$
AB	$2\varepsilon p_A p_B^2$	$2p_A p_B^2$	ε

ε denotes the combined rate of genotyping errors and mutations.

The autozygosity strength of a region is given by the sum of the single variant autozygosity LOD scores for the given region. In order to evaluate the autozygosity

score for all the SNPs along the whole genome (1 million SNPs), script in Python 2.6 was written. Since the SNP distribution is not homogeneous all over the genome and there is an inflation of the autozygosity score in the regions with high SNP density, a second algorithm was created where I take only 1 SNP (the most informative, in the mean of less probable to occur) for each region of 0.1 cM (Density Adjusted Algorithm).

5.2.9 Linkage studies

Furthermore, genotypes emerged from the Human SNP Array 6.0 (Affymetrix) were used to perform classic whole-genome linkage analysis with the Merlin software. The parametric analysis was performed using an autosomal recessive model for rare disease (allele frequency 0.001) for pedigrees with consanguinity of parents and autosomal dominant model for rare disease (allele frequency 0.00001) for the pedigree with suspected autosomal dominant inheritance.

5.2.10 Statistical analyses

χ^2-test and Fisher exact test were used for categorical variables when appropriate. Statistical analyses were performed with the R software (https://cran.r-project.org/). P-values <0.05 were considered statistically significant.

5.3 Results

5.3.1 To create a diagnostic NGS-based panel for SRNS

5.3.1.1 Panel design

The first aim of this study was to create a genetic test based on next-generation sequencing approach that could allow the detection of SNV and indel mutations, as well as susceptibility SNPs, known to be associated with kidney glomerular disorders and nephrotic syndrome. To achieve this goal, the English-language literature was reviewed. There were 79 genes with mutations repeatedly found in glomerular proteinuria and nephrotic syndrome (Table 5.3). There were also 28 non-coding SNPs in 19 genes that predispose to SRNS (n=7), MPGN (n=2), MN (n=2), IgAN (n=15) and SSNS (n=1). Overall, mutations or polymorphic risk alleles in 92 different genes were identified to be associated with proteinuric nephropathy.

Altogether the genes included were associated to SRNS, CAKUT, MPGN, MN and IgAN. Furthermore, there were genes, which genetic abnormalities cause glomerulopathy with extra-renal manifestations including amyloidosis, Alport syndrome, Nail-patella syndrome, Renal-coloboma syndrome, Fabry disease, Denys-Drash syndrome, Epstein/Fechtner syndrome, Branchio-oto-renal syndrome, coenzyme Q10 deficiency, immune-osseous dysplasia of Schimke, glycosylation disorder type Ik, nephropathy-epidermolysis bullosa-deafness syndrome, nephropathy-epidermolysis bullosa-pneumopathy syndrome, Townes-Brocks syndrome, Kallman syndrome, and hypoparathyroidism-deafness-renal dysplasia syndrome.

Two hundred forty four genes were obtained by interrogating different databases and software. After manual evaluation 21 genes were selected as candidates to cause nephrotic syndrome based on high expression in human kidney, kidney glomerular

abnormalities in animal models, protein-protein interaction with genes associated with SRNS, MPGN or MN, and low tolerance to functional genetic variation (Table 5.3).

Table 5.3. List of genes included in the proteinuric glomerulopathy panel.

Disease	Gene/SNP name	Gene n°	ncSNP n°
SRNS	*ACTN4*, *ADCK4*, *ARHGAP24*, *ARHGDIA*, *CD2AP*, *CRB2*, *CUBN*, *EMP2*, *INF2*, *MYO1E* (gene, rs113915226), *NPHS1* (gene, rs73928332), *NPHS2*, *PLCE1*, *PTPRO*, *TRPC6* (gene, rs41302375), *DGKE*, *TTC21B*, *APOL1* (rs73885319, rs60910145, rs71785313), *GPC5* (rs16946160)	17	7
SRNS syndromic	*WT1*, *COL4A3*, *COL4A4*, *COL4A5*, *MYH9*, *LMX1B*, *ALG1*, *CD151*, *COL4A6*, *ITGB4*, *SCARB2*, *ZMPSTE24*, *ANLN*, *COQ2*, *COQ6*, *ITGA3*, *LAMB2*, *PDSS2*, *SMARCAL1*	19	0
CAKUT	*PAX2* (gene, rs4523631), *EYA1*, *GATA3*, *KAL1*, *SALL1*, *SIX1*, *SIX5*, *SOX17*, *BMP7*, *CHD1L*, *DLX5*, *FREM2*, *HPSE2*, *RET*, *ROBO2*, *SIX2*, *UPK3A*, *HNF1B*, *BMP4*, *FOXC2*	20	1
MPGN/C3G	*C3*, *CD46* (gene, rs2796268), *CFB*, *CFH* (gene, rs3753394), *CFHR5*, *CFI*, *THBD*, *C1QA*, *CFP*	9	2
MN	*PLA2R1* (gene, rs4664308), *PRKCD*, *HLA-DQA1* (rs2187668)	2	2
IgAN	*VAV3* (rs17019602), *HLA-DR/HLA-DQ* (rs7763262, rs9275224, rs2856717, rs9275596), *PSMB8* (rs9357155), *TAP2-PSMB9* (rs2071543), *HLA-DP* (rs1883414), *DEFA* (rs2738048, rs10086568), CARD9 (rs4077515), *ITGAM-ITGAX* (rs11150612, rs11574637), *TNFSF3* (rs3803800), *HORMAD2* (rs2412971), *CFHR1*	1	15
Amyloidosis	*FGA*, *LYZ*, *APOA1*, *APOA2*, *APOE*, *GSN*	6	0
Metabol. dis.	*FN1*, *GLA*, *CTNS*, *LCAT*	4	0
SSNS	*HLA-DQA1* (rs1129740)	0	1
Candidates	*KIRREL*, *APCS*, *ILK*, *MIR192*, *MIR193A*, *DDN*, *TJP1*, *CRIM1*, *NCK1*, *NCK2*, *TCF21*, *IQGAP1*, *EPB41L5*, *FAT1*, *PDLIM2*, *CTGF*, *PODXL*, *FYN*, *TRIM3*, *TINAG*, *CD59*	21	0

ncSNP: non-coding Single Nucleotide Polymorphism; SRNS: Steroid-resistant nephrotic syndrome; CAKUT: Congenital Abnormalities of the Kidney and Urinary Tract; MN: Membranous nephropathy; IgAN: IgA nephropathy; MPGN: Membranoproliferative glomerulonephritis; C3G: C3 glomerulopathy; SSNS: Steroid-sensitive nephrotic syndrome. Metabol. dis.: metabolic disorders.

A custom library with 1881 primer pairs divided in two pools was designed to sequence the 1314 coding exons, their flanking regions, and the 29 non-coding SNPs. Altogether, the design covered 315,783 out of 316,167 bp (99.88%) of all genes (coding exons ± 5 bp of flanking intronic regions) and SNPs. The theoretical coverage was 99.94% (274,738 out of 274,559 bp) for the known genes and 99.51% (41,429 out of 41,224 bp) for candidate genes.

The performances of the panel were evaluated in 32 patients analyzed in 4 runs (8 patients/run) on Ion 318 Chips v2 using the Hi-Q chemistry both for template preparation and sequencing. A mean depth of 181x (±42) was obtained (Table 5.4). A coverage of ≥10x and ≥20x was reached in 98.8% and 98.4% of coding bases (±2bp of intronic regions), respectively. There were 1845 (98.1%) and 1823 (96.9%) amplicons showing a coverage of at least 10x and 20x all along the coding regions, respectively.

Table 5.4. Performance of the sequencing.

	Mean (±SD)
Mapped Reads	610,613 ± 114,768
On Target (%)	65.2 ± 6.5
Uniformity (%)	94.5 ± 1.6
Mean Depth	181.0 ± 42.4
% bases Cov≥10x	98.8 ± 0.2
% bases Cov≥20x	98.4 ± 0.2
N amplicons with coding region Cov≥10x	1845.2 ± 7.4
N amplicons with coding region Cov≥20x	1823.4 ± 19.5
% amplicons with coding region Cov≥10x	98.1 ± 0.4
% amplicons with coding region Cov≥20x	96.9 ± 1.0

Cov: coverage.

The variant identification using the Torrent Suite Variant Caller with the standard parameters showed a sensitivity of 98.2% and a false positive error rate of 5.8%. A new strategy was set up that increased the sensitivity (99.6%) without compromising the false positive error rate (5.4%). The strategy is extensively described in chapter 6.

5.3.1.2 Clinical and epidemiological features of the patients

One hundred sixty-eight patients affected by Steroid-Resistant Nephrotic Syndrome (SRNS) were recruited through the SRNS Registry of the CRCRD. Patients of 14 families that were previously found to carry mutations by Sanger sequencing (5 in *WT1*, 4 in *PAX2*, 2 in *CD2AP*, 2 in *MYO1E* and 1 in *PLCE1*) were not included. Seventy-nine out of 168 were sporadic cases, with no other family members affected by SRNS or other type of nephropathy, whereas 89 were familial cases belonging to 34 different families. The median age of onset of the patients in this cohort was 20.5 years (interquartile range 8-33 years). The distribution of the age of onset ranged from 0 to 71 years and showed a mode between 16 and 20 years of age (Figure 5.1).

Kidney biopsy showed a focal segmental glomerulosclerosis (FSGS) pattern of injury in 71% of the patients, minimal change disease (MC) in 20%, shifting from MC at the first biopsy to FSGS at a following biopsy in 3%, diffuse mesangial sclerosis in 1% and not conclusive findings in 4%, while biopsy was not available in 3% of SRNS patients.

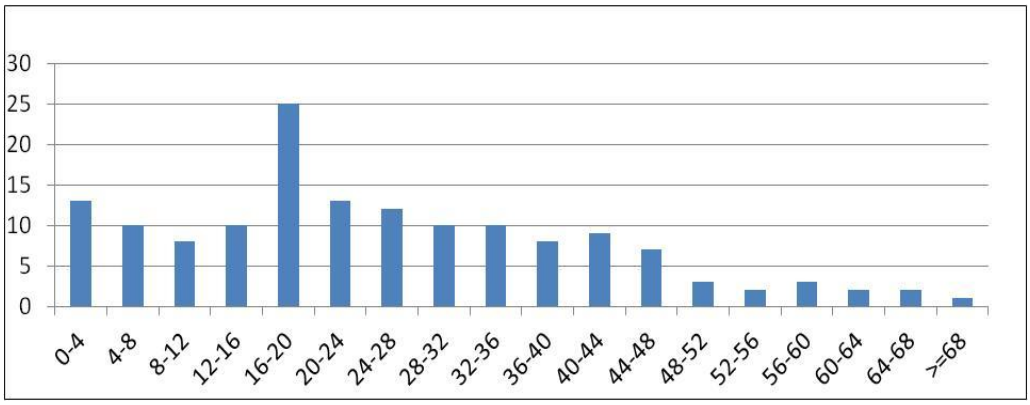


Figure 5.1. Distribution of the patient age of onset in the SRNS cohort.

About 27% of patients, who initially received the diagnosis of isolated SRNS, presented one or more extra-renal manifestations on in-depth evaluation of clinical records. Despite there was no statistically significant difference between familial and sporadic cases, familial cases had more likely ear and eye involvement suggesting, that Alport or Alport-like syndrome could be underdiagnosed (Figure 5.2).

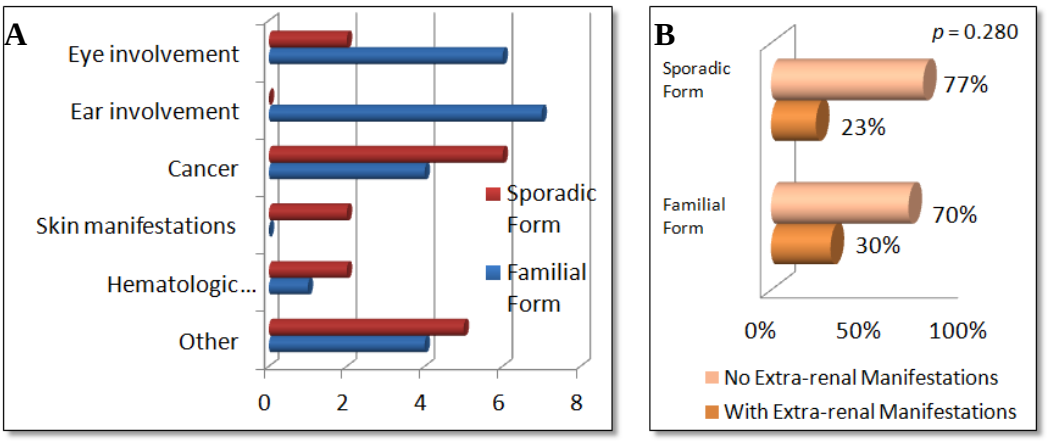


Figure 5.2. Panel A: Description of the different extra-renal manifestations observed in the SRNS patients recruited in the study. **Panel B**: Comparison of the extra-renal manifestation distribution between sporadic and familial cases (Pearson chi-squared test was performed).

5.3.1.3 Screening of a SRNS cohort with the new NGS panel

The presence of likely pathogenic (LP) variants in glomerulopathy-associated genes was investigated through the NGS panel in 111 unrelated SRNS patients. We considered only missense or in frame indel variants with Minor Allelic Frequency (MAF) ≤0.0001 and Combined Annotation Dependent Depletion (CADD) pathogenic score ≥15, as well as all nonsense and frameshift variants, since they lead to truncated proteins. For genes with autosomal dominant or X-linked inheritance, patients were considered carrying LP variants in case of at least one variant. For genes with autosomal recessive inheritance, patients were considered carriers of LP variants in case of one homozygous variant or two heterozygous variants.

LP variants in SRNS-associated genes were detected in 26 (23%) out of 111 unrelated patients (Table 5.5). The most frequently involved genes were those encoding for the collagen IV chains α3 (6 heterozygous patients), α4 (4 heterozygous patients) and α5 (2 hemizygous patients and 2 heterozygous patients). LP variants were also found in *INF2* (3 heterozygous patients), *NPHS1* (2 compound heterozygous patients), *PAX2* (2 heterozygous patients), *ARHGAP24* (1 heterozygous patient), *COL4A6* (1 heterozygous patient), *FAT1* (1 compound heterozygous patient), *MYH9* (1 heterozygous patient) and *TRPC6* (1 heterozygous patient).

Among the different LP variants in *COL4A3*, *COL4A4* and *COL4A5*, 10 are located in the corresponding triple helix domains (4 in *COL4A3*, 2 in *COL4A4* and 4 in *COL4A5*). Six of these variants cause the replacement of one Glycine in the (Gly-X-Y)n repeat of the corresponding collagen triple helixes (3 in *COL4A3*, 1 in *COL4A4* and 2 in *COL4A5*). Four LP variants are located in the corresponding globular NC1 domains (2 in *COL4A3* and 2 in *COL4A4*). In *COL4A4* the C1683Y variant affect the Cysteine at position 1683, which forms a disulfide bond with Cys-1588 or with Cys-1622 residues.

The *COL4A4* R1682W is placed next to the previous Cys-1683 residue. Also the *COL4A3* Q1663E is placed next to a disulfide bond site, the Cys-1662 that forms a disulfide bond with the Cys-1570 or with the Cys-1604 residues.

The G73S, R218Q and E220K variants affect evolutionary conserved residues of the GBD/FH3 domain of INF2. These mutations have been previously described in SRNS patients.[186, 188]

Regarding the LP variants found in the *NPHS1* gene, all of them are located at the extra-cellular C2-type Ig-like domains that allow nephrin-nephrin interactions at the slit diaphragm. The S350P, a previously described SRNS-causing mutation, lies in the Ig-like C2-type domain 4.[362] The *NPHS1* T172del, leading to the deletion of a Threonine, the S510Y, leading to a Serine to Proline substitution, and the S937N, leading to a Serine to Asparagine substitution, have not been previously described in SRNS patients and lie in the Ig-like C2-type domains 1, 5 and 8, respectively.

The *PAX2* Y185* variant leads to the formation of a truncated protein lacking the last 212 amino acids, including the octapeptide domain, the homeodomain and the transactivation domain. The PAX2 F10S variant leads to the substitution of a highly conserved Phenylalanine by a Serine, whose effect on protein function is unknown.

The *ARHGAP24* A704fs variant results in the formation of a truncated protein lacking the last 45 amino acids and thus loosing part of the coiled-coil domain.

The *MYH9* V1837G is a novel variant that affects a highly conserved amino acid located in a putative coiled-coil domain at the C-terminal of the protein.

The *TRPC6* R895C affects a residue located at a putative coiled-coil sequence in the C-terminal cytoplasmic domain. It is a known SRNS causing mutation leading to channel over-activation.[162] The *COL4A6* N1448D is a located at the triple helix domain and was

not previously reported.

A compound heterozygous patient was identified carrying the A131E and the C4113Y variants in the *FAT1* gene. The A131E leads to the substitution of an Alanine by a Glutamate at position 131, in the first Cadherin domain (Figure 5.3). This variant is not present in public databases; it has a CADD score of 15 and is predicted as damaging by 6 out of the 10 used algorithms. The C4113Y is a Cysteine to Tyrosine substitution at amino acid 4113, located in the fourth EGF-like domain. Also this variant is not present in public databases; it has a CADD score of 27 and is predicted as damaging by all 10 algorithms.

Altogether the prevalence of mutations in the above SRNS-associated genes was higher in patients compared to controls (0.108 *vs.*0.015, OR=8.04, p=3.4x10^{-5}; Table 5.8).

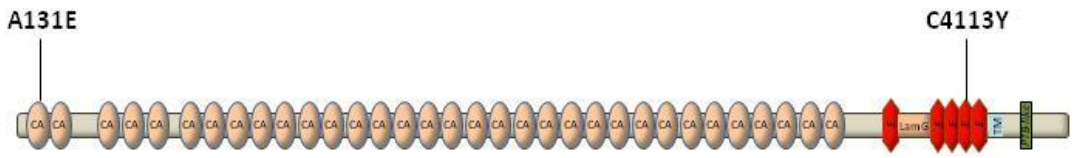


Figure 5.3. Graphical representation of the FAT1 protein. In its extracellular region contains 33 cadherin domains (CA), a laminin G domain (LamG) and 5 epidermal growth factor (EGF)-like repeat domains (E). It also contains a transmembrane region (TM) and a C-terminal cytoplasmic domain containing a PTB-like motif (PTB-like) with a PDZ-binding motif. Vertical lines represent the two LP variants identified in patient SN098.

Interestingly, I found SRNS patients carrying LP variants in genes that were previously associated with other disorders including CAKUT and amyloidois/metabolic disorders

(Tables 5.6 and 5.7). Nine patients carried LP variants in CAKUT-associated genes. The LP variants were found in *CHD1L* (3 heterozygous patients), *RET* (2 heterozygous patients), *FREM2* (1 compound heterozygous patient), *FOXC2* (1 heterozygous patient), *HNF1B* (heterozygous patient), *ROBO2* (1 heterozygous patient) and *SALL1* (1 heterozygous patient). All but *FREM2* were characterized by autosomal dominant inheritance.

The most frequently affected CAKUT-associated gene was *CHD1L*. Three LP variant were found, all of them leading to formation of truncated proteins (Q51*, K365fs and R410*). Notably, Q51* and R410* were already reported in public databases with an allele frequency of 0.0016 and 0.0002 in at least one subpopulation, respectively. The prevalence of LP variants in *CHD1L* did not differ significantly between the patients and 404 ethnically matched controls of the 1000 Genomes Project (Table 5.8); however, if the more stringent criterion of nonsense/frameshift mutations was adopted, the difference became statistically significant (3/111 *vs.* 0/404, p=0.0098).

There were two patients carrying LP variants in the *RET* gene. Both affect a highly conserved Lysine (K994N and K1003E) within the protein kinase domain.

In the *HNF1B* gene a novel A503V LP variant was identified that lies within the QSP-rich domain required for the HNF1B transactivation.[363]

The *ROBO2* M189K is a novel LP variant that occurs in the C2-type Ig-like domain 2 of the protein.

Two additional novel LP variants were found in *SALL1* (N1090S) and in *FOXC2* (Y25H). These substitutions affect amino acid whose function is unknown.

Finally, a compound heterozygous patient was identified carrying the *FREM2* R217C and A337S variants, both located at the chondroitin sulfate proteoglycan repeat 1.

The overall prevalence of mutations in the above CAKUT genes was also higher in patients compared to controls (0.081 *vs.* 0.022, R=3.87, *p*=0.006; Table 5.8).

Four patients carried LP variants in the amyloidois/metabolic disorder group. One patient carried the *FGA* L673P and another the Y821C novel variants, both located at Fibrinogen C-terminal domain. The prevalence of *FGA* LP variant was higher in patients compared to controls (0.018 *vs.* 0, *p*=0.046, Table 5.8).

One patient had two heterozygous *FN1* LP variants (S727A and V728E). Both are novel and lie within the Fibronectin type-III 2 domain. Finally one patient carried the G1026E variant, which is located within the Fibronectin type-III 5 domain of FN1. Mutations in FN1 gene cause glomerulopathy with fibronectin deposits. Both patients with FN1 LP variants described here had pathologic diagnosis of FSGS but the deposits of fibronectin were not investigated.

Altogether, LP variants were detected in 33 (30%) out of 111 index patients (Tables 5.5 and 5.7). There were 26 patients that carried LP variants in a single gene. Interestingly 7 patients carried LP variants in more than one gene. Patient SN121 carried the heterozygous G213E variant in *COL4A3* gene and the heterozygous C1683Y variant in *COL4A4* gene. Similarly to previous described patients with digenic COL4A3/COL4A4 inheritance, patient SN121 had an earlier disease onset and a more severe disease course. In three patients one of the combined LP variants fell into the *CHD1L* gene. In patient SN021, the heterozygous *PAX2* Y185* is combined with the heterozygous *CHD1L* R410*. Patient SN114 carries the compound heterozygous variants S350P and S510Y in *NPHS1* together with the K365fs variant in *CHD1L*. In patients SN148 the heterozygous *SALL1* N1090S variant is combined with the heterozygous *CHD1L* Q51* variant. LP variant in *RET* were combined with SRNS gene LP variants in two patients. Patient SN133 carries the hemizygous c.321+1G>C variant in *COL4A5* together with

the heterozygous K994N variant in *RET*. Patient SN150 carries the heterozygous *MYH9* V1837G variant combined with the heterozygous *RET* K1003E variant. Finally, in patient SN059, the heterozygous *INF2* R218Q is combined with the heterozygous *ROBO2* M189K.

To disclose whether the combination of LP variants in CAKUT-associated genes with SRNS-associated genes is a random finding or it may reflect an underlying mechanism, the prevalence of CAKUT LP variants in patients carrying or not carrying LP variants in SRNS-associated genes was compared. The prevalence of CAKUT LP variants was higher in patients carrying SRNS LP variants compared to those patients without SRNS LP variants (0.20 *vs.* 0.05, p=0.026). Moreover, the rate of patients with one CAKUT-associated gene LP variant combined to a SRNS-associated gene LP variant was higher compared to 404 ethnically matched controls (0.045 *vs.* 0.007, p=0.014). If the combinations of LP variants in two different CAKUT genes were also considered together with the CAKUT-SRSNS combinations, the significance became stronger (0.054 *vs.* 0.010, p=0.0089).

Table 5.5. Likely pathogenic variants identified in SRNS-associated genes.

ID	Genomic coordinates	Gene	AA/cDNA Change	GT	Inher.	All DB Max Freq	ESP Max Freq	CADD	PPr
SN025	chr4:86921740delCGAG	*ARHGAP24*	p.A704fs	Het	AD	0	0	35	1 / 1
SN094	chr2:228155553G>A	*COL4A3*	G1054E	Het	AD	0	0	17	9 / 10
SN121	chr2:228116080G>A	*COL4A3*	G213E	Het	AD	0	0	22	10 / 10
SN143	chr2:228175562G>A	*COL4A3*	R1609Q	Het	AD	0	0	24	10 / 10
SN173	chr2:228176560C>G	*COL4A3*	Q1663E	Het	AD	0	0	19	7 / 10
SN208	chr2:228168752G>A	*COL4A3*	G1349S	Het	AD	0	0	20	9 / 10
SN319	chr2:228159248A>G	*COL4A3*	K1127R	Het	AD	0	0	15	3 / 10
1862	chr2:227983416C>G	*COL4A4*	R145T	Het	AD	0	0	17	8 / 9
SN121	chr2:227872066C>T	*COL4A4*	C1683Y	Het	AD	0	0	20	9 / 9
SN163	chr2:227872070G>A	*COL4A4*	R1682W	Hom	AD	1.5E-05	0	16	9 / 9
SN248	chr2:227976420C>T	*COL4A4*	G190R	Het	AD	1.50E-05	0	20	9 / 9
QUGI249	chrX:107846253G>T	*COL4A5*	E736*	Het	XL	0	0	40	4 / 6
SN041	chrX:107846265G>T	*COL4A5*	G740W	Hem	XL	0	0	16	10 / 10
SN133	chrX:107811904G>C	*COL4A5*	c.321+1G>C	Hem	XL	0	0	24	4 / 4
SN152	chrX:107823772G>C	*COL4A5*	G264R	Het	XL	0	0	20	10 / 10
2089	chrX:107403705T>C	*COL4A6*	N1448D	Het	XL	0	0	17	9 / 10
SN098	chr4:187518866C>T	*FAT1*	p.C4113Y	Het	AR	0	0	27	10 / 10
SN098	chr4:187630590G>T	*FAT1*	p.A131E	Het	AR	0	0	15	6 / 10
SN003	chr14:105169782G>A	*INF2*	E220K	Het	AD	0	0	18	9 / 10
SN059	chr14:105169777G>A	*INF2*	R218Q	Het	AD	0	0	21	8 / 10
SN138	chr14:105167919G>A	*INF2*	G73S	Het	AD	0	0	36	9 / 10
SN150	chr22:36680531A>C	*MYH9*	V1837G	Het	AD	0	0	19	9 / 10
SN112	chr19:36332622C>T	*NPHS1*	S937N	Het	AR	5.00E-04	0.0005	23	9 / 10
SN112	chr19:36341872delTGG	*NPHS1*	p.Thr172del	Het	AR	6.07E-05	0	18	2 / 2
SN114	chr19:36337008G>T	*NPHS1*	S510Y	Het	AR	0	0	20	9 / 10
SN114	chr19:36339661A>G	*NPHS1*	S350P	Het	AR	8.65E-05	0	15	5 / 10
SN021	chr10:102541061C>G	*PAX2*	Y185*	Het	AD	0	0	41	5 / 6
SN086	chr10:102506046T>C	*PAX2*	F10S	Het	AD	0	0	22	8 / 9
SN272	chr11:101323799G>A	*TRPC6*	R895C	Het	AD	0	0	24	10 / 10

Hom: homozygous; Het: heterozygous; Hem: hemizygous; Inher.: inheritance; AD: autosomal dominant; AR: autosomal recessive; XL: X-linked. AA/cDNA Change: Amino acid or cDNA change. PPr: Pathogenic prediction; number of pathogenic predictions on the number of total predictions obtained with 10 different algorithms.

Table 5.6. Likely pathogenic variants identified in CAKUT-associated genes (*PAX2* was not included).

ID	Genomic coordinates	Gene	AA/cDNA Change	GT	Inher.	All DB Max Freq	ESP Max Freq	CADD	PPr
SN114	chr1:146747775ins AAGTTATT	CHD1L	p.K365fs	Het	AD	0	0	33	1 / 1
SN021	chr1:146756158C>T	CHD1L	p.R410*	Het	AD	0.0002	0	38	5 / 6
SN148	chr1:146737614C>T	CHD1L	p.Q51*	Het	AD	0.0016	0	25	4 / 6
SN170	chr16:86601014T>C	FOXC2	p.Y25H	Het	AD	1.6E-05	0	18	9 / 9
2087	chr13:39262130C>T	FREM2	p.R217C	Het	AR	1.0E-04	0	17	9 / 10
2087	chr13:39262490G>T	FREM2	p.A337S	Het	AR	1.0E-04	0	23	8 / 10
SN151	chr17:36061014G>A	HNF1B	p.A503V	Het	AD	0	0	23	8 / 10
SN133	chr10:43620373A>C	RET	p.K994N	Het	AD	1.5E-05	0	17	8 / 10
SN150	chr10:43620398A>G	RET	p.K1003E	Het	AD	0	0	21	8 / 10
SN059	chr3:77530269T>A	ROBO2	p.M189K	Het	AD	0	0	25	6 / 10
SN148	chr16:51172864T>C	SALL1	p.N1090S	Het	AD	1.0E-04	1.00E-04	24	6 / 10

Hom: homozygous; Het: heterozygous; Hem: hemizygous; Inher.: inheritance; AD: autosomal dominant; AR: autosomal recessive; XL: X-linked. AA/cDNA Change: Amino acid or cDNA change. PPr: Pathogenic prediction; number of pathogenic predictions on the number of total predictions obtained with 10 different algorithms.

Table 5.7. Likely pathogenic variants identified in amyloidosis and metabolic disorder associated genes.

ID	Genomic coordinates	Gene	AA/cDNA Change	GT	Inher.	All DB Max Freq	ESP Max Freq	CADD	PPr
SN128	chr4:155505415T>C	*FGA*	p.Y821C	Het	AD	0	0	16	10 / 10
SN177	chr4:155505859A>G	*FGA*	p.L673P	Het	AD	1.0E-04	1.0E-04	18	10 / 10
SN016	chr2:216274402A>T	*FN1*	p.V728E	Het	AD	0	0	25	9 / 10
SN016	chr2:216274406A>C	*FN1*	p.S727A	Het	AD	0	0	17	9 / 10
SN048	chr2:216269288C>T	*FN1*	p.G1026E	Het	AD	0	0	30	9 / 10

Hom: homozygous; Het: heterozygous; Hem: hemizygous; Inher.: inheritance; AD: autosomal dominant; AR: autosomal recessive; XL: X-linked. AA/cDNA Change: Amino acid or cDNA change. PPr: Pathogenic prediction; number of pathogenic predictions on the number of total predictions obtained with 10 different algorithms.

Table 5.8. Frequency of subjects carrying likely pathogenic variants in known glomerulopathy-associated genes.

Gene	Frequency of LP variant carriers		OR	*p*
	Patients	Controls		
SRNS	0.234	0.032	8.04	**1.0×10^{-12}**
ARHGAP24	0.009	0.007	1.22	1.000
COL4A3	0.054	0.010	5.71	**0.009**
COL4A4	0.036	0.010	3.74	0.070
COL4A5	0.036	0.005	7.51	**0.022**
COL4A6	0.009	0.000	11.0	0.216
FAT1	0.001	0.000	3.63	1.000
INF2	0.027	0.007	3.71	0.118
MYH9	0.009	0.010	0.91	1.000
NPHS1	0.009	0.000	11.0	0.216
PAX2	0.018	0.002	7.39	0.119
TRPC6	0.009	0.005	1.83	0.518
CAKUT	0.081	0.022	3.87	**0.006**
CHD1L	0.027	0.012	2.22	0.378
FOXC2	0.009	0.002	3.66	0.385
FREM2	0.009	0.000	11.0	0.216
HNF1B	0.009	0.000	11.0	0.216
RET	0.027	0.010	2.78	0.174
ROBO2	0.018	0.015	1.22	0.684
SALL1	0.009	0.005	1.83	0.518
Amyl./Metab. dis.	0.018	0.007	2.45	0.295
FGA	0.018	0.000	18.5	**0.046**
FN1	0.018	0.017	1.04	1.000

OR: odds-ratio. Amyl./Metab. dis.: amyloidosis or metabolic disorders.

5.3.1.4 Analysis of the candidate genes

Since information on inheritance was not available for candidate genes, all variants with a frequency ≤0.001 and a CADD score ≥15 were considered as possibly pathogenic (PP) variants. As shown in Table 5.9, there were 10 patients carrying PP variants in 6 different candidate genes (*EPB41L5* n=3, *TINAG* n=3, *KIRREL* n=1, *PDLIM2* n=1, *TJP1* n=1, and *TRIM3* n=1).

To disclose whether these variants are incidental findings, the prevalence of PP variants as defined above was evaluated in 404 ethnically matched controls (Table 5.10). The prevalence of PP variants in *EPB41L5* was higher in patients compared to controls (0.027 *vs.* 0.002, p=0.033). No differences were observed in the prevalence of the PP variants in the remaining candidate genes between patients and controls.

A deeper analysis of the *EPB41L5* variants showed that they were all carried in heterozygous state by three patients (1 familial and 2 sporadic cases). The affected relatives of the familial cases were not available for segregation studies. All three variants affect highly conserved amino acids (Figure 5.4). Two of them (Q168E and E287D) are located at the FERM domain of EPB41L5 (Figure 5.5). Through the FERM domain, EPB41L5 interacts with other partners of the Crumbs complex regulating the formation of the slit diaphragm and the maintenance of podocyte cell polarity.[364]

Table 5.9. Possibly pathogenic variants in candidate genes included in the NGS panel.

ID	Genomic coordinates	Gene	AA/cDNA Change	GT	Inher.	All DB Max Freq	ESP Max Freq	CADD	PPr
SN031	chr2:120844804A>T	*EPB41L5*	p.E287D	Het	na	1.00E-04	0	17	8 / 10
SN037	chr2:120833326C>G	*EPB41L5*	p.Q168E	Het	na	3.01E-05	0	25	9 / 10
SN296	chr2:120925069C>T	*EPB41L5*	p.R664W	Het	na	1.00E-03	0	20	10 / 10
SN248	chr1:158064158C>A	*KIRREL*	p.T592N	Het	na	4.00E-04	0	19	8 / 10
SN161	chr8:22451386G>A	*PDLIM2*	p.R591H	Het	na	0	0	28	7 / 10
SN133	chr6:54173572A>G	*TINAG*	p.Y75C	Het	na	0	0	15	6 / 10
SN138	chr6:54186175G>A	*TINAG*	p.G167E	Het	na	0	0	18	8 / 10
SN205	chr6:54173549T>A	*TINAG*	p.D67E	Het	na	0	0	18	6 / 10
SN168	chr15:30009035C>G	*TJP1*	p.E1395Q	Het	na	0	0	28	9 / 10
SN071	chr11:6477659G>A	*TRIM3*	p.R433C	Het	na	1.62E-05	0	20	7 / 10

Hom: homozygous; Het: heterozygous; Hem: hemizygous. Inher.: inheritance. PPr: Pathogenic prediction; number of pathogenic predictions on the number of total predictions obtained with 10 different algorithms. na: not applicable.

Table 5.10. Frequency of subjects carrying possibly pathogenic variants in candidate genes included in the NGS panel.

Gene	Frequency of PP variant carriers		OR	*p*
	Patients	Controls		
EPB41L5	0.027	0.002	11.2	**0.033**
TINAG	0.027	0.010	2.78	0.174
TRIM3	0.009	0.005	1.83	0.518
KIRREL	0.009	0.005	1.83	0.518
PDLIM2	0.009	0.005	1.83	0.518
TJP1	0.018	0.035	0.51	0.541

Figure 5.4. Alignment of the EPB41L5 homologues from different species around the Q168E, E287D and R664W variants of *EPB41L5* obtained by the Polyphen-2 software. The residue affected by each variant is shown in the black frame.

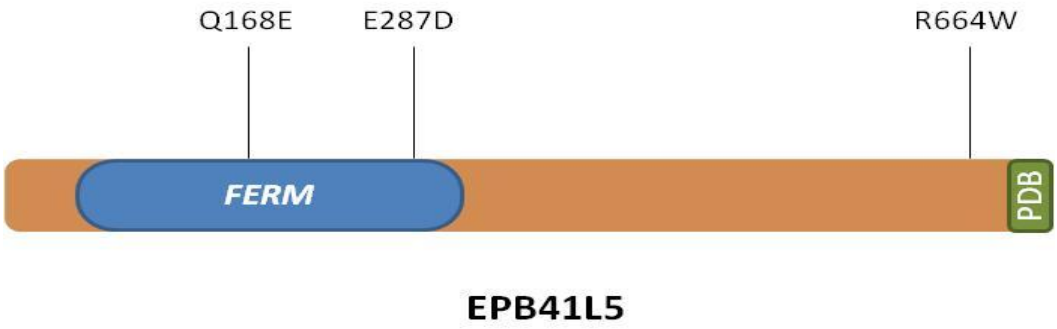


Figure 5.5. Graphical representation of the protein encoded by the candidate EPB41L5 gene. It contains a FERM domain at the N-terminus of the protein and a PDZ-binding motif (PDB) at the C-terminal. Vertical lines represent the possibly pathogenic variants found in 3 SRNS patients.

5.3.2 To search for novel genes involved in the pathogenesis of the familial form of SRNS.

Two families with SRNS patients and unknown etiology have been enrolled for genome wide analyses. For both, consanguinity of the patients' parents was hypothesized.

5.3.2.1 Novel gene hunting in patient MEBO291

Patient MEBO291 is a 24 years-old female of Italian origin. She developed SRNS at age 17 years. Proteinuria at onset was 12 g/24h. She presented no extra-renal manifestations. She underwent kidney biopsy with pathological features showing minimal change disease. Sanger sequencing revealed no mutations in *INF2*, *NPHS2*, *WT1* (exons 8 and 9) and *TRPC6*. Screening with the NGS panel revealed no mutations.

The patient pedigree is shown in Figure 5.6. She had no familial history for nephropathy, except from traces of proteins and blood in the urine in the younger brother. Consanguinity of the parents was referred.

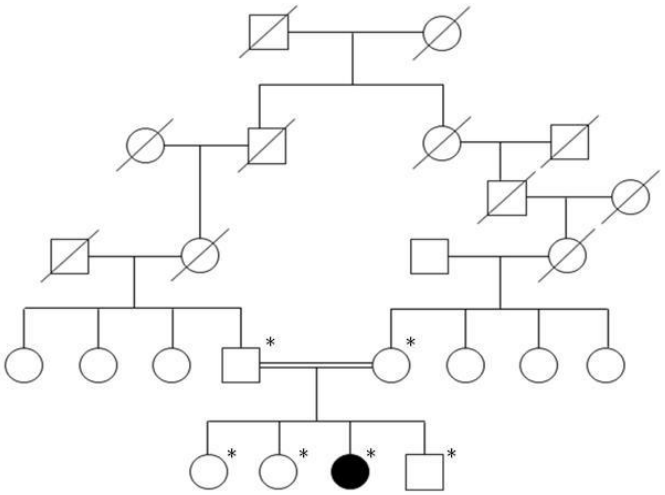


Figure 5.6. Pedigree of patient MEBO291; the asterisk denotes the subjects for whom DNA sample is available.

Autozygosity mapping using the results of Genotyping of about 1 million SNPs with the

Genome-Wide Human SNP array, showed the presence of different autozygosity regions (in chromosomes 2, 6, 13, 14 and 20) where the two alleles had high probability to be identical by descent (Figure 5.7, panels A and B). Parametric whole-genome linkage analysis assuming a autosomal recessive inheritance showed four candidate regions (in chromosomes 2, 6, 13 and 20) reaching a maximum LOD score of 1.85 (Figure 5.7, panel C).

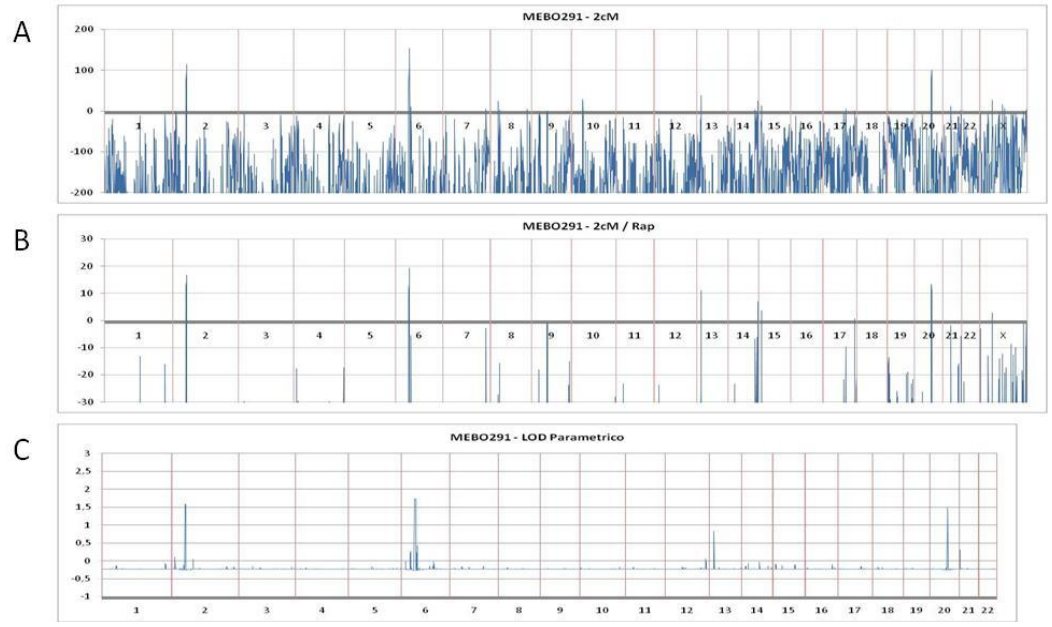


Figure 5.7. Autozygosity and linkage analysis in patient MEBO291. Panel A: classic autozygosity LOD score for regions of 2 cM (overlapped by 0.1 cM). Panel B: autozygosity LOD score for regions of 2 cM (overlapped by 0.1 cM) using the Density Adjusted Algorithm. Panel C: LOD score of parametric linkage analysis assuming an autosomal recessive inheritance.

Whole exome sequencing (WES) was performed in patient MEBO291 on an Illumina platform obtaining 101,849 variants. After filtering out common variants (minor allele frequency, MAF, >0.01), non functional variants, 355 variants survived. Considering an autosomal recessive inheritance due to consanguinity, I kept only the homozygous

variants, thus reducing the disease-associated candidates to 10. Among these, only the chr13:g.28,626,716C>T, that leads to amino acid change V194M in the *FLT3* gene, fell into one of the candidate regions identified by the autozygosity and linkage studies (Table 5.11). This variant has a MAF of 0.005, a CADD score of 10 and is predicted damaging by 4 out of 10 used algorithms. Evaluation of FLT3 with the String software showed that *FLT3* is linked to *WT1* and indirectly through *RAC1* with *ACTN4*, *MYO1E* and *ARHGDIA* (Figure 5.8).

Prioritization of all variants identified by the WES, independently from the inheritance, was performed using the Exomiser software (Table 5.12). This analysis highlighted the heterozygous P494S in the *DDR1* gene as the most probable disease-causing variant. This variant has a CADD score of 15, it is predicted damaging by 5 algorithms and is not present in the public databases.

Finally, variants in known kidney disorders were investigated. The R93W variant was identified in the CAKUT-associated GDNF gene. This variant has a MAF of 0.04, a CADD score of 16 and is predicted damaging by all the algorithms. The R93W has been previously described in CAKUT patients and it is considered of uncertain significance.

The presence of these three variants was evaluated in the family members of SN140 (Figure 5.9). The parents of SN140 were heterozygous for *FLT3* V194M; however, there were two unaffected siblings carrying V194M in homozygous state. The DDR1 P494S was inherited from the mother and was present in the first sister. Finally GDNF was inherited from the father and was also present in the second sister.

The genes FLT3, GDNF and DDR1 were sequenced in 67 SRNS patients carrying no mutations in known SRNS genes. No LP variants were identified in *FLT3* and in *GDNF*. A heterozygous G413A variant was found in the *DDR1* gene. This variant was not present in the public databases, has a CADD score of 14 and is predicted damaging

by 4 algorithms. However, the comparison with 404 controls using the LP definition for autosomal dominant inheritance showed no statistically significant difference (0.015 *vs.* 0.012, p=1.00).

Table 5.11. Number of variants that have been identified in MEBO291 during exome sequencing analysis and that have survived after the application of different filters.

	Total N° variants	… functional variants with MAF <0.01	… predicted damaging	… homozygous	… within linkage regions
N° variants	101,849	814	355	10	1

Table 5.12. Genes that ranked in the first ten positions during prioritization of the SN140 variants with the Exomiser software.

Rank	Gene	Exomiser Score	Phenotype Score	Variant Score	Phenotypic similarity	Phenotypic similarity to mouse mutant	Proximity in interactome
1	*DDR1*	0.977	0.754	1.000	.	0.754 to mouse mutant involving DDR1	FN1
2	*IFT122*	0.929	0.642	1.000	0.642 to Cranioectodermal dysplasia 1 associated with IFT122.	.	LMX1B
3	*NOS1*	0.91	0.618	1.000	.	0.618 to mouse mutant involving NOS1	ASL
4	*SLC22A5*	0.904	0.611	1.000	.	0.611 to mouse mutant involving SLC22A5	SLC26A6
5	*SLC6A19*	0.904	0.611	1.000	0.393 to Pseudomonas aeruginosa, susceptibility to chronic infection by, in cystic fibrosis associated with FCGR2A	.	SLC6A18
6	*FCGR2A*	0.901	0.740	0.850	.	0.740 to mouse mutant involving FCGR2A	APCS
7	*AGER*	0.886	0.725	0.850	.	0.725 to mouse mutant involving AGER	THBD
8	*STARD8*	0.876	0.587	0.996	.	.	TNS1
9	*REEP1*	0.874	0.581	1.000	.	.	IFT122
10	*SF3A2*	0.865	0.574	1.000	to Prader-Willi syndrome associated with SNRPN	.	SNRPN

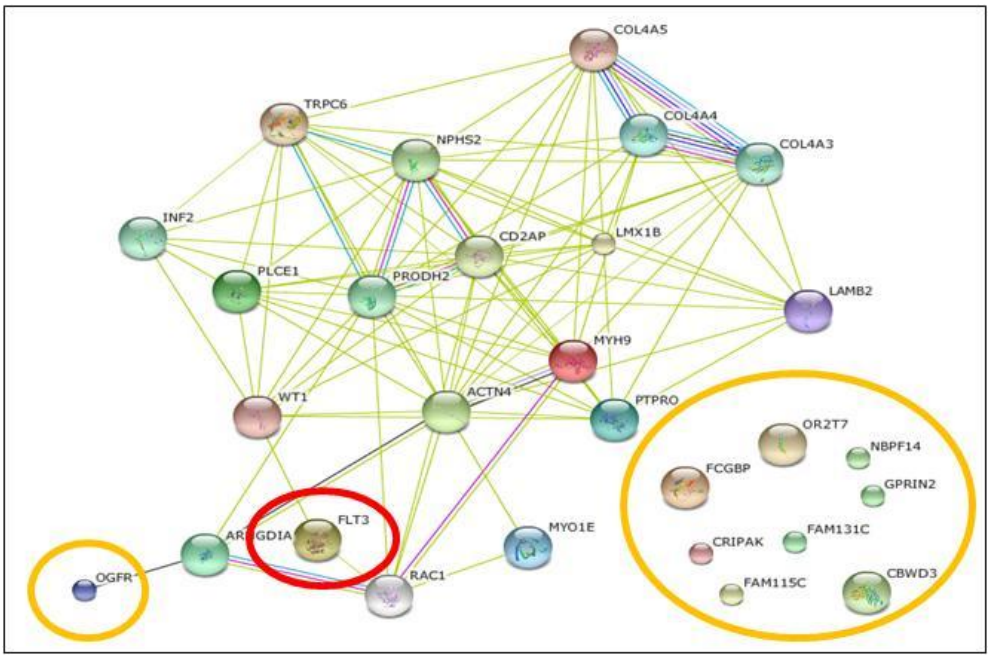
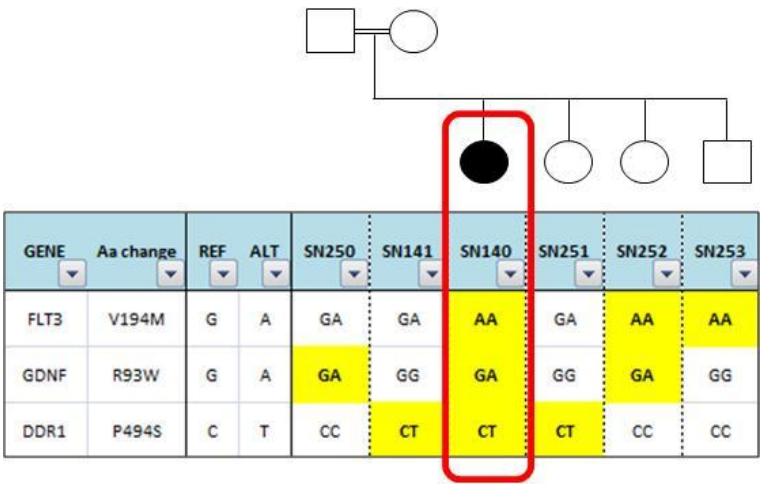


Figure 5.8. Output of the direct (physical) and indirect (functional) protein-protein interaction analysis using the String v9.05 software. These interactions may be both known or predicted and they derive from four sources: genomic context, high-throughput experiments, (conserved) coexpression, and previous knowledge (data mining). The already known podocytopathy-associated genes and the 10 exome derived candidate disease-associated genes were used as input. In the red circle is shown the candidate *FLT3* gene that is integrated in the network. In the yellow circles the other genes with homozygous variants identified in the WES.



GENE	Aa change	REF	ALT	SN250	SN141	SN140	SN251	SN252	SN253
FLT3	V194M	G	A	GA	GA	AA	GA	AA	AA
GDNF	R93W	G	A	GA	GG	GA	GG	GA	GG
DDR1	P494S	C	T	CC	CT	CT	CT	CC	CC

Figure 5.9. Segregation of the candidate variants in the SN140 pedigree.

5.3.2.2 Novel gene hunting in patients IR35 and IR40

Patients IR35 and IR40 are two male brothers of Iranian origin (Figure 5.10). Patient IR40 is the first born of the family. He is 46 years-old and is affected by SRNS since age 41 years-old. The proteinuria at onset was 3.8g/24h. He had no extra-renal manifestations. Kidney biopsy showed minimal change nephropathy with mild mesangial proliferation and mild segmental thickness of glomerular basement membrane.

IR35 is the second born of the family. He is a 40 years-old male affected by SRNS with onset at age 30 years. Proteinuria at onset was 3.4 g/24h. He had no extra-renal manifestations. Kidney biopsy showed minimal change nephropathy. He developed ESRD requiring dialysis at the age of 33 years and received a kidney transplantation at the age of 35 years. No recurrence or rejection of the graft were observed at 5 years of follow-up.

Parents consanguinity was referred and suggested by the common geographical origin of parents (parents are both from a small village), but the common ancestor was not identified. Both patients had no mutations in SRNS genes.

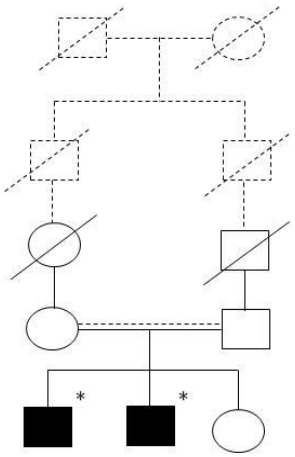


Figure 5.10. Pedigree of patients IR35 and IR40; the asterisk denotes the subjects for whom DNA sample was available.

Autozygosity mapping showed the presence a single 3.5 Mb long region in chromosome 1 where the two alleles have high probability to be identical by descent and that was shared by both affected brothers (Figure 5.11). WES was performed in patient IR35 and showed the presence of 53 homozygous variants with MAF <0.01 and predicted damaging (Table 5.13). Among these 53 variants only the chr1:g.12,836,101C>A, leading to the amino acid change L235I in the *PRAMEF12* gene, fell in the candidate region identified by the autozygosity analysis (Table 6). This variant has a CADD score of 24 and is predicted as damaging by 5 out of 10 algorithms. However, its allele frequency is rather high (MAF=0.009). Furthermore, gene *PRAMEF12* did not show any link to the podocytopathy-associated genes when evaluated with String (Figure 5.12).

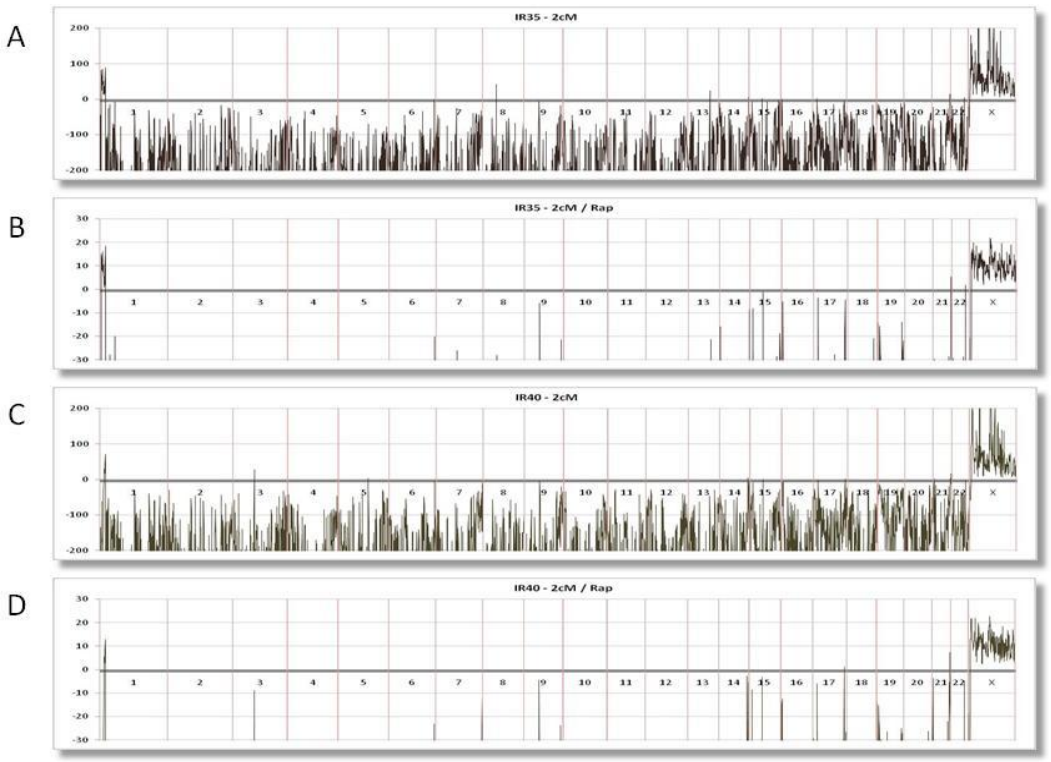


Figure 5.11. Autozygosity and linkage analysis in patients IR35 and IR40. Panel A: autozygosity LOD score for regions of 2 cM (overlapped by 0.1 cM) in patient IR35. Panel B: autozygosity LOD score for regions of 2 cM (overlapped by 0.1 cM) using the Density Adjusted Algorithm in patient IR35. Panel C: classic autozygosity LOD score for regions of 2 cM (overlapped by 0.1 cM) in patient IR40. Panel D: autozygosity LOD score for regions of 2 cM (overlapped by 0.1 cM) using the Density Adjusted Algorithm in patient IR40.

Table 5.13. Number of variants that have been identified in IR35 during exome sequencing analysis and that have survived after the application of different filters.

	Total N° variants	… functional variants with MAF <0.01	… predicted damaging	… homozygous	… in autozygosity region
N° variants	62113	572	359	53	1

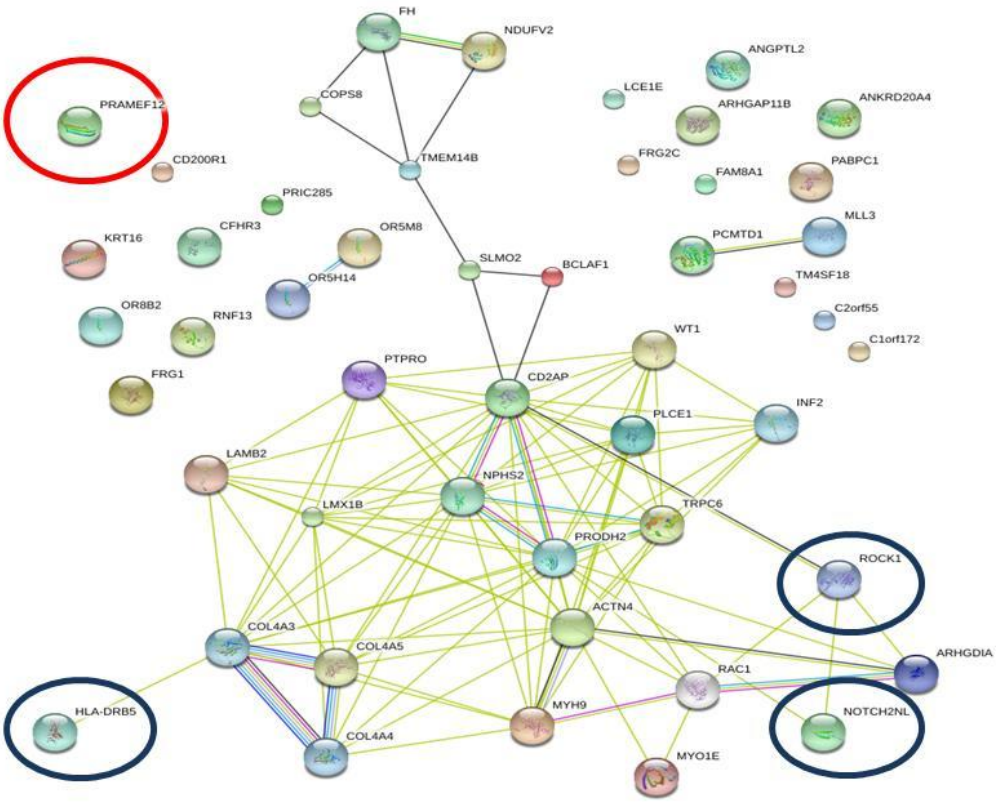


Figure 5.12. Output of the direct (physical) and indirect (functional) protein-protein interaction analysis using the String v9.05 software. The already known podocytopathy-associated genes and the 53 exome derived candidate disease-associated genes were used as input. In the dark blue circles the candidate that are integrated in the network are shown. In the red circle the only gene lying in the autozygosity region is shown.

5.3 Discussion

In this study, a next-generation sequencing panel was designed and developed, which analyses whole coding exons and SNPs in 93 different genes known to be associated with glomerulopathies including the disorders that commonly cause nephrotic syndrome. The panel was used to screen 111 index cases from a SRNS cohort and resolved 30% of overall SRNS cases. The majority of the mutations are found in already known SRNS-associated genes. Interestingly, likely pathogenic variants are also found in genes previously associated with CAKUT, as well as with amyloidosis and glomerulopathy with fibronectin deposits. Finally, through candidate screening and exome studies, the *EPB41L5* is identified as candidate SRNS-associated genes.

In this cohort, likely pathogenic variants in SRNS-associated genes are found in 23% of patients (Tables 5.5-5.8). The most frequently affected genes are *COL4A3*, *COL4A4* and *COL4A5*. Thirteen (12%) out of 111 index patients carry likely pathogenic variants in one of these three genes. Contrary to a previous large study,[152] a high prevalence of mutations in *NPHS2*, *NPHS1*, *WT1*, *PLCE1* and *LAMB2* was not observed. Regarding *WT1*, its low prevalence was probably biased by the fact that 5 patients carrying *WT1* mutations identified in a previous Sanger sequencing study, were not included in the study cohort. As for the other genes, discrepancy with literature may be due to the higher age of SRNS onset of the patients included in this study (median age 20.5 years) as compared to published studies that mainly focused on pediatric cohorts. Indeed NPHS2, NPHS1, PLCE1 and LAMB2 mutations mostly affect infants or small children with congenital nephrotic syndrome or early onset SRNS, while their prevalence drop with age.[152] Conversely, the data presented here are in line with previous findings that *COL4A4* mutations are the most frequent mutations in adult focal segmental glomerulosclerosis.[210]

Interestingly, likely pathogenic variants are also found in genes other that those already known to be associated with SRNS. In particular, 8% of patients carry LP variants in CAKUT-associated genes. This may reflect the fact that the molecular pathways underlying CAKUT and SRNS partially overlap. CAKUT genes are involved in the kidney development from the early stages to the glomerulus and podocyte formation.[219] The *PAX2* gene, which was considered as a SRNS-associated gene in this study because it is associated with adult onset FSGS, was first identified as a CAKUT-causing gene.[221, 365] Furthermore, transcription factors involved in podocyte development include proteins encoded by genes associated both with FSGS (*WT1* and *PAX2*) and CAKUT (*LMX1B* and *FOXC2*). Interestingly, in the present cohort the CAKUT LP variants are frequently combined with SRNS variants and these associations do not seem incidental. It is possible that these variants concur to the development of the disease or act as genetic modifiers resulting in more severe disease presentation.

This study also aimed to identify novel SRNS-causing variants. Two different strategies have been adopted. The first was to identify candidate genes on the basis of their expression in the human kidney and glomerulus, their association with glomerular abnormalities in animal models, their interaction with proteins associated with SRNS, MPGN and MN, and their tolerance to functional genetic variation. Software and algorithms that prioritize genes on the basis of multiple features, such as ToppGene, Endeavour and GeneDistiller, or on the low tolerance to functional genetic variation, were also used. All these data were used to obtain a list of top genes followed by manual evaluation based on the literature. On September 2015, 21 candidates were selected and included in the NGS panel. Since then, five new SRNS-causing genes have been reported. One of these, *FAT1*, which encodes an atypical cadherin present at the podocyte slit diaphragm, was included among the candidates of the panel here presented. On the other hand, the four nucleoporin genes (*NUP93*, *NUP107*, *NUP205*

and *XPO5*), which have been involved in a new SRNS-associated pathway, were not included. Therefore, candidate gene analysis may be a viable approach for genes involved in known pathways, but would probably fail to catch those belonging to unknown mechanisms.

Among the other candidate genes analyzed here, the *EPB41L5* is the most promising to be involved in SRNS pathogenesis (Tables 5.9 and 5.10). The *EPB41L5* encodes for the Erythrocyte Membrane Protein Band 4.1 Like 5, a protein involved in the formation of slit diaphragm and in the maintenance of podocyte cell polarity.[366] It is a component of the mammalian Crumbs complex and antagonizes its signaling.[366] Studies in zebrafish show that *moe*, the zebrafish *EPB41L5* ortholog, is expressed in podocytes and it is required for proper formation and maintenance of slit-diaphragms. Podocytes of *moe* -/- mutants lack slit-diaphragms on basement membrane-associated foot processes.[364] In the mutant embryos, it is possible that slit-diaphragms may form at sites distant from the basement membrane, since some double cross-strand structures spanning the space between processes extending into Bowman's space are observed.[364] Absence of *moe* expression results in loss of glomerular barrier permselectivity and aberrant passage of high molecular weight molecules.[364] In addition, through its FERM domain, EPB41L5 directly binds CRB2 that is a known SRNS-causing gene.[235, 236, 367, 368] Zebrafish crb2b is required for podocyte foot process arborization, slit diaphragm formation, and proper nephrin trafficking.[235] In the present SRNS cohort, three possibly pathogenic variants of *EPB41L5* are present. All the three variants affect evolutionary conserved amino acids (Figures 5.4 and 5.5) and two of them are located in the FERM domain, through which EPB41L5 interacts with other proteins of the Crumb complex.[366] Notably, these two variants fulfill the criteria of likely pathogenic variants adopted in this study. Finally, the prevalence of possibly pathogenic variants was significantly higher in patients compared to controls. However, the significance level is not robust and the effect of the

identified variant was evaluated only *in silico*. Analysis of further SRNS cohorts and functional studies on the role of the identified variants are necessary.

A second strategy, based on whole genome autozygosity mapping and whole exome sequencing allowed the identification of *DDR1 as* a potential SRNS candidate gene (Table 5.12). This gene encodes for the Epithelial discoidin domain-containing receptor 1, which is a collagen receptor that bind collagen IV in podocytes.[369] It also interacts with myosin IIA (MYH9), another SRNS-associated protein, thus regulating cell spreading and motility.[370, 371] DDR1-deficient mice show localized thickening of the glomerular basement membrane, focal loss of slit diaphragms and proteinuria.[369] Through the exome study a heterozygous missense variant of DDR1 was identified (P494S). Further screening of 67 patients disclosed the presence of a second variant (G413A). These variants are located at the second and first Glycine/Proline-rich domains of DDR1, respectively, and both affect evolutionary conserved residues. However, the variant was found in two unaffected relatives and the prevalence of likely pathogenic variants is not significantly higher in patients compared to controls. Altogether, the data available here do not support a role in the pathogenesis of SRNS, although such a role cannot be excluded. Further screening of patients and functional studies are warranted to define the role of *DDR1*.

Some limitations of the present study have to be considered. The number of patients included in this cohort is low. Moreover, the control subjects were obtained from a public database. They were not recruited from the same territory of the patients and they were not analyzed with the same technique. Finally, the pathogenicity of many variants is based on prediction software and algorithms, and functional studies are not available. Therefore, a cautious interpretation of the results is necessary. This is a pilot study to evaluate the performance of the NGS panel. Recruitment and analysis of further SRNS

patients, as well as screening of control subjects, is in the future perspectives. However, some elements enforce the findings presented here. The high rate of combined CAKUT LP variants support these are not incidental findings. Moreover, the absence of LP variants in genes that present a different pathogenesis, such as the MPGN-associated genes, provide further support.

In conclusion, this study shows that next-generation sequencing can be successfully applied in the screening of SRNS. It also suggests that CAKUT genes may be involved on the pathogenesis of SRNS. Finally, *EPB41L45* is proposed as a candidate gene for SRNS.

6. OPTIMIZATION OF VARIANT IDENTIFICATION IN THE ION TORRENT PGM PLATFORM

6.1 Introduction

Next Generation Sequencing (NGS) has been a revolutionary tool for discovering causal variants in human diseases and unraveling new disease mechanisms.[372] Compared to classical Sanger sequencing, NGS can detect genetic abnormalities from a significantly larger number of genes or samples in parallel, in a more rapid and cost effective manner.[373] Based on the platforms and kits, the sequencing may yield hundreds to millions of variants compared to a human reference sequence. In these large data sets, the identification of causal variants may be arduous because of the presence of many false positive variants.

Distinguishing true variants from false positive variants is an outstanding challenge in variant discovery. High sensitivity and specificity are essential for clinical genetic tests, since high error rates may compromise confirmation of diagnosis, treatment strategies, genetic counseling and carrier screening.[374] Inaccurate genetic tests have significant impact on validation time and costs.[375] Nonetheless, several studies on currently available NGS sequencers have highlighted that the accuracy of mutation detection should be improved to ensure that NGS become a viable option for clinical genetic testing.[354-356, 373, 375]

Several NGS sequencers are available and they are characterized by distinct error profiles, which are related to the respective sequencing technology.[376] The overall error

rate ranges from 0.3 per 100 sequenced bases for the Illumina platform to 16.2% of the Pacific biosciences platform (Table 6.1). The Ion Torrent platform that emerged as a viable alternative to Illumina in the last years showed 1.6% of error rate.[376, 377] Regarding the type of error, Illumina and Complete Genomics show more substitution than indel errors, while indel errors are predominant in all other platforms.[376]

Table 6.1. Error rates (per 100 sequenced bases) of high-throughput sequencing platforms.

Platform	Subs (±SD)	Indels (±SD)	All (±SD)[a]
Illumina HiSeq	0.26 (±0.11)	0.03 (±0.02)	0.28 (±0.12)
Illumina MiSeq	0.25 (±0.11)	0.01 (±0.01)	0.30 (±0.19)
Ion Torrent PGM	0.17 (±0.17)	1.46 (±1.22)[b]	1.63 (±c 1.24)[b]
454 GS FLX	0.09 (NA[c])	0.90 (NA[c])	0.99 (NA[c])
454 GS Junior	0.05 (NA[c])	0.39 (NA[c])	0.46 (NA[c])
Complete Genomics	2.30 (NA[c])	0.02 (NA[c])	2.32 (NA[c])
Pacific Biosciences RS	1.10 (±0.45)	15.57 (±3.29)	16.19 (±3.17)

Modified from Laehnemann, 2016.[376] Subs: substitution errors per 100 bases; Indels: insertion and deletion errors per 100 bases; SD: standard deviation.
[a]Some studies contain only aggregated measures for Indels and/or total error; therefore, the value of All is not necessarily the sum of Subs and Indels.
[b]One study with three samples used indel-tolerant mapping, resulting in almost 100% of reads being mapped, but also producing much higher indel error rates and high SD.
[c]Only one sample was analyzed, thus SDs are not available.

The false positive variants are commonly machine artifacts or alignment errors.[378-380] The artifacts are introduced during sample preparation, clonal amplification, sequencing cycles, and data collection.[381, 382] Amplification bias during PCR can result in underrepresentation of one of the two alleles in heterozygous variants.[383] Polymerase errors, such as base misincorporations and rearrangements secondary to template switching, may also result in incorrect variant calls.

Although the overall error rates are relatively low, some properties of nucleic acid

sequences are known to raise the error rates.[372, 376, 384, 385] These sequence-specific error sources include homopolymers and regions with high GC content. Homopolymers are stretches of identical bases, which are frequently associated with errors at the bases next to the homopolymer. As short as two nucleotide homopolymers may increase the sequencing error at the base following the homopolymer.[385] Moreover, certain DNA sequence motifs, such as GGA or GGT, are more susceptible to sequencing errors.[372, 385] High GC content regions are another source of error in NGS because they negatively affect read coverage and because of their low complexity leads to polymerase errors.[376, 386] Furthermore, single nucleotides show different behavior. In the Illumina platform the A↔C and the G↔T transversions are much more common than other false substitutions.[376] Finally, a source of error common in many platforms is the decay of base signal along each read.[376] Toward the end of the read, the frequency of erroneous calls significantly increase.[376]

Alignment errors are a further source of false positive variants. They are typically associated with repetitive or homologous sequence regions, and they can lead to substitution, insertion and/or deletion errors involving one or more bases.[372, 387] Since both sequencing and alignment errors can be associated with particular DNA sequence motifs, these errors can be consistent between samples analyzed using the same sequencing chemistry, run on the same instrument and analyzed with the same bioinformatic pipeline, as shown in multiple studies.[384, 388-390]

In the last years, new biochemical protocols and bioinformatic solutions have been introduced to improve sequencing accuracy.[382] Since the large majority of the studies have used Illumina sequencers, most technical optimization efforts have been focused on this platform.[377, 391-393] Accordingly, most bioinformatic methods are optimized for Illumina-derived data.[376, 377, 394-398] Several basic parameters on variant calling and

alignment have been developed within the different variant calling softwares to identify false positive variants. These parameters, such as the probabilistic base quality score, the alignment mapping quality score, the aligned read coverage for possible alleles and the base alignment quality score, can be used to apply hard filters for variants to discard putative false positives.[372, 380, 399, 400] They can also be used in a statistical model to generate the variant quality score (QUAL), an estimated Phred-scaled probability that the variant exists given sequencing data. However, this score underperforms at DNA regions prone to sequencing and/or alignment errors. To address QUAL score limitations, further parameters have been developed, including several bias values that compare variant containing reads versus reference containing reads for systematic differences in base quality, mapping quality, mapping position, and mapping direction.

A more sophisticated implementation of the above parameters in Illumina-derived data is the variant quality score recalibration (VQSR), a quality score that is supposed to be better calibrated compared to QUAL. To obtain VQSR, machine learning algorithms are employed under a multidimensional Gaussian mixture model to learn from each dataset what is the parameter profile of true variants *vs.* the profile of false positive variants.[380] However, VQSR does not fully capture site specific issues, because it evaluates variants from many different regions simultaneously and the site-specific characteristics are difficult to codify.[372]

Other methods incorporate additional information to more accurately identify sites with higher variant calling error rate, such as methods that utilize cross-sample error information. This approach assesses the characteristics of the errors by pooling information across samples and positions, since specific region likely contain similar errors in different samples.[388, 389, 401, 402]

Recently, VarBin has been proposed as a novel analytical method to distinguish true

from false positive variants in Illumina NGS data. VarBin is a variant likelihood binning method that uses Genotype likelihood scores (PL) to generate the variant likelihood ratio over coverage depth (PLRD), a value affected by alignment and sequencing error. Then it compares the PLRD calculated for the variant of interest to those calculated for multiple background samples for each variant change and position. Similarly to other cross-sample methods, it is based on the fact that error characteristics show similar trends between samples sequenced with the same chemistry on the same platform and with the same bioinformatic data pipeline. Its advantage is that it evaluates each variant site individually.

While there are numerous scientific publications aiming to address false positive variation on Illumina platform, only few studies are available for the Ion Torrent platform. This may be due to the fact that it is a young technology launched in 2011. The Ion Torrent NGS platforms generate DNA sequencing reads by detecting pH changes caused by ion release during incorporation of deoxyribonucleotide triphosphates into a growing DNA chain on a semiconductor device.[403] In general, Ion Torrent NGS platform is characterized by high accuracy at single nucleotide variant (SNV) discovery, but has a high false positive rate for the identification of small indels.[355, 377, 404] This is particularly pronounced in homopolymer regions.[403]

Several improvements have been delivered by Life Technologies for the Ion Torrent platform. In 2015, the HiQ sequencing chemistry was released to improve accuracy of indel detection. According to the manufacturer, indel error rates decrease up to 90% compared to previous Ion Torrent sequencing chemistries.[405] Moreover, to overcome the problem of high indel false positive rates, multiple upgrades of the Torrent Suite, the proprietary analytical workflow for the Ion Torrent platform, have been released.[403] More recent versions of the Torrent Suite Variant Caller (TSVC) include flow signal

information during variant calling, which is expected to improve the accuracy.

In the scientific literature, a first solution was proposed within a workflow for clinical genetic testing of the *BRCA1* and *BRCA2* genes on Ion Torrent PGM.[406] The workflow used a filter of 'presence of variants in <15% of samples' thus discarding frequently recurring errors and achieving a sensitivity ≥98.6% and a specificity of 96.9%. However, such approach may result in low sensitivity for samples derived from the same family or that contain highly conserved functional mutations.

A study on improving the indel detection specificity of *BRCA1* and *BRCA2* genes using PGM implemented two simple filtering criteria: B-allele frequency (BAF), that is the non-reference allele frequency, and variation of the width of deletions and insertions (VARW), which is the variance of the width of the inserted/deleted sequences.[375] After optimizing the filtering thresholds, the authors succeeded to achieve ≥99.99% specificity retaining 100% sensitivity.[375]

In a follow-up study the same group adopted the Quality by Depth (QD) parameter, which is the ratio of QUAL score on read depth, as an alternative to the BAF threshold, since BAF estimation is unreliable in regions of low read depth.[403, 407] They showed that the combination of QD with the VARW can substantially improve variant detection specificity without compromising sensitivity.

A more sophisticated strategy was developed by Damiati and coauthors for data from whole exome sequencing on Ion Proton sequencer.[377] For SNV and indel calls separately, they identified five parameters calculated by the TSVC that could better discriminate erroneous calls. Both for SNVs and indels these parameters are: QUAL, GQ (the genotype quality score), FDP (the flow-space corrected read depth) and FAO (the flow-space corrected alternate allele observations). The fifth parameter is STB, the strand bias ratio, for SNVs and HRUN, the length of homopolymer, for indels. Three

sets of thresholds for these parameters are provided, corresponding to low, medium and high-stringency levels of filtering, which filter out 20-68% of false SNVs and 10-40% of false indels retaining 99%, 95% and 90% of true positive calls, respectively.

Interestingly, manual inspection of NGS data in the Integrative Genomics Viewer (IGV) browser allows to differentiate likely true variants from false positive variants. [372] Comparison of the reads from different samples allow the visual identification of common false positive variant features such as the low variant allele percentage, the presence of the variant in reads from only one direction, the sequence context, and the presence of the variant in many background samples even if it may not have passed the variant calling filters.[372] Ideally, visual inspection allows the simultaneous evaluation of multiple error features thus gaining further information compared to hard filtering. For these reasons, prediction of true variants from more likely false positive variants has been already used in scientific literature and has been included in the guidelines of the Italian Society of Human Genetics.[372, 375, 408]

Contrary to Illumina platform, more sophisticated metrics that implement basic parameters on variant calling and alignment have not been developed for the Ion Torrent platform. Such metrics could allow a more robust identification of true variants over false positive variants. This would increase confidence in the results of variant discovery and decreases analysis efforts.[372] Here, a score that combines multiple variant calling parameters is presented that allow increased sensitivity and low false positive detection rate with the Ion Torrent PGM.

6.2 Specific aims

1. To create a strategy that allows to increase the sensitivity of the Ion Torrent PGM platform.

2. To create a score that combines multiple variant calling parameters in order to increase sensitivity and reduce false positive detection rate with the Ion Torrent PGM platform.

6.3 Methods

6.3.1 Patients

We analyzed 26 patients recruited through the Registry of Steroid Resistant Nephrotic Syndrome, the Registry of Membranoproliferative Glomerulonephritis and the International Registry of atypical Hemolytic Uremic Syndrome. All participants provided informed written consent. The study was approved by the Ethics Committee of the Azienda Sanitaria Locale of Bergamo (Italy).

6.3.2 Ion Torrent PGM DNA Sequencing

Genomic DNA was extracted from peripheral blood using the NucleoSpin® Blood kit (Macherey-Nagel). Analysis of DNA regions from 112 genes was performed by highly multiplex PCR using the Ion AmpliSeq™ Library Kit 2.0 and the Ion AmpliSeq Custom WG_IAD76560 panel (subdivided into two pools), designed by the White-Glove service of Life Technologies. After clonal amplification using the Ion PGM™ Template HiQ Kit, sequencing was performed on Ion PGM Sequencer (Life Technologies). Variants were identified using TorrentSuite Software 5.2. Regions with a depth of amplicon coverage ≥10x reads were considered appropriately analyzed. Genetic variants were functionally annotated using the ANNOVAR software (June 17th 2015 release). The 1000 Genomes Project, the Exome Sequencing Project and the ExAC databases were adopted as reference to assess allelic frequencies.

6.3.3 Sequencing with other techniques

For the sequencing by the Sanger method, DNA regions were amplified by PCR amplification followed by purification using ExoSAP-IT® (GE Healthcare, Little

Chalfont, UK). The direct sequencing was performed through sequence reaction using a BigDye® terminator kit v.3.1 (Applied Biosystems, Foster City, CA, USA) followed by purification of the sequencing reaction products using the BigDye XTerminator® Purification Kit (ThermoFisher Scientific) and then run on an ABIPRISM® 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA).

Sequencing with the Illumina platform was outsourced. Exome sequencing was performed using the TruSeq Exome Enrichment Kit (Illumina) and sequenced on a HiSeq 2000 sequencer (Illumina). Target sequencing of 32 SRNS-associated genes was performed using a custom MASTR kit (Multiplicom) and sequenced on HiSeq sequencer (Illumina).

6.3.4 Statistical analyses

One-way ANOVA, Kruskal-Wallis and multivariate regression tests were used for continuous variables. Multivariate analyses included all the covariates that reached statistical significance at the univariate analysis. Statistical analyses were performed with the R software (https://cran.r-project.org/) using the glm, ggplot2 and givitiR packages. P-values <0.05 were considered statistically significant.

6.4 Results

6.4.1 Performance of the proteinuric glomerulopathy NGS panel

This study was performed in the context of a wider project that aimed to create a NGS based test that would allow the detection of SNV and indel mutations known to occur in glomerular proteinuria and nephrotic syndrome, as well as SNP alleles known to predispose to disorders associated with glomerular proteinuria and nephrotic syndrome, as detailed in Chapter 5.

6.4.2 Performance of the panel

The performance of this panel was evaluated in 11 patients with proteinuric nephropathy (6 patients with SRNS, 3 patients with MPGN and 2 patients with atypical Hemolytic Uremic Syndrome, aHUS). These patients have been previously analyzed by whole exome sequencing (n=4), Sanger sequencing of six SRNS-associated genes (n=4) or by a previously described and validated NGS genetic panel for MPGN/aHUS (n=3) (ref and chapters 3 and 4 of this thesis). Overall, the previous sequencing in the above patients identified 2664 true variants suitable for the comparison with the new *proteinuric glomerulopathy NGS panel.*

Using the default parameters for variant calling on the TSVC, 605±22 variants were identified in each sample for a total of 6655 different variants

These parameters showed an overall sensitivity of 98.2% (Table 6.2). The performance was better for SNVs, for which 98.7% of true variants were captured, compared to indel variants, for which only 91.7% of true variants were captured. Altogether, there were 47 variants (31 SNVs and 16 indels) that were not identified by TSVC. On the other hand, 5.8% of the identified variants were false positives (Table 6.2). When SNVs and indels

were analyzed independently, the false positive detection rate was 3.3% and 30.2%, respectively.

Table 6.2. Sensitivity and false positive detection rate.

Parameter profile	All		SNV		Indel	
	Sens	FDR	Sens	FDR	Sens	FDR
Default*	0.982	0.058	0.987	0.033	0.917	0.302
Custom	0.997	0.109	0.999	0.037	0.969	0.551
Combined Default & Custom	0.9996	0.139	0.9996	0.061	1.000	0.580

SNV: single nucleotide variant; indel: insertion/deletion variant; All: both SNV and indel; Sens: Sensitivity (= true positive / [true positive + false negative]); FDR: False detection rate (false positive / [true positive + false positive])
*default parameter profile provided by the TSVC 5.2 version.

To increase sensitivity, the causes leading to false negatives were investigated using the hotspot option of the TSVC and by manual inspection on IGV browser. Most of the missed variants were filtered out during hard filtering, therefore a new set of thresholds for the TSVC parameters was used to filter out false positive calls without compromising sensitivity, as reported in Table 6.3.

Table 6.3. Optimized thresholds of the TSVC parameters to increase sensitivity.

Parameter	Default	Custom
relative_strand_bias	0.7	0.85
snp_strand_bias	0.98	1.00
filter_insertion_predictions	0.30	0.40
data_quality_stringency	5.00	4.50
filter_unusual_predictions	0.30	0.70
indel_min_cov_each_strand	4.00	3.00
hotspot_strand_bias	0.98	1
hotspot_min_variant_score	10	6

With the new custom thresholds for variant calling, a mean of 646 (±26) variants were identified for each sample. The overall sensitivity improved compared to the analysis with the default parameters (99.7% *vs.* 98.2%) at the cost of more false positive variants (FDR 10.9% *vs.* 5.8%) (Table 6.2). Such performance was observed both for SNVs and indels, when they were analyzed separately. The sensitivity was 99.9% and 96.9% for SNVs and indels, respectively. The FDR was 3.7% and 55.1% for SNVs and indels, respectively.

Altogether, there were 9 variants (3 SNVs and 6 indels) that were not identified by TSVC using the custom setting. Interestingly, 8 out of these 9 false negative were correctly called with the default settings. Therefore, the two analyses are combined achieving to reduce the false positives to 1 and obtaining a sensitivity of 99.96%. However, this procedure further increased the FDR (13.9%, 6.1% and 58.0% for overall, SNV and indel, respectively) (Table 6.2).

6.4.3 Strategy to reduce false positive variants

In order to reduce the high false positive detection rate of the above analyses, it was investigated whether the combination of TSVC parameters in a probabilistic model could outperform hard filtering. At this purpose, we used the above dataset containing 2664 true variants and 441 false positive variants. Since this dataset was enriched for variants from a subgroup of genes (i.e. SRNS-associated for Sanger sequencing variants and complement disorder-associated genes for the NGS complement panel), to avoid a possible bias, all the variants with no information by an alternative technique were manually evaluated on IGV. In this way, 3576 likely true variants and 415 likely false positive variants were identified. For 73 variants, it was not possible to reach a high confidence prediction. The sensitivity and the FDR of this enlarged dataset (overall 6240 true variants and 856 false positive variant) were comparable to the above more restrict dataset (Tables 6.2 and 6.4).

Table 6.4. Sensitivity and false positive detection rate in the larger dataset containing likely true and likely false variants.

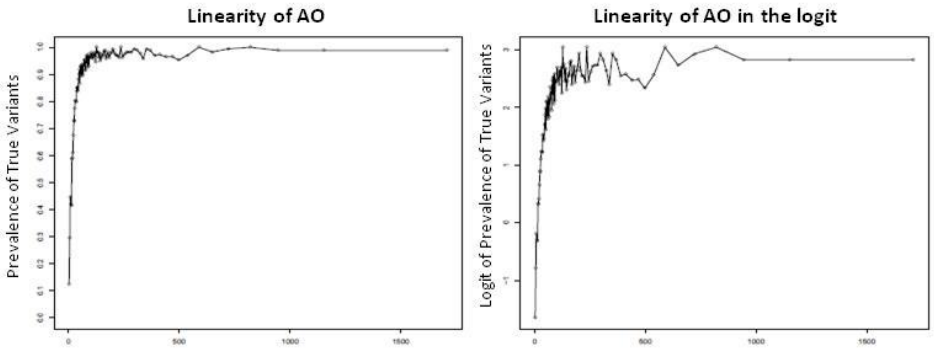
Parameter profile	All		SNV		Indel	
	Sens	FDR	Sens	FDR	Sens	FDR
Default*	0.985	0.044	0.992	0.021	0.904	0.278
Custom	0.997	0.098	0.999	0.035	0.970	0.506
Combined Default & Custom	0.9998	0.116	0.9998	0.049	1.000	0.530

SNV: single nucleotide variant; indel: insertion/deletion variant; All: both SNV and indel; Sens: Sensitivity (= true positive / [true positive + false negative]); FDR: False detection rate (false positive / [true positive + false positive])
*default parameter profile provided by the TSVC 5.2 version.

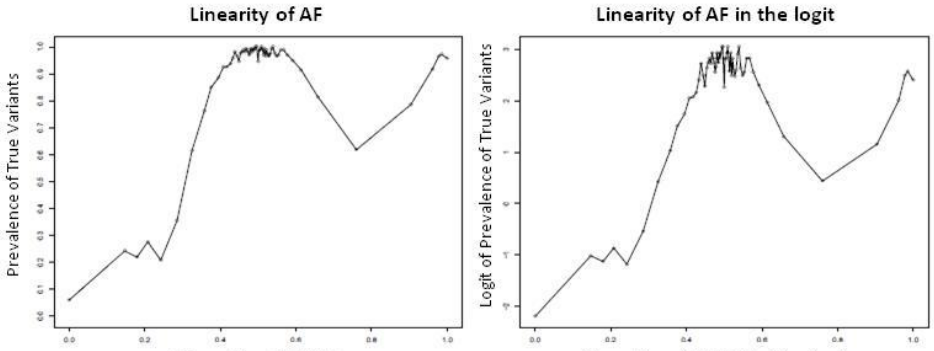
To implement the TSVC parameters in a probabilistic model, the linear or non-linear relationship between true variants and all parameters (as is and in the logit) were evaluated. Some examples are reported in Figure 6.1. Interestingly, several parameters showed linear (n=6), logarithmic (n=12) or polynomial (n=2) rather than categorical relationship (n=15). Then univariate analysis was performed. As expected, all the parameters were significantly associated with true variants separately (Table 6.5).

Figure 6.1. Linearity of the prevalence of true variants in correlation with the parameters of Torrent Suite Variant Caller that reached significance in the multivariate logistic regression. AO: Alternate allele observations; AF: Allele frequency based on Flow Evaluator observation counts; MLLD: Mean log-likelihood delta per read; SSSB: Strand-specific strand bias for allele; VARB: Variant Hypothesis bias in prediction; REVB: Reverse strand bias in prediction; RBI: Distance of bias parameters from zero; FXX: Flow Evaluator failed read ratio; STB: Strand bias in variant relative to reference; LEN: allele length; HRUN: Run length: the number of consecutive repeats of the alternate allele in the reference genome; QD: QualityByDepth as 4*QUAL/FDP (analogous to GATK).

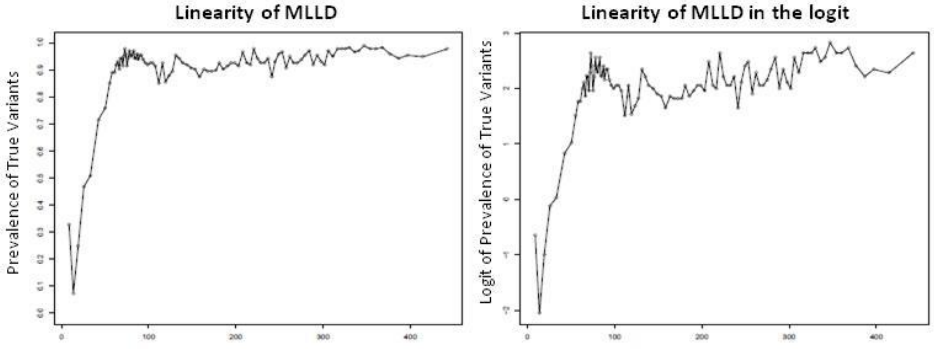
AO



AF



MLLD



SSSB

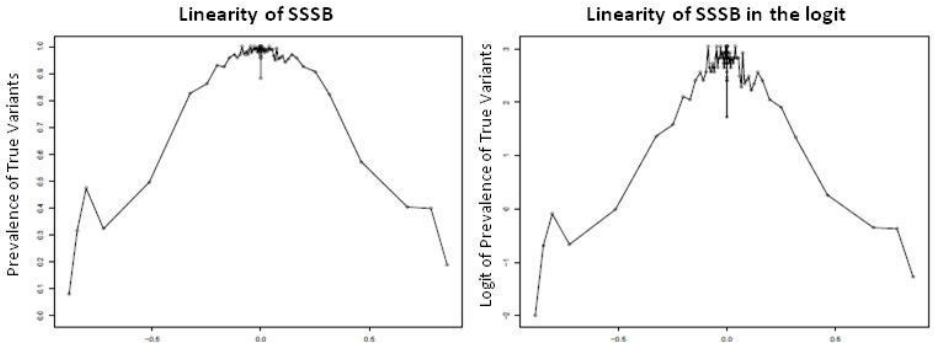


Figure 6.1. *Continue.*

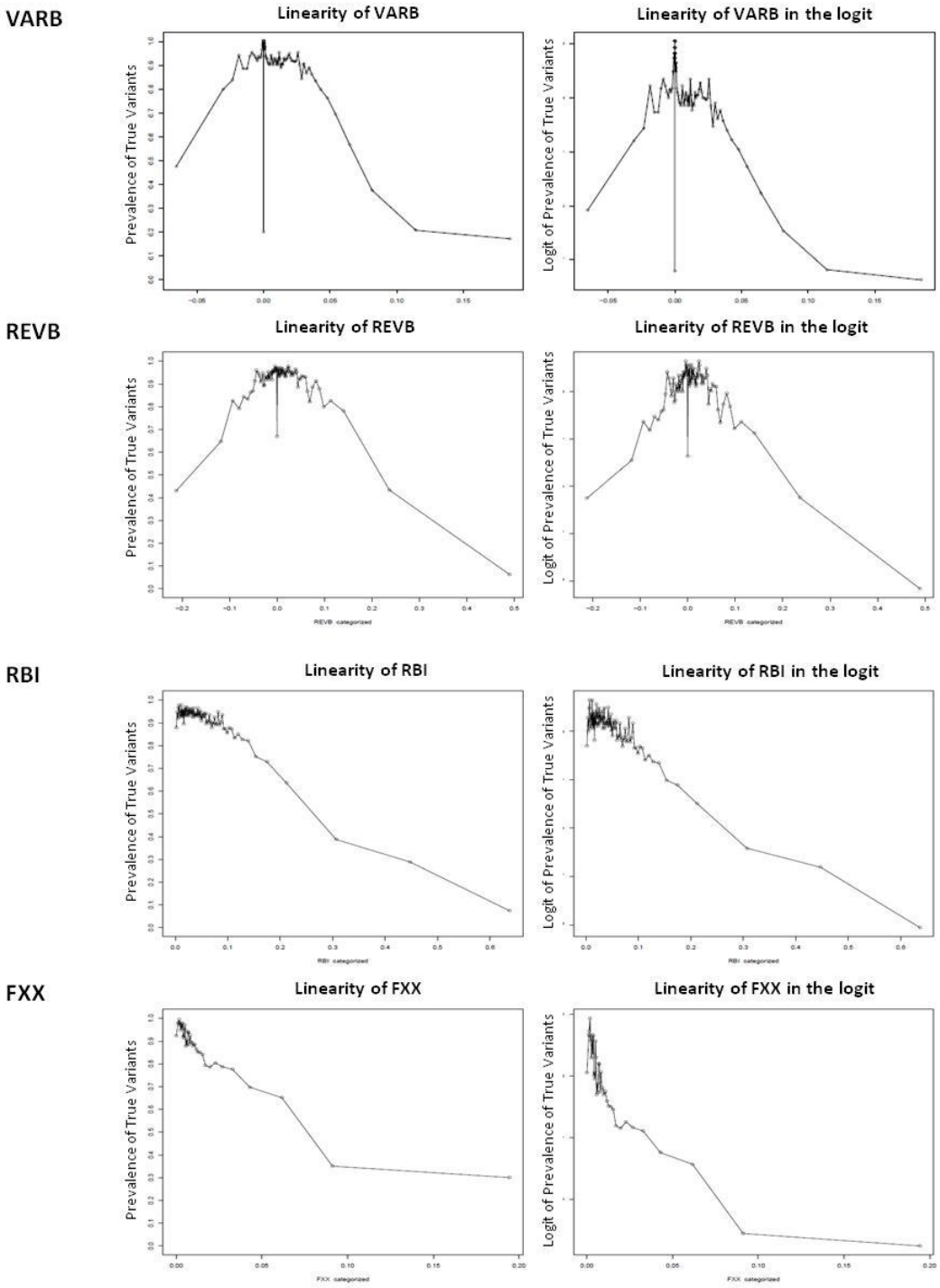


Figure 6.1. *Continue.*

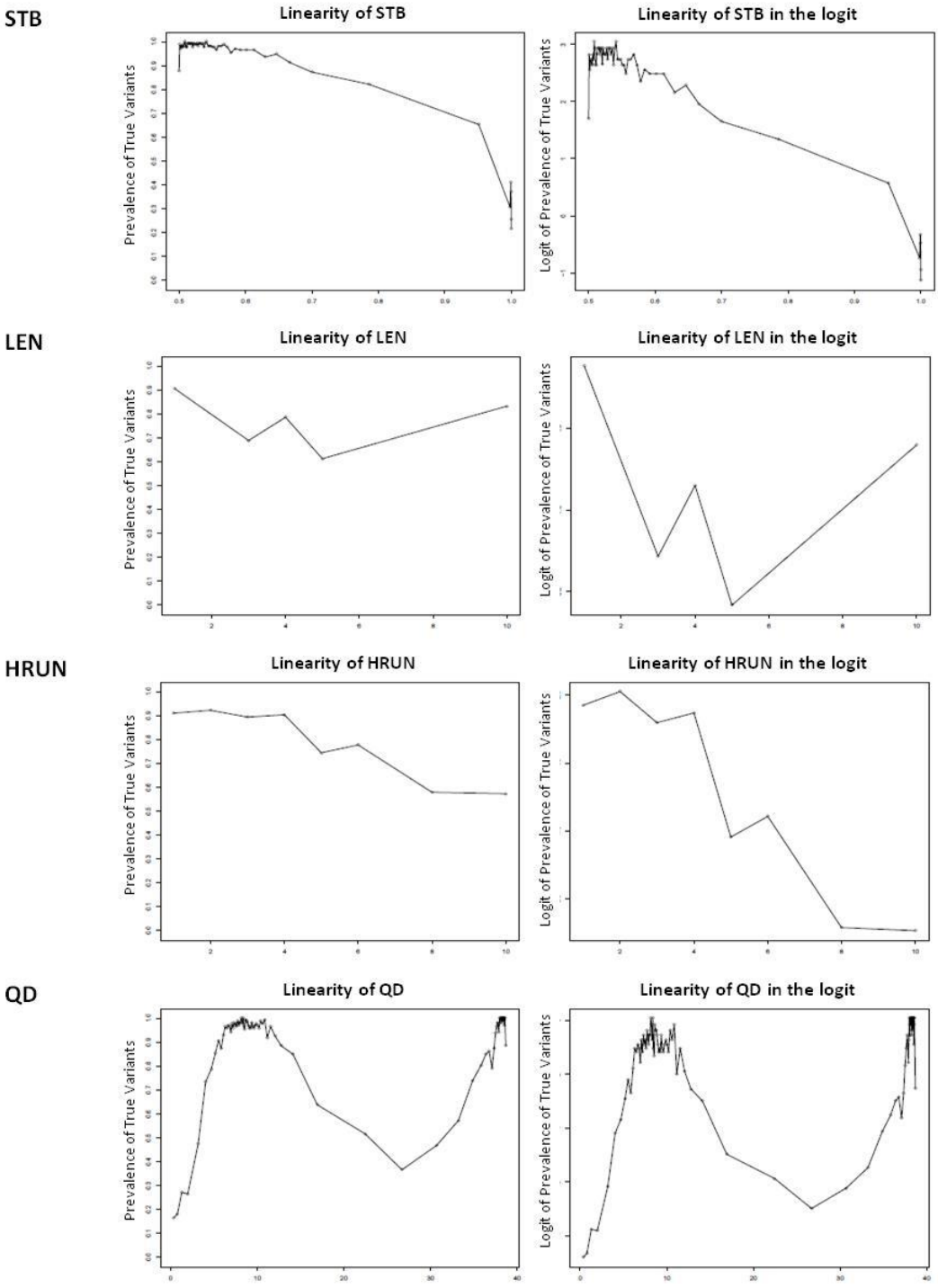


Figure 6.1. *Continue.*

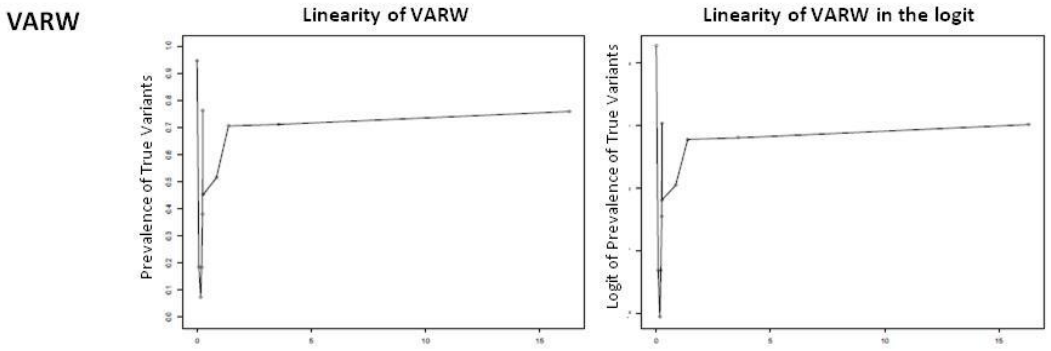


Table 6.5. Univariate analysis showing the correlation of the TSVC parameters with the true variants.

Parameters	**OR (95% CI)**	**Coeff.**	***p*-value**
$\log_{10}$(DP)	6.0 (5.1-7.0)	1.79	2.8E-114
$\log_{10}$(FDP)	6.5 (5.6-7.6)	1.87	5.8E-122
RO > 2	0.24 (0.19-0.30)	-1.43	1.2E-38
FRO > 68	2.0 (1.7-2.3)	0.70	6.4E-19
$\log_{10}$(AO)	18.2 (15.3-21.7)	2.90	4.4E-228
$\log_{10}$(FAO)	11.4 (9.8-13.4)	2.44	2.6E-204
SRF > 1	0.43 (0.36-0.51)	-0.85	9.3E-22
SRR > 55	2.0 (1.7-2.4)	0.71	4.8E-17
AF	168 (114-247)	5.12	5.04E-148
$\log_{10}$(SAF)	5.1 (4.7-5.7)	1.64	4.4E-250
$\log_{10}$(SAR)	4.8 (4.4-5.3)	1.57	2.4E-230
FSRF > 36	1.9 (1.6-2.2)	0.65	4.3E-16
$\log_{10}$(FSAF)	5.5 (5.0-6.0)	1.70	1.2E-273
FSRR > 38	1.3 (1.2-1.4)	0.27	2.4E-15
$\log_{10}$(FSAR)	4.8 (4.4-5.3)	1.57	2.5E-246
$\log_{10}$(QUAL)	7.2 (6.3-8.2)	1.97	5.0E-190
$\log_{10}$(GQ)	10.1 (8.8-11.8)	2.32	7.2E-207
REFB > -0.0536	3.4 (2.8-4.2)	1.23	1.4E-30
VARB > 0.044	0.06 (0.05-0.08)	-2.78	9.9E-160
FWDB > 0.1013	0.12 (0.09-0.15)	-2.14	4.7E-80
REVB	$8x10^{-3}$ ($4x10^{-3}$-0.016)	-4.86	1.8E-40
FXX	$4x10^{-12}$ ($2x10\text{-}13^{-1}x10^{-10}$)	-26.17	2.3E-54
RBI	$3x10^{-5}$ ($1x10^{-5}$-$7x10^{-5}$)	-10.38	1.1E-110
$\log_{10}$(MLLD)	26.1 (20.9-32.7)	3.26	1.7E-177
SSEN > 0	0.07 (0.03-0.14)	-2.69	6.9E-13
SSEP > 0	0.09 (0.06-0.15)	-2.37	1.7E-22
SSSB > -0.2083	14.7 (12.1-17.9)	2.69	9.3E-160
STB	$2x10^{-3}$ ($2x10^{-3}$-$4x10^{-3}$)	-6.04	6.4E-154
$\log_{10}$(STBP)	1.9 (1.3-2.8)	0.64	0.0016
LEN	0.91 (0.88-0.95)	-0.09	5.6E-07
HRUN	0.78 (0.75-0.81)	-0.25	7.3E-34
QD: QD < 5.1804	0.023 (0.019-0.029)	-3.74	3.4E-240
14 < QD ≤ 36	0.038 (0.031-0.048)	-3.25	6.5E-194
VARW > 0	0.06 (0.05-0.07)	-2.79	8.2E-217
GT	7.3 (6.1-8.8)	1.99	6.0E-102
Variant type: TYPE = "ins"	0.057 (0.045-0.072)	-2.86	4.1E-124
TYPE = "del"	0.024 (0.019-0.029)	-3.75	1.5E-259
TYPE = "mnp"	0.009 (0.005-0.016)	-4.74	3.0E-55

Since it is possible that there is a correlation between the different parameters, multivariate analysis in a backward stepwise procedure was performed, to identify the parameters that independently dissected true from false variants (Table 6.6). The variants included in the multivariate model were the number of reads containing the alternate allele (AO), the alternate allele frequency (AF), the mean Log-Likelihood Difference between predictions under the reference and the alternate allele hypothesis (MLLD), strand-specific strand bias (SSSB), the Euclidean distance between ideal predictions for the alternate alleles and the average measured values for the alternate alleles (VARB), the Euclidean distance between ideal predictions for the locus and the average measured values for the locus on the reverse strand (REVB), the Euclidean distance between ideal predictions for the locus and the average measured values for the locus (RBI), the proportion of reads that fail to fit any candidate variant (FXX), the strand bias measured in proportion of variant on each strand using flow evaluated allele counts (STB), the length of the alternate allele (LEN), the length of the first homopolymer in the reference that changes in length (HRUN), QUAL on DP ration (QD), variation of the width of deletions and insertions (VARW) and the type of variant (insertion, deletion or multiple nucleotide variation).

The parameters AO, MLLD, STB and LEN were associated with higher probability of the variants to be true. The parameters REVB, RBI, FXX, HRUN, VARW, QD <5.2 or within 14-36, and being a deletion, insertion or Multi Nucleotide Polymorphism (mnp) were associated with lower probability of the variants to be true. The AF showed the maximum probability of true variant for the 0.5 and 1 values and the lowest for the 0 value. Finally, variants with SSSB and VARB values around 0 had the maximum probability to be true.

Table 6.6. Multivariate analysis.

Parameter	OR (95% CI)	*p*-value	*β*
log_{10}(AO)	8.5 (6-12)	1.2E-33	2.1
Polynomial AF: AF^1	$1x10^{21}$ ($1x10^{11}$-$2x10^{31}$)	4.3E-05	48.7
AF^2	$3x10^{-27}$ ($4x10^{-34}$-$3x10^{-20}$)	5.0E-14	-61.0
AF^3	$5x10^{8}$ (1900-$2x10^{14}$)	0.002	20.1
log_{10}(MLLD)	3.3 (1.8-6.1)	1.5E-04	1.2
Polynomial SSSB: $SSSB^1$	$2x10^{5}$ (58.0-$6x10^{8}$)	0.003	12.1
$SSSB^2$	$7x10^{-44}$ ($8x10^{-54}$-$6x10^{-34}$)	1.6E-17	-99.4
Polynomial VARB: $VARB^1$	$3x10^{-16}$ ($6x10^{-22}$-$2x10^{-10}$)	1.3E-07	-35.7
$VARB^2$	$4x10^{-15}$ ($5x10^{-25}$-$4x10^{-5}$)	0.005	-33.0
REVB	$4x10^{-16}$ ($2x10^{-21}$-$6x10^{-11}$)	5.9E-09	-35.6
RBI	0.089 (0.011-0.725)	0.024	-2.4
FXX	0.010 ($2x10^{-4}$-0.419)	0.016	-4.6
STB	23.3 (2.4-223)	0.006	3.1
LEN	1.6 (1.4-1.9)	3.1E-08	0.5
HRUN	0.905 (0.836-0.979)	0.013	-0.1
QD: QD < 5.1804	0.470 (0.248-0.888)	0.020	-0.8
14 < QD ≤ 36	0.497 (0.287-0.86)	0.012	-0.7
VARW	0.952 (0.913-0.992)	0.020	-0.05
Variant type: TYPE = "ins"	0.269 (0.156-0.463)	2.2E-06	-1.3
TYPE = "del"	0.020 (0.012-0.034)	7.9E-48	-3.9
TYPE = "mnp"	0.023 (0.004-0.121)	9.2E-06	-3.8
Constant	0.024 (0.003-0.198)	5.4E-04	-3.7

The β coefficients of the logistic regression model were used to calculate the probability of being a true variant (p_{TRUE}) for each one of the 2664 true and the 441 false positive variants in the first dataset. This probability was used to identify and discard likely false positive calls. Adopting a p_{TRUE} >0.50 cut-off to accept a variant as true, the sensitivity slightly improved compared to the Default analysis (0.986 *vs.* 0.982) and the false positive variant error rare was reduced by almost 3-fold (0.022 *vs.* 0.058) (Tables 6.2 and 6.7). When a p_{TRUE} >0.10 cut-off was adopted, FDR was comparable to the Default analysis (0.054 *vs.* 0.058), while sensitivity was increased (0.996 *vs.* 0.982).

Table 6.7. Sensitivity and false positive detection rate in the combined Default and Custom dataset using the p_{TRUE}.

Analysis	All		SNV		indel	
	Sens	FDR	Sens	FDR	Sens	FDR
p_{TRUE} > 0.80	0.970	0.009	0.981	0.006	0.818	0.048
p_{TRUE} > 0.50	0.986	0.022	0.992	0.016	0.911	0.107
p_{TRUE} > 0.20	0.994	0.037	0.998	0.024	0.943	0.188
p_{TRUE} > 0.10	0.996	0.054	0.998	0.029	0.969	0.293
p_{TRUE} > 0.01	0.999	0.090	0.999	0.044	0.990	0.440
p_{TRUE} > 0.001	0.9996	0.117	0.9996	0.052	1.000	0.533

6.3.4 Frequency-based validation of p_{TRUE}

Since false positives often occur randomly, they are enriched for novel variants.[409] On the other hand, variants already reported in public databases should be enriched for true positives, since false positives are not commonly expected in these databases.[372] Based on these facts, it would be expected that p_{TRUE} is higher in common variants. To verify this hypothesis, fifteen new samples were sequenced and analyzed with the default and the custom settings. A total of 9536 variants were obtained. The p_{TRUE} was compared between novel variants and common variants that show an allelic frequency >1% in the 1000 Genome Project, Exome Sequencing project and the ExAC databases. The mean p_{TRUE} was significantly lower in novel variants compared to common variants (0.440 ± 0.48 *vs.* 0.967 ± 0.18, $p<2x10^{-16}$) (Figure 6.2).

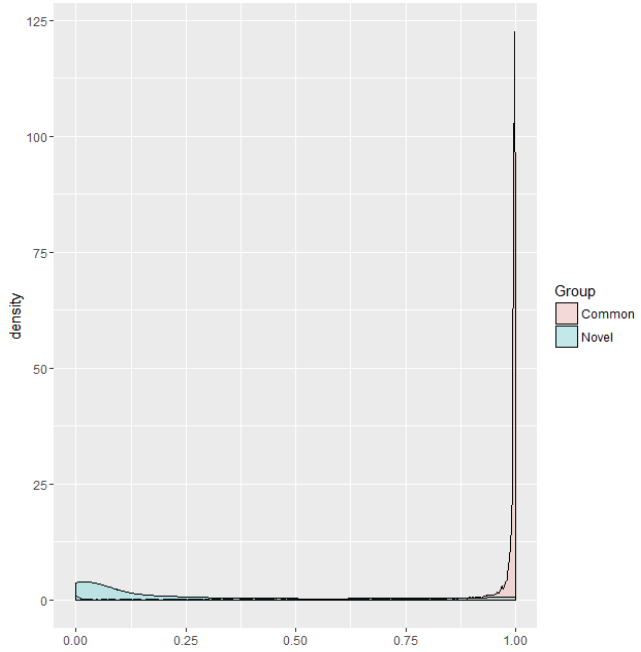


Figure 6.2. Distribution of p_{TRUE} for novel and common variants.

6.3.5 Validation of p_{TRUE} in rare variants

To further validate the strategy, 140 variants with allelic frequency <1% were selected and analyzed with another technique. The frequency criterion was adopted in order to reproduce the mutational screening setting, which is one of the most frequent applications of the target NGS panels on Ion Torrent PGM. The selected variants contained both SNV (n=100) and indels (n=40), to ensure that the approach presented here could accurately detect a spectrum of mutations. There were 86 likely true variants and 54 likely false positive variants defined on the basis of the p_{TRUE} score. Sanger sequencing and Illumina NGS, which were used as validation techniques, showed that 83 out of 86 likely true variants were really true. On the other hand, all the likely false positive variants were confirmed as false. In this confirmation dataset the sensitivity was 100% and the FDR 3.5%.

6.4 Discussion

In this study, the sensitivity and the false positive detection rate of the Ion Torrent PGM platform were evaluated. Altogether relatively low error rates are observed that, however, may affect its clinical application. Here, it is shown that the performances of the platform can be improved by replacing the hard-filtering strategy with a unique score, p_{TRUE}, that combines multiple variant calling and alignment parameters.

Using the standard protocol with the default Torrent Suite Variant Caller parameters, variants are detected with a sensitivity of 98.2% and a false detection rate of 5.8% (Table 6.2). Since accurate mutation identification is of paramount importance for confirmation of the diagnosis, risk assessment, genetic counseling, and carrier screening in patients and families affected by genetic disorders,[374] an effort was made to increase sensitivity. A higher sensitivity was obtained at the expense of more false positive variants by adopting more relaxed thresholds for the parameters of the variant calling process. Currently, the TSVC, the proprietary software commonly used to identify variants in the Ion Torrent platform, calculates a set of parameters that correlate with the variant quality.[375, 377] A set of thresholds for these parameters is defined and then, by a hard-filtering approach, variants that do not pass all the thresholds are discarded as putative false positive calls.[377] However, many of these parameters are quantitative variables and, as shown here, different values may be characterized by a different prevalence of true variants. Reducing these variants in dichotomous results in loss of valuable information. Moreover, it is possible that variants that do not fulfill multiple parameter thresholds have higher probability to be false positive callings compared to variants that do not fulfill just one.

To overcome these issues, the variant caller parameters were combined in a logistic regression probabilistic model (Tables 6.5 and 6.6). There are 14 parameters that

independently predict whether a variant is true or false. The probability derived from this model, p_{TRUE}, was used to distinguish true variants from erroneous calls, thus increasing sensitivity and reducing the error rate. In addition, since different applications may require different levels of sensitivity and error rate, it is possible to tailor the analysis by adopting different thresholds for p_{TRUE}. The score was validated with two different datasets. Firstly, the validity of p_{TRUE} is suggested by the fact that it is significantly higher in common variants compared to novel variants, since novel variants are enriched for false positives, while common variants already reported in public databases are enriched for true positives[372, 409]. Secondly, 140 true and false variants were selected and verified by an independent technique showing 100% of sensitivity and 3.5% of false detection rate.

Some limitations of the present study must be taken into account. First, the p_{TRUE} was obtained in a limited panel of genes, therefore it may not be able to separate true and false positive variants in other genes. However, the probabilistic model was trained with 1564 different types of variants derived from 112 different genes, therefore it is likely that this dataset may represent most of the variants in other genes. Moreover, the validity of the variant caller parameters used in the model is widely tested singularly both by Life Technologies and in the literature.[377, 403] Another criticism of this study is that a part of the variants included in the training dataset were obtained by distinguishing likely true from false positive variants by manual inspection of reads in the IGV. This may have introduced a bias, however, a high confidence of such predictions has been previously shown.[372, 375, 408] The good performances in the two validation datasets suggest that the introduction of few errors may be balanced by the advantage of including more variants in different genes.

In conclusion, a new score and a new approach are proposed in this study that increase

the sensitivity and reduce the false positive callings in the Ion Torrent platform. Such an approach may allow a more accurate application of the NGS in the clinical practice for the benefit of patients and researchers.

7. CONCLUSIONS

In the present thesis two groups of kidney disorders characterized by the presence of nephrotic syndrome are studied: the immune complex-mediated membranoproliferative glomerulonephritis (IC-MPGN) and C3 glomerulopathy (C3G), and the steroid-resistant nephrotic syndrome (SRNS). Genetic screening was performed and genetic data were combined with clinical, histopathologic and biochemical features to improve the knowledge in the pathogenesis of the diseases.

The main technique used to study these diseases was next-generation sequencing. In particular, the Ion Torrent PGM platform was employed. Altogether relatively low error rates are observed with this platform. However, optimization is required, since high sensitivity and low false positive detection rate are essential for clinical application of genetic tests. To improve the performances of the platform, it was created a new score that through a probabilistic model combines 14 different parameters among those currently used to identify true variants. By relaxing the thresholds in the hard-filtering strategy and using the new score to exclude putative false positives, it was possible to increase the sensitivity and reduce the false positive callings. This may allow a more accurate application of the Ion Torrent PGM platform in the clinical practice for the benefit of patients and researchers. Of note, this approach was adopted for the genetic assay used to analyze the SRNS cohort, an assay interrogating 112 different genes. In the MPGN cohort, a genetic assay of only six genes was used with relaxed filtering criteria and all the rare variants that could be involved in the pathogenesis of the disease were verified by Sanger sequencing.

In the first part of this thesis focused on IC-MPGN/C3G, it is shown that mutations in the complement alternative pathway genes *C3*, *CFB*, *CFH*, *CD46* and *CFI* account for a significant part of IC-MPGN/C3G patients in accordance with previous studies.[256, 262] In addition, two rare variants in the *THBD* gene in two unrelated C3G patients are reported in this study. These variants are already described as mutations in patients affected by atypical hemolytic uremic syndrome, an allelic disorder of IC-MPGN/C3G.[309] Moreover, functional studies *in vitro* have previously shown that these two *THBD* variants cause impaired FI-mediated C3b inactivation leading to defective suppression of the complement alternative pathway.[309] Altogether these data support that pathogenic variants in *THBD* are involved in the pathogenesis of C3G.

Another interesting finding from the study of the IC-MPGN/C3G cohort is that the presence of mutations alone does not significantly increase the risk of developing IC-MPGN or C3G. Indeed, the risk of developing IC-MPGN or C3G strongly increases only when mutations are combined with common susceptibility variants, while the risk of developing the disease is comparable between subjects carrying mutations but no susceptibility variants and subjects lacking both mutations and susceptibility variants. These data underline that IC-MPGN and C3G are complex genetic diseases.

Remarkably, the majority of IC-MPGN patients present alternative pathway dysregulation in this study. The prevalence of alternative pathway abnormalities (mutations, C3NeFs, low C3 levels with normal C4) was comparable in IC-MPGN and C3G. In particular, 56% and 65% of IC-MPGN and C3G patients, respectively, carried mutations and/or C3NeFs. These findings, which are consistent with a previous report,[256] suggest that IC-MPGN and C3G have more commonalities than previously recognized, and do not support the current classification that restricts the definition of alternative pathway-mediated glomerulopathies to patients with "dominant C3"

glomerular staining.[253] Moreover, in accordance with previous findings,[255, 256] in this study the different histologic groups have similar risk of developing ESRD. On the other hand, it is shown here that patients without complement gene mutations in the six analyzed genes and without C3NeFs have a higher risk of developing ESRD than patients carrying mutations and/or C3NeFs. This observation may indicate the existence of at least two different pathogenic mechanisms of renal disease progression in IC-MPGN/C3G.

To better understand the pathogenesis of the diseases, the presence of more homogeneous subgroups within IC-MPGN/C3G was investigated through unsupervised hierarchical clustering. This is a statistical approach where the classification criteria are not defined by the operator based on *a priori* hypotheses but arise from the data. Both etiologic features, such as presence of C3NeF or complement alternative pathway gene mutations, and phenotypic features, such as histologic findings, serum complement profile and clinical features, were included in the analysis. Four relatively homogeneous subgroups (clusters) were identified. The first 3 clusters were characterized by high prevalence of mutations and/or C3NeF contrary to cluster 4. Clusters 1 and 2 both showed the highest prevalence of subendothelial deposits and the highest SC5b-9 levels, whereas cluster 2 was distinguished from cluster 1 for the stronger C1q and IgG staining, and for the higher prevalence of nephrotic syndrome at onset. Cluster 3 showed the highest prevalence of intramembranous electron-dense deposits. Cluster 4 had the highest serum C3 levels and the highest risk of developing ESRD. Altogether, these clusters of IC-MPGN/C3G patients seem to reflect the underlying molecular mechanisms and are useful to predict the risk of developing ESRD.

In the second part of this thesis, the genetic causes of SRNS were investigated. The major role of *COL4A3-5* mutations in the pathogenesis of adult-onset SRNS and/or

FSGS, which was reported in a previous study,[210] is confirmed. Moreover, this study shows that mutations in CAKUT-associated genes are enriched in SRNS patients. In particular, for the first time to the best of my knowledge, likely pathogenic (LP) variants in *CHD1L*, *RET*, *FREM2*, *FOXC2*, *ROBO2* and *SALL1* in patients with isolated SRNS are reported. This finding is not surprising, if it is considered that mutations in *PAX2* and *LMX1B* cause both CAKUT and isolated SRNS.[221, 410] In addition, LP variants in CAKUT-associated genes combine with LP variants in SRNS-associated genes more frequently than would be expected by chance. These data suggest that the molecular pathways underlying CAKUT and SRNS partially overlap and that a complex genetic pathogenesis should be considered at least in some patients.

Finally, a new candidate gene for SRNS is proposed, namely *EPB41L5*. They protein encoded by this gene is a component of the mammalian Crumbs complex, which drives the apical-basal cell polarity.[366] EPB41L5 physically interacts with CRB2, a core protein of the Crumbs complex that is a known SRNS-causing gene.[235, 236, 367] Moreover, *EPB41L5* is expressed in zebrafish podocytes and it is required for proper formation and maintenance of slit-diaphragms, since *EPB41L5* deficient zebrafishes present absence of slit-diaphragms and loss of glomerular barrier permselectivity.[364] In three unrelated patients of the present SRNS cohort, three possibly pathogenic variants of *EPB41L5* affecting evolutionary conserved amino acids were identified and at least two them affect the FERM domain, which is crucial for the protein function.[366] Analysis of further SRNS patients and functional studies on the identified variants are necessary to disclose the role of *EPB41L5*.

In conclusion, this thesis provides new insights in IC-MPGN/C3G and SRNS. It introduces new players in the pathogenesis of these diseases. Moreover, in IC-MPGN/C3G provides evidence of the presence of subgroups of patients with common

underlying mechanisms as well as common clinical and histopathologic features. Therefore, the present thesis may be a contribution to better understand the pathogenesis of IC-MPGN/C3G and SRNS, and may be useful in future basic and clinical research.

REFERENCES

1. Tryggvason K, Wartiovaara J. How does the kidney filter plasma? Physiology (Bethesda) 2005;20:96-101.
2. Haraldsson B, et al. Properties of the glomerular barrier and mechanisms of proteinuria. Physiol Rev 2008;88:451-87.
3. Kurts C, et al. The immune system and kidney disease: basic concepts and clinical implications. Nat Rev Immunol 2013;13:738-53.
4. Leeuwis JW, et al. Targeting podocyte-associated diseases. Adv Drug Deliv Rev 2010;62:1325-36.
5. Salmon AH, et al. New aspects of glomerular filtration barrier structure and function: five layers (at least) not three. Curr Opin Nephrol Hypertens 2009;18:197-205.
6. Rostgaard J, Qvortrup K. Sieve plugs in fenestrae of glomerular capillaries--site of the filtration barrier? Cells Tissues Organs 2002;170:132-8.
7. Savage CO. The biology of the glomerulus: endothelial cells. Kidney Int 1994;45:314-9.
8. Kuligowski MP, et al. Leukocyte recruitment to the inflamed glomerulus: a critical role for platelet-derived P-selectin in the absence of rolling. J Immunol 2006;176:6991-9.
9. Noris M, et al. STEC-HUS, atypical HUS and TTP are all diseases of complement activation. Nat Rev Nephrol 2012;8:622-33.
10. Miner JH. The glomerular basement membrane. Exp Cell Res 2012;318:973-8.
11. Yurchenco PD, Patton BL. Developmental and pathogenic mechanisms of basement membrane assembly. Curr Pharm Des 2009;15:1277-94.
12. Deen WM, et al. Structural determinants of glomerular permeability. Am J Physiol Renal Physiol 2001;281:F579-96.
13. St John PL, Abrahamson DR. Glomerular endothelial cells and podocytes jointly synthesize laminin-1 and -11 chains. Kidney Int 2001;60:1037-46.
14. Goldberg S, et al. Glomerular filtration is normal in the absence of both agrin and perlecan-heparan sulfate from the glomerular basement membrane. Nephrol Dial Transplant 2009;24:2044-51.
15. Utriainen A, et al. Structurally altered basement membranes and hydrocephalus in a type XVIII collagen deficient mouse line. Hum Mol Genet 2004;13:2089-99.
16. Kanwar YS, et al. Increased permeability of the glomerular basement membrane to ferritin after removal of glycosaminoglycans (heparan sulfate) by enzyme digestion. J Cell Biol 1980;86:688-93.
17. Kim BS, Goligorsky MS. Role of VEGF in kidney development, microvascular maintenance and pathophysiology of renal disease. Korean J Intern Med 2003;18:65-75.
18. Farquhar MG. The glomerular basement membrane: not gone, just forgotten. J Clin Invest 2006;116:2090-3.
19. Butkowski RJ, et al. Basement membrane collagen in the kidney: regional localization of novel chains related to collagen IV. Kidney Int 1989;35:1195-202.
20. Hudson BG, et al. Alport's syndrome, Goodpasture's syndrome, and type IV collagen. N Engl J Med 2003;348:2543-56.
21. Kashtan CE. Alport Syndrome and Thin Basement Membrane Nephropathy.

1993.
22. Hausmann R, et al. The glomerular filtration barrier function: new concepts. Curr Opin Nephrol Hypertens 2012;21:441-9.
23. Shankland SJ. The podocyte's response to injury: role in proteinuria and glomerulosclerosis. Kidney Int 2006;69:2131-47.
24. Dressler GR. The cellular basis of kidney development. Annu Rev Cell Dev Biol 2006;22:509-29.
25. Faul C, et al. Actin up: regulation of podocyte structure and function by components of the actin cytoskeleton. Trends Cell Biol 2007;17:428-37.
26. Mundel P, Reiser J. Proteinuria: an enzymatic disease of the podocyte? Kidney Int;77:571-80.
27. Ronco P. Proteinuria: is it all in the foot? J Clin Invest 2007;117:2079-82.
28. Kreidberg JA, et al. Alpha 3 beta 1 integrin has a crucial role in kidney and lung organogenesis. Development 1996;122:3537-47.
29. Raats CJ, et al. Expression of agrin, dystroglycan, and utrophin in normal renal tissue and in experimental glomerulopathies. Am J Pathol 2000;156:1749-65.
30. Drenckhahn D, Franke RP. Ultrastructural organization of contractile and cytoskeletal proteins in glomerular podocytes of chicken, rat, and man. Lab Invest 1988;59:673-82.
31. Dai C, et al. Essential role of integrin-linked kinase in podocyte biology: Bridging the integrin and slit diaphragm signaling. J Am Soc Nephrol 2006;17:2164-75.
32. Sachs N, et al. Kidney failure in mice lacking the tetraspanin CD151. J Cell Biol 2006;175:33-9.
33. Pozzi A, Zent R. Integrins in kidney disease. J Am Soc Nephrol 2013;24:1034-9.
34. El-Aouni C, et al. Podocyte-specific deletion of integrin-linked kinase results in severe glomerular basement membrane alterations and progressive glomerulosclerosis. J Am Soc Nephrol 2006;17:1334-44.
35. Wei C, et al. Modification of kidney barrier function by the urokinase receptor. Nat Med 2008;14:55-63.
36. Regele HM, et al. Glomerular expression of dystroglycans is reduced in minimal change nephrosis but not in focal segmental glomerulosclerosis. J Am Soc Nephrol 2000;11:403-12.
37. Larrucea S, et al. Expression of podocalyxin enhances the adherence, migration, and intercellular communication of cells. Exp Cell Res 2008;314:2004-15.
38. Doyonnas R, et al. Anuria, omphalocele, and perinatal lethality in mice lacking the CD34-related protein podocalyxin. J Exp Med 2001;194:13-27.
39. Takeda T, et al. Loss of glomerular foot processes is associated with uncoupling of podocalyxin from the actin cytoskeleton. J Clin Invest 2001;108:289-301.
40. Thomas PE, et al. GLEPP1, a renal glomerular epithelial cell (podocyte) membrane protein-tyrosine phosphatase. Identification, molecular cloning, and characterization in rabbit. J Biol Chem 1994;269:19953-62.
41. Ozaltin F, et al. Disruption of PTPRO causes childhood-onset nephrotic syndrome. Am J Hum Genet;89:139-47.
42. Tryggvason K, et al. Hereditary proteinuria syndromes and mechanisms of proteinuria. N Engl J Med 2006;354:1387-401.
43. Reiser J, et al. The glomerular slit diaphragm is a modified adherens junction. J Am Soc Nephrol 2000;11:1-8.
44. Fukasawa H, et al. Slit diaphragms contain tight junction proteins. J Am Soc Nephrol 2009;20:1491-503.
45. Lavin PJ, et al. Therapeutic targets in focal and segmental glomerulosclerosis. Curr Opin Nephrol Hypertens 2008;17:386-92.
46. Chuang PY, He JC. Signaling in regulation of podocyte phenotypes. Nephron

Physiol 2009;111:p9-15.
47. Kwoh C, et al. Pathogenesis of nonimmune glomerulopathies. Annu Rev Pathol 2006;1:349-74.
48. Patrakka J, Tryggvason K. Nephrin--a unique structural and signaling protein of the kidney filter. Trends Mol Med 2007;13:396-403.
49. Kerjaschki D. Caught flat-footed: podocyte damage and the molecular bases of focal glomerulosclerosis. J Clin Invest 2001;108:1583-7.
50. Schwarz K, et al. Podocin, a raft-associated component of the glomerular slit diaphragm, interacts with CD2AP and nephrin. J Clin Invest 2001;108:1621-9.
51. Antignac C. Genetic models: clues for understanding the pathogenesis of idiopathic nephrotic syndrome. J Clin Invest 2002;109:447-9.
52. Patrakka J, Tryggvason K. New insights into the role of podocytes in proteinuria. Nat Rev Nephrol 2009;5:463-8.
53. Garg P, et al. Neph1 cooperates with nephrin to transduce a signal that induces actin polymerization. Mol Cell Biol 2007;27:8698-712.
54. Donoviel DB, et al. Proteinuria and perinatal lethality in mice lacking NEPH1, a novel protein with homology to NEPHRIN. Mol Cell Biol 2001;21:4829-36.
55. Nilius B, et al. Transient receptor potential cation channels in disease. Physiol Rev 2007;87:165-217.
56. Spassova MA, et al. A common mechanism underlies stretch activation and receptor activation of TRPC6 channels. Proc Natl Acad Sci U S A 2006;103:16586-91.
57. Moller CC, et al. Sensitizing the Slit Diaphragm with TRPC6 ion channels. J Am Soc Nephrol 2009;20:950-3.
58. Hisatsune C, et al. Regulation of TRPC6 channel activity by tyrosine phosphorylation. J Biol Chem 2004;279:18887-94.
59. Faul C, et al. The actin cytoskeleton of kidney podocytes is a direct target of the antiproteinuric effect of cyclosporine A. Nat Med 2008;14:931-8.
60. Verma R, et al. Fyn binds to and phosphorylates the kidney slit diaphragm component Nephrin. J Biol Chem 2003;278:20716-23.
61. Yaoita E, et al. Role of Fat1 in cell-cell contact formation of podocytes in puromycin aminonucleoside nephrosis and neonatal kidney. Kidney Int 2005;68:542-51.
62. Ciani L, et al. Mice lacking the giant protocadherin mFAT1 exhibit renal slit junction abnormalities and a partially penetrant cyclopia and anophthalmia phenotype. Mol Cell Biol 2003;23:3575-82.
63. Gee HY, et al. FAT1 mutations cause a glomerulotubular nephropathy. Nat Commun 2016;7:10822.
64. Verma R, et al. Nephrin ectodomain engagement results in Src kinase activation, nephrin phosphorylation, Nck recruitment, and actin polymerization. J Clin Invest 2006;116:1346-59.
65. Jones N, et al. Nck adaptor proteins link nephrin to the actin cytoskeleton of kidney podocytes. Nature 2006;440:818-23.
66. Shih NY, et al. Congenital nephrotic syndrome in mice lacking CD2-associated protein. Science 1999;286:312-5.
67. Benzing T. The promise of well-being: stay in shape with N(i)ck. J Am Soc Nephrol 2009;20:1425-7.
68. Jones N, et al. Nck proteins maintain the adult glomerular filtration barrier. J Am Soc Nephrol 2009;20:1533-43.
69. Simons M, et al. Podocyte polarity signalling. Curr Opin Nephrol Hypertens 2009;18:324-30.
70. Hartleben B, et al. Neph-Nephrin proteins bind the Par3-Par6-atypical protein kinase C (aPKC) complex to regulate podocyte cell polarity. Journal of Biological Chemistry 2008;283:23033-8.

71. Schlondorff D, Banas B. The mesangial cell revisited: no cell is an island. J Am Soc Nephrol 2009;20:1179-87.
72. Kreisberg JI, et al. Contractile properties of cultured glomerular mesangial cells. Am J Physiol 1985;249:F457-63.
73. Schlondorff D. The glomerular mesangial cell: an expanding role for a specialized pericyte. Faseb J 1987;1:272-81.
74. Mene P, et al. Physiology of the mesangial cell. Physiol Rev 1989;69:1347-424.
75. Ohyama K, et al. Extracellular matrix phenotype of rat mesangial cells in culture. Biosynthesis of collagen types I, III, IV, and V and a low molecular weight collagenous component and their regulation by dexamethasone. J Lab Clin Med 1990;116:219-27.
76. Abboud HE, et al. Production of platelet-derived growth factorlike protein by rat mesangial cells in culture. J Clin Invest 1987;80:675-83.
77. Aron DC, et al. Synthesis and binding of insulin-like growth factor I by human glomerular mesangial cells. J Clin Endocrinol Metab 1989;68:585-91.
78. Zoja C, et al. Interleukin-1 beta and tumor necrosis factor-alpha induce gene expression and production of leukocyte chemotactic factors, colony-stimulating factors, and interleukin-6 in human mesangial cells. Am J Pathol 1991;138:991-1003.
79. Julian BA, Novak J. IgA nephropathy: an update. Curr Opin Nephrol Hypertens 2004;13:171-9.
80. Lai AS, Lai KN. Molecular basis of IgA nephropathy. Curr Mol Med 2005;5:475-87.
81. Iatropoulos P, et al. Complement gene variants determine the risk of immunoglobulin-associated MPGN and C3 glomerulopathy and predict long-term renal outcome. Mol Immunol 2016;71:131-42.
82. Ohse T, et al. The enigmatic parietal epithelial cell is finally getting noticed: a review. Kidney Int 2009;76:1225-38.
83. Appel D, et al. Recruitment of podocytes from glomerular parietal epithelial cells. J Am Soc Nephrol 2009;20:333-43.
84. Ronconi E, et al. Regeneration of glomerular podocytes by human renal progenitors. J Am Soc Nephrol 2009;20:322-32.
85. Noris M, Remuzzi G. Glomerular Diseases Dependent on Complement Activation, Including Atypical Hemolytic Uremic Syndrome, Membranoproliferative Glomerulonephritis, and C3 Glomerulopathy: Core Curriculum 2015. Am J Kidney Dis 2015;66:359-75.
86. Sarma JV, Ward PA. The complement system. Cell Tissue Res 2011;343:227-35.
87. Amara U, et al. Interaction between the coagulation and complement system. Adv Exp Med Biol 2008;632:71-9.
88. Noris M, Remuzzi G. Overview of complement activation and regulation. Semin Nephrol 2013;33:479-92.
89. Merle NS, et al. Complement System Part II: Role in Immunity. Front Immunol 2015;6:257.
90. Valoti E, et al. A novel atypical hemolytic uremic syndrome-associated hybrid CFHR1/CFH gene encoding a fusion protein that antagonizes factor H-dependent complement regulation. J Am Soc Nephrol 2015;26:209-19.
91. Goicoechea de Jorge E, et al. Dimerization of complement factor H-related proteins modulates complement activation in vivo. Proc Natl Acad Sci U S A 2013;110:4685-90.
92. Tschopp J, et al. Clusterin, the human apolipoprotein and complement inhibitor, binds to complement C7, C8 beta, and the b domain of C9. J Immunol 1993;151:2159-65.

93. Hadders MA, et al. Assembly and regulation of the membrane attack complex based on structures of C5b6 and sC5b9. Cell Rep 2012;1:200-7.
94. Wurzner R. Modulation of complement membrane attack by local C7 synthesis. Clin Exp Immunol 2000;121:8-10.
95. Noris M, Remuzzi G. Atypical hemolytic–uremic syndrome. New England Journal of Medicine 2009;361:1676-87.
96. Longo DH, TR. Harrison's principles of internal medicine. 2012.
97. Trautmann A, et al. Spectrum of steroid-resistant and congenital nephrotic syndrome in children: the PodoNet registry cohort. Clin J Am Soc Nephrol 2015;10:592-600.
98. Cadnapaphornchai MA, et al. The nephrotic syndrome: pathogenesis and treatment of edema formation and secondary complications. Pediatr Nephrol 2014;29:1159-67.
99. Ordonez JD, et al. The increased risk of coronary heart disease associated with nephrotic syndrome. Kidney Int 1993;44:638-42.
100. Uncu N, et al. Primary peritonitis in children with nephrotic syndrome: results of a 5-year multicenter study. Eur J Pediatr 2010;169:73-6.
101. Kaysen GA. Plasma composition in the nephrotic syndrome. Am J Nephrol 1993;13:347-59.
102. Tsimihodimos V, et al. Dyslipidemia associated with chronic kidney disease. Open Cardiovasc Med J 2011;5:41-8.
103. Eddy AA, Symons JM. Nephrotic syndrome in childhood. Lancet 2003;362:629-39.
104. Barisoni L, et al. A proposed taxonomy for the podocytopathies: a reassessment of the primary nephrotic diseases. Clin J Am Soc Nephrol 2007;2:529-42.
105. Kodner C. Nephrotic syndrome in adults: diagnosis and management. Am Fam Physician 2009;80:1129-34.
106. Barnes PJ. Mechanisms and resistance in glucocorticoid control of inflammation. J Steroid Biochem Mol Biol;120:76-85.
107. Meyrier A. Treatment of focal segmental glomerulosclerosis. Expert Opin Pharmacother 2005;6:1539-49.
108. Musante L, et al. Humoral permeability factors in the nephrotic syndrome: a compendium and prospectus. J Nephrol 2001;14 Suppl 4:S48-50.
109. Schulman SL, et al. Predicting the response to cytotoxic therapy for childhood nephrotic syndrome: superiority of response to corticosteroid therapy over histopathologic patterns. J Pediatr 1988;113:996-1001.
110. Gbadegesin R, et al. Pathogenesis and therapy of focal segmental glomerulosclerosis: an update. Pediatr Nephrol 2011;26:1001-15.
111. Gbadegesin R, et al. Pathogenesis and therapy of focal segmental glomerulosclerosis: an update. Pediatr Nephrol;26:1001-15.
112. Segarra A, et al. Combined therapy of tacrolimus and corticosteroids in cyclosporin-resistant or -dependent idiopathic focal glomerulosclerosis: a preliminary uncontrolled study with prospective follow-up. Nephrol Dial Transplant 2002;17:655-62.
113. Tejani A, Ingulli E. Current concepts of pathogenesis of nephrotic syndrome. Contrib Nephrol 1995;114:1-5.
114. Neuhaus TJ, et al. Alternative treatment to corticosteroids in steroid sensitive idiopathic nephrotic syndrome. Arch Dis Child 1994;71:522-6.
115. Ueda N, et al. Beneficial effect of chlorambucil in steroid-dependent and cyclophosphamide-resistant minimal change nephrotic syndrome. J Nephrol 2009;22:610-5.
116. Bagga A, Mantan M. Nephrotic syndrome in children. Indian J Med Res

2005;122:13-28.
117. Sellier-Leclerc AL, et al. Rituximab efficiency in children with steroid-dependent nephrotic syndrome. Pediatr Nephrol;25:1109-15.
118. Trachtman H, Gauthier B. Effect of angiotensin-converting enzyme inhibitor therapy on proteinuria in children with renal disease. J Pediatr 1988;112:295-8.
119. Milliner DS, Morgenstern BZ. Angiotensin converting enzyme inhibitors for reduction of proteinuria in children with steroid-resistant nephrotic syndrome. Pediatr Nephrol 1991;5:587-90.
120. Delucchi A, et al. Enalapril and prednisone in children with nephrotic-range proteinuria. Pediatr Nephrol 2000;14:1088-91.
121. Arneil GC, Lam CN. Long-term assessment of steroid therapy in childhood nephrosis. Lancet 1966;2:819-21.
122. Primary nephrotic syndrome in children: clinical significance of histopathologic variants of minimal change and of diffuse mesangial hypercellularity. A Report of the International Study of Kidney Disease in Children. Kidney Int 1981;20:765-71.
123. Takeda A, et al. Risk factors for relapse in childhood nephrotic syndrome. Pediatr Nephrol 1996;10:740-1.
124. Yap HK, et al. Risk factors for steroid dependency in children with idiopathic nephrotic syndrome. Pediatr Nephrol 2001;16:1049-52.
125. Schachter AD, Harmon WE. Single-center analysis of early recurrence of nephrotic syndrome following renal transplantation in children. Pediatr Transplant 2001;5:406-9.
126. Salomon R, et al. Intravenous cyclosporine therapy in recurrent nephrotic syndrome after renal transplantation in children. Transplantation 2003;75:810-4.
127. Ruf RG, et al. Patients with mutations in NPHS2 (podocin) do not respond to standard steroid treatment of nephrotic syndrome. J Am Soc Nephrol 2004;15:722-32.
128. Weber S, et al. NPHS2 mutation analysis shows genetic heterogeneity of steroid-resistant nephrotic syndrome and low post-transplant recurrence. Kidney Int 2004;66:571-9.
129. Bertelli R, et al. Recurrence of focal segmental glomerulosclerosis after renal transplantation in patients with mutations of podocin. Am J Kidney Dis 2003;41:1314-21.
130. Orth SR, Ritz E. The nephrotic syndrome. N Engl J Med 1998;338:1202-11.
131. Bonilla-Felix M, et al. Changing patterns in the histopathology of idiopathic nephrotic syndrome in children. Kidney Int 1999;55:1885-90.
132. Korbet SM, et al. The racial prevalence of glomerular lesions in nephrotic adults. Am J Kidney Dis 1996;27:647-51.
133. Haas M, et al. Changing etiologies of unexplained adult nephrotic syndrome: a comparison of renal biopsy findings from 1976-1979 and 1995-1997. Am J Kidney Dis 1997;30:621-31.
134. Floege J, Amann K. Primary glomerulonephritides. Lancet 2016;387:2036-48.
135. McGrogan A, et al. The incidence of primary glomerulonephritis worldwide: a systematic review of the literature. Nephrol Dial Transplant 2011;26:414-30.
136. Cameron JS. Nephrotic syndrome in the elderly. Semin Nephrol 1996;16:319-29.
137. D'Agati VD, et al. Pathologic classification of focal segmental glomerulosclerosis: a working proposal. Am J Kidney Dis 2004;43:368-82.
138. Thomas DB, et al. Clinical and pathologic characteristics of focal segmental glomerulosclerosis pathologic variants. Kidney Int 2006;69:920-6.
139. D'Agati VD, et al. Focal segmental glomerulosclerosis. N Engl J Med 2011;365:2398-411.
140. Chun MJ, et al. Focal segmental glomerulosclerosis in nephrotic adults:

presentation, prognosis, and response to therapy of the histologic variants. J Am Soc Nephrol 2004;15:2169-77.
141. Shiiki H, Dohi K. Primary focal segmental glomerulosclerosis: clinical course, predictors of renal outcome and treatment. Intern Med 2000;39:606-11.
142. Ponticelli C, Passerini P. Treatment of the nephrotic syndrome associated with primary glomerulonephritis. Kidney Int 1994;46:595-604.
143. Kestila M, et al. Positionally cloned gene for a novel glomerular protein--nephrin--is mutated in congenital nephrotic syndrome. Mol Cell 1998;1:575-82.
144. Vivante A, Hildebrandt F. Exploring the genetic basis of early-onset chronic kidney disease. Nat Rev Nephrol 2016;12:133-46.
145. Piscione TD, Licht C. Genetics of proteinuria: an overview of gene mutations associated with nonsyndromic proteinuric glomerulopathies. Adv Chronic Kidney Dis 2011;18:273-89.
146. D'Agati VD. Pathobiology of focal segmental glomerulosclerosis: new developments. Curr Opin Nephrol Hypertens 2012;21:243-50.
147. Hinkes BG, et al. Nephrotic syndrome in the first year of life: two thirds of cases are caused by mutations in 4 genes (NPHS1, NPHS2, WT1, and LAMB2). Pediatrics 2007;119:e907-19.
148. Heeringa SF, et al. Thirteen novel NPHS1 mutations in a large cohort of children with congenital nephrotic syndrome. Nephrol Dial Transplant 2008;23:3527-33.
149. Schoeb DS, et al. Nineteen novel NPHS1 mutations in a worldwide cohort of patients with congenital nephrotic syndrome (CNS). Nephrol Dial Transplant 2010;25:2970-6.
150. Santin S, et al. Clinical utility of genetic testing in children and adults with steroid-resistant nephrotic syndrome. Clin J Am Soc Nephrol 2011;6:1139-48.
151. Philippe A, et al. Nephrin mutations can cause childhood-onset steroid-resistant nephrotic syndrome. J Am Soc Nephrol 2008;19:1871-8.
152. Sadowski CE, et al. A single-gene cause in 29.5% of cases of steroid-resistant nephrotic syndrome. J Am Soc Nephrol 2015;26:1279-89.
153. Buscher AK, et al. Mutations in podocyte genes are a rare cause of primary FSGS associated with ESRD in adult patients. Clin Nephrol 2012;78:47-53.
154. Megremis S, et al. Nucleotide variations in the NPHS2 gene in Greek children with steroid-resistant nephrotic syndrome. Genet Test Mol Biomarkers 2009;13:249-56.
155. Caridi G, et al. Prevalence, genetics, and clinical features of patients carrying podocin mutations in steroid-resistant nonfamilial focal segmental glomerulosclerosis. J Am Soc Nephrol 2001;12:2742-6.
156. Hinkes B, et al. Specific podocin mutations correlate with age of onset in steroid-resistant nephrotic syndrome. J Am Soc Nephrol 2008;19:365-71.
157. Berdeli A, et al. NPHS2 (podicin) mutations in Turkish children with idiopathic nephrotic syndrome. Pediatr Nephrol 2007;22:2031-40.
158. Bouchireb K, et al. NPHS2 mutations in steroid-resistant nephrotic syndrome: a mutation update and the associated phenotypic spectrum. Hum Mutat 2014;35:178-86.
159. Boute N, et al. NPHS2, encoding the glomerular protein podocin, is mutated in autosomal recessive steroid-resistant nephrotic syndrome. Nat Genet 2000;24:349-54.
160. Tory K, et al. Mutation-dependent recessive inheritance of NPHS2-associated steroid-resistant nephrotic syndrome. Nat Genet 2014;46:299-304.
161. Roselli S, et al. Plasma membrane targeting of podocin through the classical exocytic pathway: effect of NPHS2 mutations. Traffic 2004;5:37-44.
162. Reiser J, et al. TRPC6 is a glomerular slit diaphragm-associated channel required for normal renal function. Nat Genet 2005;37:739-44.
163. Moller CC, et al. Induction of TRPC6 channel in acquired forms of proteinuric kidney disease. J Am Soc Nephrol 2007;18:29-36.

164. Mottl AK, et al. A novel TRPC6 mutation in a family with podocytopathy and clinical variability. BMC Nephrol 2013;14:104.
165. Santin S, et al. TRPC6 mutational analysis in a large cohort of patients with focal segmental glomerulosclerosis. Nephrol Dial Transplant 2009;24:3089-96.
166. Gigante M, et al. TRPC6 mutations in children with steroid-resistant nephrotic syndrome and atypical phenotype. Clin J Am Soc Nephrol 2011;6:1626-34.
167. Liakopoulos V, et al. Familial collapsing focal segmental glomerulosclerosis. Clin Nephrol 2011;75:362-8.
168. Kalwa H, et al. Phospholipase C epsilon (PLCepsilon) induced TRPC6 activation: a common but redundant mechanism in primary podocytes. J Cell Physiol 2015;230:1389-99.
169. Hinkes B, et al. Positional cloning uncovers mutations in PLCE1 responsible for a nephrotic syndrome variant that may be reversible. Nat Genet 2006;38:1397-405.
170. Gilbert RD, et al. Mutations in phospholipase C epsilon 1 are not sufficient to cause diffuse mesangial sclerosis. Kidney Int 2009;75:415-9.
171. Boyer O, et al. Mutational analysis of the PLCE1 gene in steroid resistant nephrotic syndrome. J Med Genet 2010;47:445-52.
172. Shih NY, et al. CD2AP localizes to the slit diaphragm and binds to nephrin via a novel C-terminal domain. Am J Pathol 2001;159:2303-8.
173. Lehtonen S, et al. CD2-associated protein directly interacts with the actin cytoskeleton. Am J Physiol Renal Physiol 2002;283:F734-43.
174. Al-Hamed MH, et al. A molecular genetic analysis of childhood nephrotic syndrome in a cohort of Saudi Arabian families. J Hum Genet 2013;58:480-9.
175. Lowik MM, et al. Focal segmental glomerulosclerosis in a patient homozygous for a CD2AP mutation. Kidney Int 2007;72:1198-203.
176. Kim JM, et al. CD2-associated protein haploinsufficiency is linked to glomerular disease susceptibility. Science 2003;300:1298-300.
177. Gigante M, et al. CD2AP mutations are associated with sporadic nephrotic syndrome and focal segmental glomerulosclerosis (FSGS). Nephrol Dial Transplant 2009;24:1858-64.
178. Lowik M, et al. Bigenic heterozygosity and the development of steroid-resistant focal segmental glomerulosclerosis. Nephrol Dial Transplant 2008;23:3146-51.
179. Mele C, et al. MYO1E mutations and childhood familial focal segmental glomerulosclerosis. N Engl J Med 2011;365:295-306.
180. Akilesh S, et al. Arhgap24 inactivates Rac1 in mouse podocytes, and a mutant form is associated with familial focal segmental glomerulosclerosis. J Clin Invest 2011;121:4127-37.
181. Chhabra ES, et al. INF2 is an endoplasmic reticulum-associated formin protein. J Cell Sci 2009;122:1430-40.
182. Ichimura K, et al. Actin filament organization of foot processes in rat podocytes. J Histochem Cytochem 2003;51:1589-600.
183. Nabet B, et al. Identification of a putative network of actin-associated cytoskeletal proteins in glomerular podocytes defined by co-purified mRNAs. PLoS One 2009;4:e6491.
184. Garcia-Mata R, et al. The 'invisible hand': regulation of RHO GTPases by RHOGDIs. Nat Rev Mol Cell Biol 2011;12:493-504.
185. Gbadegesin RA, et al. Mutations in the gene that encodes the F-actin binding protein anillin cause FSGS. J Am Soc Nephrol 2014;25:1991-2002.
186. Brown EJ, et al. Mutations in the formin gene INF2 cause focal segmental glomerulosclerosis. Nat Genet 2010;42:72-6.
187. Boyer O, et al. Mutations in INF2 are a major cause of autosomal dominant focal segmental glomerulosclerosis. J Am Soc Nephrol 2011;22:239-45.

188. Barua M, et al. Mutations in the INF2 gene account for a significant proportion of familial but not sporadic focal and segmental glomerulosclerosis. Kidney Int 2013;83:316-22.
189. Gbadegesin RA, et al. Inverted formin 2 mutations with variable expression in patients with sporadic and hereditary focal and segmental glomerulosclerosis. Kidney Int 2012;81:94-9.
190. Lee HK, et al. Variable renal phenotype in a family with an INF2 mutation. Pediatr Nephrol 2011;26:73-6.
191. Boyer O, et al. INF2 mutations in Charcot-Marie-Tooth disease with glomerulopathy. N Engl J Med 2011;365:2377-88.
192. Sanchez-Ares M, et al. A novel mutation, outside of the candidate region for diagnosis, in the inverted formin 2 gene can cause focal segmental glomerulosclerosis. Kidney Int 2013;83:153-9.
193. Kaplan JM, et al. Mutations in ACTN4, encoding alpha-actinin-4, cause familial focal segmental glomerulosclerosis. Nat Genet 2000;24:251-6.
194. Weins A, et al. Mutational and Biological Analysis of alpha-actinin-4 in focal segmental glomerulosclerosis. J Am Soc Nephrol 2005;16:3694-701.
195. Pollak MR, et al. A case of familial kidney disease. Clin J Am Soc Nephrol 2007;2:1367-74.
196. Choi HJ, et al. Familial focal segmental glomerulosclerosis associated with an ACTN4 mutation and paternal germline mosaicism. Am J Kidney Dis 2008;51:834-8.
197. Yao J, et al. Alpha-actinin-4-mediated FSGS: an inherited kidney disease caused by an aggregated and rapidly degraded cytoskeletal protein. PLoS Biol 2004;2:e167.
198. Giglio S, et al. Heterogeneous genetic alterations in sporadic nephrotic syndrome associate with resistance to immunosuppression. J Am Soc Nephrol 2015;26:230-6.
199. Michaud JL, et al. Focal and segmental glomerulosclerosis in mice with podocyte-specific expression of mutant alpha-actinin-4. J Am Soc Nephrol 2003;14:1200-11.
200. Michaud JL, et al. FSGS-associated alpha-actinin-4 (K256E) impairs cytoskeletal dynamics in podocytes. Kidney Int 2006;70:1054-61.
201. Gupta IR, et al. ARHGDIA: a novel gene implicated in nephrotic syndrome. J Med Genet 2013;50:330-8.
202. Gee HY, et al. ARHGDIA mutations cause nephrotic syndrome via defective RHO GTPase signaling. J Clin Invest 2013;123:3243-53.
203. Auguste D, et al. Disease-causing mutations of RhoGDIalpha induce Rac1 hyperactivation in podocytes. Small GTPases 2016;7:107-21.
204. Zenker M, et al. Congenital nephrosis, mesangial sclerosis, and distinct eye abnormalities with microcoria: an autosomal recessive syndrome. Am J Med Genet A 2004;130A:138-45.
205. Matejas V, et al. A syndrome comprising childhood-onset glomerular kidney disease and ocular abnormalities with progressive loss of vision is caused by mutated LAMB2. Nephrol Dial Transplant 2006;21:3283-6.
206. Hasselbacher K, et al. Recessive missense mutations in LAMB2 expand the clinical spectrum of LAMB2-associated disorders. Kidney Int 2006;70:1008-12.
207. Choi HJ, et al. Variable phenotype of Pierson syndrome. Pediatr Nephrol 2008;23:995-1000.
208. Kagan M, et al. A milder variant of Pierson syndrome. Pediatr Nephrol 2008;23:323-7.
209. Stokman MF, et al. The expanding phenotypic spectra of kidney diseases: insights from genetic studies. Nat Rev Nephrol 2016;12:472-83.
210. Gast C, et al. Collagen (COL4A) mutations are the most frequent mutations underlying adult focal segmental glomerulosclerosis. Nephrol Dial Transplant

2016;31:961-70.
211. Gee HY, et al. Mutations in EMP2 cause childhood-onset nephrotic syndrome. Am J Hum Genet 2014;94:884-90.
212. Shukrun R, et al. A human integrin-alpha3 mutation confers major renal developmental defects. PLoS One 2014;9:e90879.
213. Lopez LC, et al. Leigh syndrome with nephropathy and CoQ10 deficiency due to decaprenyl diphosphate synthase subunit 2 (PDSS2) mutations. Am J Hum Genet 2006;79:1125-9.
214. Ashraf S, et al. ADCK4 mutations promote steroid-resistant nephrotic syndrome through CoQ10 biosynthesis disruption. J Clin Invest 2013;123:5179-89.
215. Diomedi-Camassei F, et al. COQ2 nephropathy: a newly described inherited mitochondriopathy with primary renal involvement. J Am Soc Nephrol 2007;18:2773-80.
216. Heeringa SF, et al. COQ6 mutations in human patients produce nephrotic syndrome with sensorineural deafness. J Clin Invest 2011;121:2013-24.
217. Jakobs BS, et al. A novel mutation in COQ2 leading to fatal infantile multisystem disease. J Neurol Sci 2013;326:24-8.
218. Korkmaz E, et al. ADCK4-Associated Glomerulopathy Causes Adolescence-Onset FSGS. J Am Soc Nephrol 2016;27:63-8.
219. Song R, Yosypiv IV. Genetics of congenital anomalies of the kidney and urinary tract. Pediatr Nephrol 2011;26:353-64.
220. Chugh SS. Transcriptional regulation of podocyte disease. Transl Res 2007;149:237-42.
221. Barua M, et al. Mutations in PAX2 associate with adult-onset FSGS. J Am Soc Nephrol 2014;25:1942-53.
222. Gessler M, et al. Homozygous deletion in Wilms tumours of a zinc-finger gene identified by chromosome jumping. Nature 1990;343:774-8.
223. Pelletier J, et al. WT1 mutations contribute to abnormal genital system development and hereditary Wilms' tumour. Nature 1991;353:431-4.
224. Klamt B, et al. Frasier syndrome is caused by defective alternative splicing of WT1 leading to an altered ratio of WT1 +/-KTS splice isoforms. Hum Mol Genet 1998;7:709-14.
225. Miller RW, et al. Association of Wilms's Tumor with Aniridia, Hemihypertrophy and Other Congenital Malformations. N Engl J Med 1964;270:922-7.
226. Jeanpierre C, et al. Identification of constitutional WT1 mutations, in patients with isolated diffuse mesangial sclerosis, and analysis of genotype/phenotype correlations by use of a computerized mutation database. Am J Hum Genet 1998;62:824-33.
227. Schumacher V, et al. Spectrum of early onset nephrotic syndrome associated with WT1 missense mutations. Kidney Int 1998;53:1594-600.
228. Ito S, et al. Nephrotic syndrome and end-stage renal disease with WT1 mutation detected at 3 years. Pediatr Nephrol 1999;13:790-1.
229. Denamur E, et al. WT1 splice-site mutations are rarely associated with primary steroid-resistant focal and segmental glomerulosclerosis. Kidney Int 2000;57:1868-72.
230. Lipska BS, et al. Genotype-phenotype associations in WT1 glomerulopathy. Kidney Int 2014;85:1169-78.
231. Sanyanusin P, et al. Mutation of the PAX2 gene in a family with optic nerve colobomas, renal anomalies and vesicoureteral reflux. Nat Genet 1995;9:358-64.
232. Iatropoulos P, et al. Discordant phenotype in monozygotic twins with renal coloboma syndrome and a PAX2 mutation. Pediatr Nephrol 2012;27:1989-93.
233. Boyer O, et al. LMX1B mutations cause hereditary FSGS without extrarenal involvement. J Am Soc Nephrol 2013;24:1216-22.

234. Edwards N, et al. A novel LMX1B mutation in a family with end-stage renal disease of 'unknown cause'. Clin Kidney J 2015;8:113-9.
235. Ebarasi L, et al. Defects of CRB2 cause steroid-resistant nephrotic syndrome. Am J Hum Genet 2015;96:153-61.
236. Slavotinek A, et al. CRB2 mutations produce a phenotype resembling congenital nephrosis, Finnish type, with cerebral ventriculomegaly and raised alpha-fetoprotein. Am J Hum Genet 2015;96:162-9.
237. Miyake N, et al. Biallelic Mutations in Nuclear Pore Complex Subunit NUP107 Cause Early-Childhood-Onset Steroid-Resistant Nephrotic Syndrome. Am J Hum Genet 2015;97:555-66.
238. Braun DA, et al. Mutations in nuclear pore genes NUP93, NUP205 and XPO5 cause steroid-resistant nephrotic syndrome. Nat Genet 2016;48:457-65.
239. Ozaltin F, et al. Disruption of PTPRO causes childhood-onset nephrotic syndrome. Am J Hum Genet 2011;89:139-47.
240. Wharram BL, et al. Altered podocyte structure in GLEPP1 (Ptpro)-deficient mice associated with hypertension and low glomerular filtration rate. J Clin Invest 2000;106:1281-90.
241. Taschner M, et al. Architecture and function of IFT complex proteins in ciliogenesis. Differentiation 2012;83:S12-22.
242. Ishikawa H, Marshall WF. Ciliogenesis: building the cell's antenna. Nat Rev Mol Cell Biol 2011;12:222-34.
243. Davis EE, et al. TTC21B contributes both causal and modifying alleles across the ciliopathy spectrum. Nat Genet 2011;43:189-96.
244. Otto EA, et al. Mutation analysis of 18 nephronophthisis associated ciliopathy disease genes using a DNA pooling and next generation sequencing strategy. J Med Genet 2011;48:105-16.
245. McInerney-Leo AM, et al. Whole exome sequencing is an efficient, sensitive and specific method for determining the genetic cause of short-rib thoracic dystrophies. Clin Genet 2015;88:550-7.
246. Tran PV, et al. THM1 negatively modulates mouse sonic hedgehog signal transduction and affects retrograde intraflagellar transport in cilia. Nat Genet 2008;40:403-10.
247. Huynh Cong E, et al. A homozygous missense mutation in the ciliary gene TTC21B causes familial FSGS. J Am Soc Nephrol 2014;25:2435-43.
248. Sethi S, Fervenza FC. Membranoproliferative glomerulonephritis--a new look at an old entity. N Engl J Med 2012;366:1119-31.
249. Sethi S, et al. Membranoproliferative glomerulonephritis and C3 glomerulopathy: resolving the confusion. Kidney Int 2012;81:434-41.
250. Sethi S, Fervenza FC. Membranoproliferative glomerulonephritis: pathogenetic heterogeneity and proposal for a new classification. Semin Nephrol 2011;31:341-8.
251. Cook HT, Pickering MC. Histopathology of MPGN and C3 glomerulopathies. Nat Rev Nephrol 2015;11:14-22.
252. Fakhouri F, et al. C3 glomerulopathy: a new classification. Nat Rev Nephrol 2010;6:494-9.
253. Pickering MC, et al. C3 glomerulopathy: consensus report. Kidney Int 2013;84:1079-89.
254. Pickering MT, JM. Membranoproliferative Glomerulonephritis. In: Prabhakar S, ed. An Update on Glomerulopathies - Clinical and Treatment Aspects: InTech; 2011
255. Medjeral-Thomas NR, et al. C3 glomerulopathy: clinicopathologic features and predictors of outcome. Clin J Am Soc Nephrol 2014;9:46-53.
256. Servais A, et al. Acquired and genetic complement abnormalities play a critical role in dense deposit disease and other C3 glomerulopathies. Kidney Int 2012;82:454-

64.
257. Appel GB, et al. Membranoproliferative glomerulonephritis type II (dense deposit disease): an update. J Am Soc Nephrol 2005;16:1392-403.
258. Fervenza FC, et al. Idiopathic membranoproliferative glomerulonephritis: does it exist? Nephrol Dial Transplant 2012;27:4288-94.
259. Goodship TH, et al. Factor H autoantibodies in membranoproliferative glomerulonephritis. Mol Immunol 2012;52:200-6.
260. Strobel S, et al. Anti-factor B autoantibody in dense deposit disease. Mol Immunol 2010;47:1476-83.
261. Imamura H, et al. Familial C3 glomerulonephritis associated with mutations in the gene for complement factor B. Nephrol Dial Transplant 2015;30:862-4.
262. Bu F, et al. High-Throughput Genetic Testing for Thrombotic Microangiopathies and C3 Glomerulopathies. J Am Soc Nephrol 2015.
263. Martinez-Barricarte R, et al. Human C3 mutation reveals a mechanism of dense deposit disease pathogenesis and provides insights into complement activation and regulation. J Clin Invest 2010;120:3702-12.
264. Abrera-Abeleda MA, et al. Allelic variants of complement genes associated with dense deposit disease. J Am Soc Nephrol 2011;22:1551-9.
265. Heurich M, et al. Common polymorphisms in C3, factor B, and factor H collaborate to determine systemic complement activity and disease risk. Proc Natl Acad Sci U S A 2011;108:8761-6.
266. Ponticelli C. Membranous nephropathy. J Nephrol 2007;20:268-87.
267. Fogo AB, et al. AJKD Atlas of Renal Pathology: Membranous Nephropathy. Am J Kidney Dis 2015;66:e15-7.
268. Ronco P, Debiec H. Pathophysiological advances in membranous nephropathy: time for a shift in patient's care. Lancet 2015;385:1983-92.
269. Beck LH, Jr., et al. M-type phospholipase A2 receptor as target antigen in idiopathic membranous nephropathy. N Engl J Med 2009;361:11-21.
270. Dai H, et al. Diagnostic accuracy of PLA2R autoantibodies and glomerular staining for the differentiation of idiopathic and secondary membranous nephropathy: an updated meta-analysis. Sci Rep 2015;5:8803.
271. Tomas NM, et al. Thrombospondin type-1 domain-containing 7A in idiopathic membranous nephropathy. N Engl J Med 2014;371:2277-87.
272. Murtas C, et al. Coexistence of different circulating anti-podocyte antibodies in membranous nephropathy. Clin J Am Soc Nephrol 2012;7:1394-400.
273. Stanescu HC, et al. Risk HLA-DQA1 and PLA(2)R1 alleles in idiopathic membranous nephropathy. N Engl J Med 2011;364:616-26.
274. Bullich G, et al. HLA-DQA1 and PLA2R1 polymorphisms and risk of idiopathic membranous nephropathy. Clin J Am Soc Nephrol 2014;9:335-43.
275. Jiang S, et al. The primary glomerulonephritides: a systems biology approach. Nat Rev Nephrol 2013;9:500-12.
276. Wyatt RJ, Julian BA. IgA nephropathy. N Engl J Med 2013;368:2402-14.
277. Cattran DC, et al. The Oxford classification of IgA nephropathy: rationale, clinicopathological correlations, and classification. Kidney Int 2009;76:534-45.
278. Kim JK, et al. Clinical features and outcomes of IgA nephropathy with nephrotic syndrome. Clin J Am Soc Nephrol 2012;7:427-36.
279. Kiryluk K, et al. Discovery of new risk loci for IgA nephropathy implicates genes involved in immunity against intestinal pathogens. Nat Genet 2014;46:1187-96.
280. Kiryluk K, Novak J. The genetics and immunobiology of IgA nephropathy. J Clin Invest 2014;124:2325-32.
281. Silva FG, et al. Disappearance of glomerular mesangial IgA deposits after renal allograft transplantation. Transplantation 1982;33:241-6.

282. Kiryluk K, et al. Pathogenesis of immunoglobulin A nephropathy: recent insight from genetic studies. Annu Rev Med 2013;64:339-56.
283. Fabiano RC, et al. Immunoglobulin A nephropathy: a pathophysiology view. Inflamm Res 2016;65:757-70.
284. Stanescu HK, A; Kleta, R. Risk alleles in idiopathic membranous nephropathy; Author reply. New England Journal of Medicine 2011;364:2073-4.
285. Leslie S, et al. A statistical method for predicting classical HLA alleles from SNP data. Am J Hum Genet 2008;82:48-56.
286. Coppo R, et al. Upregulation of the immunoproteasome in peripheral blood mononuclear cells of patients with IgA nephropathy. Kidney Int 2009;75:536-41.
287. McCarthy DD, et al. Mice overexpressing BAFF develop a commensal flora-dependent, IgA-associated nephropathy. J Clin Invest 2011;121:3991-4002.
288. Castigli E, et al. TACI is mutant in common variable immunodeficiency and IgA deficiency. Nat Genet 2005;37:829-34.
289. Hadjicharalambous C, et al. Mechanisms of alpha-defensin bactericidal action: comparative membrane disruption by Cryptdin-4 and its disulfide-null analogue. Biochemistry 2008;47:12626-34.
290. Kunisawa J, et al. Microbe-dependent CD11b+ IgA+ plasma cells mediate robust early-phase intestinal IgA responses in mice. Nat Commun 2013;4:1772.
291. Imielinski M, et al. Common variants at five new loci associated with early-onset inflammatory bowel disease. Nat Genet 2009;41:1335-40.
292. Hughes AE, et al. A common CFH haplotype, with deletion of CFHR1 and CFHR3, is associated with lower risk of age-related macular degeneration. Nat Genet 2006;38:1173-7.
293. Manenti L, et al. Atypical haemolytic uraemic syndrome with underlying glomerulopathies. A case series and a review of the literature. Nephrol Dial Transplant 2013;28:2246-59.
294. Noris M, et al. Relative role of genetic complement abnormalities in sporadic and familial aHUS and their impact on clinical phenotype. Clin J Am Soc Nephrol 2010;5:1844-59.
295. Bircan Z, et al. Factor H deficiency and fibrillary glomerulopathy. Nephrol Dial Transplant 2004;19:727-30.
296. Sethi S, et al. C4d as a diagnostic tool in proliferative GN. Journal of the American Society of Nephrology 2015:ASN. 2014040406.
297. Yabuuchi J, et al. Sjögren Syndrome-Related Membranous Glomerulonephritis Progressing to Membranoproliferative Glomerulonephritis. Case Reports in Nephrology and Dialysis 2016;6:133-42.
298. Fremeaux-Bacchi V, et al. Selective disappearance of C3NeF IgG autoantibody in the plasma of a patient with membranoproliferative glomerulonephritis following renal transplantation. Nephrol Dial Transplant 1994;9:811-4.
299. Kircher M, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014;46:310-5.
300. Hou J, et al. Toward a working definition of C3 glomerulopathy by immunofluorescence. Kidney Int 2014;85:450-6.
301. Maga TK, et al. Mutations in alternative pathway complement proteins in American patients with atypical hemolytic uremic syndrome. Hum Mutat 2010;31:E1445-60.
302. Perez-Caballero D, et al. Clustering of missense mutations in the C-terminal region of factor H in atypical hemolytic uremic syndrome. Am J Hum Genet 2001;68:478-84.
303. Zhang Y, et al. Causes of alternative pathway dysregulation in dense deposit disease. Clin J Am Soc Nephrol 2012;7:265-74.

304. Kavanagh D, et al. Atypical hemolytic uremic syndrome. Semin Nephrol 2013;33:508-30.
305. Dragon-Durey MA, et al. Heterozygous and homozygous factor h deficiencies associated with hemolytic uremic syndrome or membranoproliferative glomerulonephritis: report and genetic analysis of 16 cases. J Am Soc Nephrol 2004;15:787-95.
306. Noris M, et al. Dynamics of complement activation in aHUS and how to monitor eculizumab therapy. Blood 2014;124:1715-26.
307. Caprioli J, et al. Complement factor H mutations and gene polymorphisms in haemolytic uraemic syndrome: the C-257T, the A2089G and the G2881T polymorphisms are strongly associated with the disease. Hum Mol Genet 2003;12:3385-95.
308. Pechtl IC, et al. Disease-associated N-terminal complement factor H mutations perturb cofactor and decay-accelerating activities. Journal of Biological Chemistry 2011;286:11082-90.
309. Delvaeye M, et al. Thrombomodulin mutations in atypical hemolytic-uremic syndrome. N Engl J Med 2009;361:345-57.
310. Caprioli J, et al. Genetics of HUS: the impact of MCP, CFH, and IF mutations on clinical presentation, response to treatment, and outcome. Blood 2006;108:1267-79.
311. Sethi S, et al. C3 glomerulonephritis associated with complement factor B mutation. Am J Kidney Dis 2015;65:520-1.
312. Rooijakkers SH, et al. Structural and functional implications of the alternative complement pathway C3 convertase stabilized by a staphylococcal inhibitor. Nat Immunol 2009;10:721-7.
313. Schramm EC, et al. Functional mapping of the interactions between complement C3 and regulatory proteins using atypical hemolytic uremic syndrome-associated mutations. Blood 2015.
314. Wu J, et al. Structure of complement fragment C3b-factor H and implications for host protection by complement regulators. Nat Immunol 2009;10:728-33.
315. Forneris F, et al. Structures of C3b in complex with factors B and D give insight into complement convertase formation. Science 2010;330:1816-20.
316. Nilsson SC, et al. Analysis of binding sites on complement factor I that are required for its activity. J Biol Chem 2010;285:6235-45.
317. Iwata K, et al. Diversity of sites for measles virus binding and for inactivation of complement C3b and C4b on membrane cofactor protein CD46. J Biol Chem 1995;270:15148-52.
318. Daina E, et al. Eculizumab in a patient with dense-deposit disease. N Engl J Med 2012;366:1161-3.
319. Mizutani M, et al. Glomerular localization of thrombomodulin in human glomerulonephritis. Lab Invest 1993;69:193-202.
320. Niemir ZI, et al. Can von Willebrand factor, platelet-endothelial cell adhesion molecule-1 and thrombomodulin be used as alternative markers of endothelial cell injury in human glomerulonephritis? Rocz Akad Med Bialymst 2004;49:213-8.
321. Van de Wouwer M, et al. Thrombomodulin-protein C-EPCR system: integrated to regulate coagulation and inflammation. Arterioscler Thromb Vasc Biol 2004;24:1374-83.
322. Tortajada A, et al. The disease-protective complement factor H allotypic variant Ile62 shows increased binding affinity for C3b and enhanced cofactor activity. Hum Mol Genet 2009;18:3452-61.
323. Pickering MC, et al. Spontaneous hemolytic uremic syndrome triggered by complement factor H lacking surface recognition domains. J Exp Med 2007;204:1249-56.

324. Montes T, et al. Functional basis of protection against age-related macular degeneration conferred by a common polymorphism in complement factor B. Proceedings of the National Academy of Sciences 2009;106:4366-71.
325. Kavanagh D, Goodship TH. Atypical hemolytic uremic syndrome, genetic basis, and clinical manifestations. ASH Education Program Book 2011;2011:15-20.
326. Paixao-Cavalcante D, et al. Sensitive and specific assays for C3 nephritic factors clarify mechanisms underlying complement dysregulation. Kidney Int 2012;82:1084-92.
327. Licht C, et al. MPGN II–genetically determined by defective complement regulation? Pediatric Nephrology 2007;22:2-9.
328. Little MA, et al. Severity of primary MPGN, rather than MPGN type, determines renal survival and post-transplantation recurrence risk. Kidney Int 2006;69:504-11.
329. Bomback AS, Appel GB. Pathogenesis of the C3 glomerulopathies and reclassification of MPGN. Nat Rev Nephrol 2012;8:634-42.
330. Jozsi M, et al. Factor H-related proteins determine complement-activating surfaces. Trends Immunol 2015;36:374-84.
331. Sprague S, et al. Multicenter collaboration in observational research: improving generalizability and efficiency. J Bone Joint Surg Am 2009;91 Suppl 3:80-6.
332. Strom CM, et al. Development and Validation of a Next-Generation Sequencing Assay for BRCA1 and BRCA2 Variants for the Clinical Laboratory. PLoS One 2015;10:e0136419.
333. Figueres ML, et al. Heterogeneous histologic and clinical evolution in 3 cases of dense deposit disease with long-term follow-up. Hum Pathol 2014;45:2326-33.
334. Habbig S, et al. C3 deposition glomerulopathy due to a functional factor H defect. Kidney Int 2009;75:1230-4.
335. Licht C, et al. Deletion of Lys224 in regulatory domain 4 of Factor H reveals a novel pathomechanism for dense deposit disease (MPGN II). Kidney Int 2006;70:42-50.
336. van Rooden SM, et al. The identification of Parkinson's disease subtypes using cluster analysis: a systematic review. Mov Disord 2010;25:969-78.
337. Burgel PR, et al. Clinical COPD phenotypes: a novel approach using principal component and cluster analyses. Eur Respir J 2010;36:531-9.
338. Howrylak JA, et al. Classification of childhood asthma phenotypes and long-term clinical responses to inhaled anti-inflammatory medications. J Allergy Clin Immunol 2014;133:1289-300, 300 e1-12.
339. Iatropoulos P, et al. Complement gene variants determine the risk of immunoglobulin-associated MPGN and C3 glomerulopathy and predict long-term renal outcome. Mol Immunol 2016;71:131-42.
340. Kruizenga HM, et al. Development and validation of a hospital screening tool for malnutrition: the short nutritional assessment questionnaire (SNAQ). Clin Nutr 2005;24:75-82.
341. Cook HT. C4d Staining in the Diagnosis of C3 Glomerulopathy. J Am Soc Nephrol 2015;26:2609-11.
342. Sethi S, et al. C3 glomerulonephritis: clinicopathological findings, complement abnormalities, glomerular proteomic profile, treatment, and follow-up. Kidney Int 2012;82:465-73.
343. Zhang Y, et al. Defining the complement biomarker profile of C3 glomerulopathy. Clin J Am Soc Nephrol 2014;9:1876-82.
344. Bomback AS, et al. Eculizumab for dense deposit disease and C3 glomerulonephritis. Clin J Am Soc Nephrol 2012;7:748-56.
345. Mekahli D, et al. Long-term outcome of idiopathic steroid-resistant nephrotic syndrome: a multicenter study. Pediatr Nephrol 2009;24:1525-32.
346. Rheault MN, Gbadegesin RA. The genetics of nephrotic syndrome. Journal of

Pediatric Genetics 2016;5:015-24.
347. Mele C, et al. Characterization of a New DGKE Intronic Mutation in Genetically Unsolved Cases of Familial Atypical Hemolytic Uremic Syndrome. Clin J Am Soc Nephrol 2015;10:1011-9.
348. Ozaltin F, et al. DGKE variants cause a glomerular microangiopathy that mimics membranoproliferative GN. J Am Soc Nephrol 2013;24:377-84.
349. Zhu L, et al. Variants in Complement Factor H and Complement Factor H-Related Protein Genes, CFHR3 and CFHR1, Affect Complement Activation in IgA Nephropathy. J Am Soc Nephrol 2015;26:1195-204.
350. Rugarli C. Medicina interna. MASSON 1997.
351. Schachter ME, et al. Recurrent focal segmental glomerulosclerosis in the renal allograft: single center experience in the era of modern immunosuppression. Clin Nephrol 2010;74:173-81.
352. Gbadegesin R, et al. A new locus for familial FSGS on chromosome 2p. J Am Soc Nephrol 2010;21:1390-7.
353. Sanders SJ, et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature 2012;485:237-41.
354. Loman NJ, et al. Performance comparison of benchtop high-throughput sequencing platforms. Nat Biotechnol 2012;30:434-9.
355. Junemann S, et al. Updating benchtop sequencing performance comparison. Nat Biotechnol 2013;31:294-6.
356. Quail MA, et al. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics 2012;13:341.
357. Krupp M, et al. RNA-Seq Atlas--a reference database for gene expression profiling in normal tissue by next-generation sequencing. Bioinformatics 2012;28:1184-5.
358. Sampson MG, et al. Gene-level integrated metric of negative selection (GIMS) prioritizes candidate genes for nephrotic syndrome. PLoS One 2013;8:e81062.
359. Petrovski S, et al. Genic intolerance to functional variation and the interpretation of personal genomes. PLoS Genet 2013;9:e1003709.
360. Ju W, et al. Defining cell-type specificity at the transcriptional level in human disease. Genome Res 2013;23:1862-73.
361. Broman KW, Weber JL. Long homozygous chromosomal segments in reference families from the centre d'Etude du polymorphisme humain. Am J Hum Genet 1999;65:1493-500.
362. Lenkkeri U, et al. Structure of the gene for congenital nephrotic syndrome of the finnish type (NPHS1) and characterization of mutations. Am J Hum Genet 1999;64:51-61.
363. Heidet L, et al. Spectrum of HNF1B mutations in a large cohort of patients who harbor renal diseases. Clin J Am Soc Nephrol 2010;5:1079-90.
364. Kramer-Zucker AG, et al. Organization of the pronephric filtration apparatus in zebrafish requires Nephrin, Podocin and the FERM domain protein Mosaic eyes. Dev Biol 2005;285:316-29.
365. Hwang DY, et al. Mutations in 12 known dominant disease-causing genes clarify many congenital anomalies of the kidney and urinary tract. Kidney Int 2014;85:1429-33.
366. Gosens I, et al. FERM protein EPB41L5 is a novel member of the mammalian CRB-MPP5 polarity complex. Exp Cell Res 2007;313:3959-70.
367. Lamont RE, et al. Expansion of phenotype and genotypic data in CRB2-related syndrome. Eur J Hum Genet 2016;24:1436-44.
368. Laprise P, et al. The FERM protein Yurt is a negative regulatory component of

the Crumbs complex that controls epithelial polarity and apical membrane size. Developmental cell 2006;11:363-74.
369. Gross O, et al. DDR1-deficient mice show localized subepithelial GBM thickening with focal loss of slit diaphragms and proteinuria. Kidney Int 2004;66:102-11.
370. Huang Y, et al. The collagen receptor DDR1 regulates cell spreading and motility by associating with myosin IIA. J Cell Sci 2009;122:1637-46.
371. Meyer Zum Gottesberge AM, Hansen S. The collagen receptor DDR1 co-localizes with the non-muscle myosin IIA in mice inner ear and contributes to the cytoarchitecture and stability of motile cells. Cell Tissue Res 2014;358:729-36.
372. Durtschi J, et al. VarBin, a novel method for classifying true and false positive variants in NGS data. BMC Bioinformatics 2013;14 Suppl 13:S2.
373. Chan M, et al. Development of a next-generation sequencing method for BRCA mutation screening: a comparison between a high-throughput and a benchtop platform. J Mol Diagn 2012;14:602-12.
374. Chin EL, et al. Assessment of clinical analytical sensitivity and specificity of next-generation sequencing for detection of simple and complex mutations. BMC Genet 2013;14:6.
375. Yeo ZX, et al. Improving indel detection specificity of the Ion Torrent PGM benchtop sequencer. PLoS One 2012;7:e45798.
376. Laehnemann D, et al. Denoising DNA deep sequencing data-high-throughput sequencing errors and their correction. Brief Bioinform 2016;17:154-79.
377. Damiati E, et al. Amplicon-based semiconductor sequencing of human exomes: performance evaluation and optimization strategies. Hum Genet 2016;135:499-511.
378. Ledergerber C, Dessimoz C. Base-calling for next-generation sequencing platforms. Brief Bioinform 2011;12:489-97.
379. Coonrod EM, et al. Developing genome and exome sequencing for candidate gene identification in inherited disorders: an integrated technical and bioinformatics approach. Arch Pathol Lab Med 2013;137:415-33.
380. DePristo MA, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011;43:491-8.
381. Glenn TC. Field guide to next-generation DNA sequencers. Mol Ecol Resour 2011;11:759-69.
382. Fox EJ, et al. Accuracy of Next Generation Sequencing Platforms. Next Gener Seq Appl 2014;1.
383. Zavodna M, et al. The accuracy, feasibility and challenges of sequencing short tandem repeats using next-generation sequencing platforms. PLoS One 2014;9:e113862.
384. Abnizova I, et al. Analysis of context-dependent errors for illumina sequencing. J Bioinform Comput Biol 2012;10:1241005.
385. Meacham F, et al. Identification and correction of systematic error in high-throughput sequence data. BMC Bioinformatics 2011;12:451.
386. Benjamini Y, Speed TP. Summarizing and correcting the GC content bias in high-throughput sequencing. Nucleic Acids Res 2012;40:e72.
387. Treangen TJ, Salzberg SL. Repetitive DNA and next-generation sequencing: computational challenges and solutions. Nat Rev Genet 2012;13:36-46.
388. Bansal V, et al. Accurate detection and genotyping of SNPs utilizing population sequencing data. Genome Res 2010;20:537-45.
389. Muralidharan O, et al. A cross-sample statistical model for SNP detection in short-read sequencing data. Nucleic Acids Res 2012;40:e5.
390. Margraf RL, et al. Multi-sample pooling and illumina genome analyzer sequencing methods to determine gene sequence variation for database development. J Biomol Tech 2010;21:126-40.

391. Chilamakuri CS, et al. Performance comparison of four exome capture systems for deep sequencing. BMC Genomics 2014;15:449.
392. Head SR, et al. Library construction for next-generation sequencing: overviews and challenges. Biotechniques 2014;56:61-4, 6, 8, passim.
393. van Dijk EL, et al. Library preparation methods for next-generation sequencing: tone down the bias. Exp Cell Res 2014;322:12-20.
394. Hatem A, et al. Benchmarking short sequence mapping tools. BMC Bioinformatics 2013;14:184.
395. Ross MG, et al. Characterizing and measuring bias in sequence data. Genome Biol 2013;14:R51.
396. Pabinger S, et al. A survey of tools for variant analysis of next-generation genome sequencing data. Brief Bioinform 2014;15:256-78.
397. Ghoneim DH, et al. Comparison of insertion/deletion calling algorithms on human next-generation sequencing data. BMC Res Notes 2014;7:864.
398. Yi M, et al. Performance comparison of SNP detection tools with illumina exome sequencing data--an assessment using both family pedigree information and sample-matched SNP array data. Nucleic Acids Res 2014;42:e101.
399. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 2011;27:2987-93.
400. Li H. Improving SNP discovery by base alignment quality. Bioinformatics 2011;27:1157-8.
401. Simola DF, Kim J. Sniper: improved SNP discovery by multiply mapping deep sequenced reads. Genome Biol 2011;12:R55.
402. Shen Y, et al. A SNP discovery method to assess variant allele probability from next-generation resequencing data. Genome Res 2010;20:273-80.
403. Yeo ZX, et al. Evaluation and optimisation of indel detection workflows for ion torrent sequencing of the BRCA1 and BRCA2 genes. BMC Genomics 2014;15:516.
404. Boland JF, et al. The new sequencer on the block: comparison of Life Technology's Proton sequencer to an Illumina HiSeq for whole-exome sequencing. Hum Genet 2013;132:1153-63.
405. Ion Hi-Q View Chemistry for Ion Torrent Next-Generation Sequencing.
406. Costa JL, et al. Nonoptical massive parallel DNA sequencing of BRCA1 and BRCA2 genes in a diagnostic setting. Hum Mutat 2013;34:629-35.
407. Kim SY, et al. Estimation of allele frequency and association mapping using next-generation sequencing data. BMC Bioinformatics 2011;12:231.
408. Ferlini AG, M; Genuardi, M; Grammatico, P; Seri, M; Torricelli, F; Zollino, M; Zuffardi, O. Il sequenziamento del DNA di nuova generazione: indicazioni per l'impiego clinico. Documenti SIGU, Linee Guida e Raccomandazioni 2016.
409. Warden CD, et al. Detailed comparison of two popular variant calling packages for exome and targeted exon studies. PeerJ 2014;2:e600.
410. Harita Y, et al. Spectrum of LMX1B mutations: from nail-patella syndrome to isolated nephropathy. Pediatr Nephrol 2016.

MATERIAL PUBLISHED CONTAINING WORK DESCRIBED IN THE THESIS

Iatropoulos P, Noris M, Mele C, Piras R, Valoti E, Bresin E, Curreri M, Mondo E, Zito A, Gamba S, Bettoni S, Murer L, Fremeaux-Bacchi V, Vivarelli M, Emma F, Daina E, Remuzzi G. *"Complement gene variants determine the risk of immunoglobulin-associated MPGN and C3 glomerulopathy and predict long-term renal outcome."* Mol Immunol. 2016;71:131-42.